



Evaluating the performance of demographic targeting using gender in sponsored search

Bernard J. Jansen*, Kathleen Moore, Stephen Carman

College of Information Sciences and Technology, The Pennsylvania State University, United States

ARTICLE INFO

Article history:

Received 8 December 2011
 Received in revised form 12 May 2012
 Accepted 25 June 2012
 Available online 17 July 2012

Keywords:

Sponsored search
 Keyword advertising
 Pay-per-click
 PPC
 Online advertising
 Search engine marketing
 Gender targeting
 Demographic profiling

ABSTRACT

In this research, we evaluate the effect of gender targeted advertising on the performance of sponsored search advertising. We analyze nearly 7,000,000 records spanning 33 consecutive months of a keyword advertising campaign from a major US retailer. In order to determine the effect of demographic targeting, we classify the campaign's key phrases by a probability of being targeted for a specific gender, and we then compare the key performance indicators among these groupings using the critical sponsored search metrics of impressions, clicks, cost-per-click, sales revenue, orders, and items, and return on advertising. Findings from our research show that the gender-orientation of the key phrase is a significant determinant in predicting behaviors and performance, with statistically different consumer behaviors for all attributes as the probability of a male or female keyword phrase changes. However, gender neutral phrases perform the best overall, generating 20 times the return of advertising than any gender targeted category. Insight from this research could result in sponsored advertising efforts being more effectively targeted to searchers and potential consumers.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Online shopping is a growing phenomenon with both off-line retailers expanding their business opportunities to the digital marketplace and the number of online-only businesses increasing. As the internet-based marketplace surges, online retailers are more interested in finding ways to leverage the Web to enhance their business models, including incorporating keyword advertising. The challenge that online retailers face is not only getting visitors to their site but also having these potential customers convert (e.g., to make a purchase). By understanding the search behavior of online consumers, these retailers can find opportunities to optimize keyword advertising strategies.

This effort increasingly requires the understanding of Web demographics, with many practitioners interested and willing to pay more for demographically target advertisements (SEMPO, 2011). Based on the theory of social categories (Deaux, 1984) that postulates that *like people act alike*, demographically targeted advertising has attracted the interest of online marketers in a variety of areas. Gender has emerged as one of the defining demographics of focus. Gender is a recurrent determinant in devising marketing and advertising strategies, with electronic commerce research indicating that gender is a key attribute and predictor of intent to purchase. Research findings indicate that females might have negative views of online shopping (Rodgers & Harris, 2003), although subsequent survey data suggests females are as likely as males to purchase online (Horrigan, 2008).

* Corresponding author. Tel.: +1 814 865 6459.

E-mail addresses: jjansen@acm.org (B.J. Jansen), kmoore@ist.psu.edu (K. Moore), shcarman@gmail.com (S. Carman).

There has been very limited published research concerning the effect of gender demographic targeting in evaluating the successes or failures of sponsored search campaigns. This is because companies generally have not published their sponsored search data, and most web analytics methods used in practice do not incorporate gender attributes. Consequently, we have incomplete insight into how the gender demographics of searchers affect their interaction with this increasingly important segment of Web search engine results and ecommerce. So, although online marketers have indicated interest in demographic targeting (SEMPO, 2011), we have no evidence on whether or not this is a fruitful advertising endeavor within the domain of keyword advertising.

In this research, we use actual keyword advertising data from a major US retailer to examine searches from a gender-orientation and their corresponding online consumer behavior. The research goal is to determine whether gender is a factor that online retailers should consider when implementing their advertising strategies, with direct implications for ecommerce personalization and web searching.

For readers who may not be familiar with keyword advertising campaigns on major web search engines, there are several important sponsored search terms that are frequently used in the industry that one must have a working knowledge of in order to understand the research presented in this paper. When a sponsored result is displayed on a search engine results page (SERP) in response to a query that matches a given key phrase, this is called an *impression*. When a searcher clicks on the hyperlink pointing to an advertiser's landing page, this is a *click*. The search engine bills the advertiser for this click, an amount known as the *cost-per-click* (CPC). Once at the landing page, if the consumer makes a purchase (a.k.a., places an *order*), this act is known as a *conversion*. An order can be composed of one or more *items*. The *sales revenue* generated from this conversion defines the value of the customer. The effectiveness of the keyword advertising effort is measured by revenue generated less the advertising cost of that visitor. This is a brief overview of a very complex process. For further discussions of the keyword advertising, one can see (Fain & Pedersen, 2006; Jansen & Mullen, 2008) that are part of a small but growing body of literature on keyword advertising.

We begin with a literature review, outlining the prior work concerning the effect of gender on searching and online purchasing. We then present our research questions and associated hypotheses, with justifications. We present a description of our data and methods of analysis. We then discuss results and implications for advertisers, online advertising platforms, and consumers. We end with directions for future research.

2. Literature review

Human information processing is the theoretical basis for this research, which is the method of acquiring, interpreting, manipulating, storing, retrieving, and classifying recorded information (Putrevu, 2002; Wilson, 2000). A number of personal factors influence information processing in an individual, including, economic class, culture, education, and knowledge.

2.1. The role gender in information processing

Gender is also one of these factors, which is the attribute of interest in this research. Gender has been shown to have an influence on information searching tactics (Lorigo et al., 2006). Prior research has identified important gender differences in information processing strategies. Table 1 depicts a summary of prior work examining information processing between genders.

While certainly a generalization, prior research in information processing shows that men and women differ in the way they internalize information stimulation and respond, thus impacting judgment of said information. The summary from Table 1 of research into these differences show women generally consider all information whereas men are more selective

Table 1
Major findings concerning gender influences from information processing research.

Information processing	Male	Female
Carsky and Zuckerman (1991)	Selective information processors	Comprehensive information processors
Meyers-Levy and Maheswaran (1991)	Utilize schema-based processing strategy	
Meyers-Levy and Sternthal (1991)		Utilize detailed processing strategy
Kirouac and Dore (1983)		Better able to process non-verbal cues
Prakash and Flores (1985)		Process information more subjectively
Meyers-Levy	Persuaded by information choices, value self-generated information	Persuaded by communal sentiment, value both self and other-generated information
Meyers-Levy	Process information selectively	Process information elaborately
Darley and Smith	Do not prefer objective claims	At low risk, consider both objective and subjective claims but favor subjective; at moderate risk, favor objective claims

(Carsky & Zuckerman, 1991), considering the most conspicuous pieces of information and ignoring the rest. This processing difference makes women more integrative, seeking more data in their decision making while men prefer a much less complicated approach (Meyers-Levy & Maheswaran, 1991; Meyers-Levy & Sternthal, 1991). Kirouac and Dore (1983) show that women are better able to process non-verbal information cues, and Prakash and Flores (1985) show that women are more subjective in their information processing. Meyers-Levy (1988) state that men are persuaded if there are information choices and value self-generated information. Meyers-Levy (1988) also notes that women are persuaded by community sentiment and value a range of information sources. Meyers-Levy (1989) also claim that men are selective information processors, while women are more complex in their information processing. Darley and Smith (1995) state that men are persuaded by objective claims, while women's place value in both objective and subjective claims depending on the context. Since children are socialized along gender lines, it is likely that the differences in socialization reinforce the differences in information processes from childhood into adulthood (Carpenter & Huston-Stein, 1980). Given these differences in information processing, one would expect differences in online searching behaviors.

Studies in online information seeking behavior, particularly web searching, have shown differences between men and women in line with accepted information processing theories. In a study of school children, boys used fewer words in query building relative to girls, and boys tended to scan pages quickly (i.e., selective processing) and leave result pages to follow other search result links. The girls preferred to read more thoroughly through the page of results (i.e., comprehensive processing) (Large, Beheshti, & Rahman, 2002). A study of college students also showed that women spent more time on the Internet for communication purposes and spent longer than men searching for information (Roy & Chi, 2003). Additionally, women reported more anxiety regarding Internet use than men (Jackson, Ervin, Gardner, & Schmitt, 2001). Lorigo et al. (2006) noted that females fixated less on individual web documents and submitted longer queries than males. With keyword advertising tied closely to web searching, gender may be a key demographic attribute for sponsored search efforts.

2.2. The role gender in advertising

Gender has certainly been a critical factor in market segmentation strategy for advertisers (Wolin & Korgaonkar, 2003). Given the different ways men and women process information, researchers have recommend that retailers employ specific strategies to appeal to the differences between genders (Darley & Smith, 1995; McMahan, Hovland, & McMillan, 2009; Meyers-Levy, 1989; Wolin & Korgaonkar, 2003). Results from studies in traditional advertising mediums (e.g., television, radio, print) have been consistent with expectations from information processing theories; where men preferred simple ads with straightforward comparative appeals, while women responded better to more verbal, complex, and informative ads (Putrevu, 2002). Men also respond more favorably to ads that reference self where as women responded better to ads referencing self and others (Meyers-Levy, 1988). Both genders responded favorably towards ads that corresponded with respective gender-roles (Morrison & Shaffer, 2003; Putrevu, 2002); however, in a study of participants self-referencing as either traditional (gender-stereotyped) or non-traditional (androgynous, feminine-men, masculine-women), when encouraged to self-reference, both traditional and non-traditional participants were more responsive to non-traditional or gender-neutral advertisements (Morrison & Shaffer, 2003). Again, these findings were based on traditional advertising mediums (i.e., off-line).

Verifying these same dispositions towards advertising in an online environment can be difficult because the highly dynamic online environment make it not just an advertising medium but a customer communications forum and channel of distribution (Wolin & Korgaonkar, 2003). The capacity of multi-media messages in the web, both solicited and unsolicited by consumers, allows retailers broad reach in their advertising abilities. However, there have been some studies that examine the information processing similarities and differences between males and females, as presented in Table 2.

The studies that have examined gender differences in response to online advertising on the web, shown in Table 2, show marked distinctions between men and women in their response to advertising information. Bezjian-Avery, Alexa, and Iacobucci (1998) note that men prefer interactive online ads. Overall, men have more positive reactions towards online ads than women (Wolin & Korgaonkar, 2003), and attitudes reflect gender-respective processing traits in that men will quickly filter information to make decisions while women will regard information more comprehensively. Kempf, Laczniak, and Smith

Table 2
Major findings concerning gender influences from online advertising research.

Online advertising response	Male	Female
Bezjian-Avery et al. (1998) (Wolin & Korgaonkar, 2003)	Prefer interactivity of web ads Have more positive beliefs and attitudes about web advertising	Have more negative beliefs and attitudes about web advertising
Wolin and Korgaonkar (2003)	Based on medium, more likely to make web purchase, more likely to choose function and entertainment sites	Based on medium, more likely to choose shopping site but likely for information gathering
Kempf et al. (2006)	Used more readily available judgment, less likely to notice incomplete information	More sensitive to comprehensive product information and manipulation of descriptive information
Bae and Lee (2011)	Less affected by online reviews	More affected by online reviews

(2006) note that men rely more on judgment and might not notice incomplete information while women are sensitive to missing or manipulation of information. Bae and Lee (2011) state that men are less affected by online review than women. Most important to note, however, is how men are more likely make a web purchase based on a favorable online ad while women are more likely to only visit the shopping site of that ad for information gathering (Wolin & Korgaonkar, 2003), although this is based on the particular medium.

2.3. The role gender in online shopping

Similar results to that in online advertising are also found in gender-differentiated behavior during online shopping, as shown in Table 3.

Studies have shown mixed and inconsistent results in the differences between genders when it comes to online shopping (Chang, Cheung, & Lai, 2005; Zhou, Dai, & Zhang, 2007) or no differences at all (Ulbrich, Christenson, & Stankhus, 2011). This could be explained that as the Internet became more integrated into people's daily lives, gender differences in online activity have narrowed. Women may be ambivalent about online shopping, but they are spending more dollars online and their participation in shopping activities is increasing (Abraham, Morn, & Vollman, 2010). van Slyke, Comunale, and Belanger (2002) claim that men have more preferred products online relative to women. Garbarino and Strahilevitze (2004) state that women have a higher perceived risk of shopping online. Dittmar, Long, and Meek (2004) write that men view online shopping the same as off-line shopping, which women view online shopping as less enjoyable than off-line shopping. Sanchez-Franco (2006) states that attitude has a strong role for women when shopping. McMahan et al. (2009) write that men prefer interactive attributes and women prefer more state online features. Hasan (2010) claims that men are more engaged online than are women.

As with online ads, men prefer the efficiency and convenience of online shopping and are shown to have more trust and satisfaction in the online experience than women. Given women's propensity as data collectors, the lack of dimension in online shopping (that is, the lack interaction with the product where they can effectively evaluate it) makes online shopping a less enjoyable activity (Cho, 2004; Citrin, Stem, Spangenberg, & Clark, 2003; Cyr & Bonanni, 2005; Dittmar et al., 2004; Rodgers & Harris, 2003). Aside from the act of shopping, however, there are also differences between genders regarding interactivity with an online shopping site. Men are more likely to explore a website's interactive features, like videos and customized views of products, while women are more likely to explore a website's communication features and promotional information, which is consistent with the more general Internet activities performed by different genders (Jackson et al., 2001). Yeh, Hsiao, and Yang (2012) noted differences in online shopping behavior by gender.

Additionally, research into online shopping shows differences in cognitive, affective and behavioral attitudes by gender (Hasan, 2010) when in overall Internet activity, cognitive attitudes are the same (McMahan et al., 2009). This may explain why website features specifically targeting gender appear unsuccessful (McMahan et al., 2009; Palanisamy, 2005).

At the end of the day, we are left with a mixed bag of research results. Studies in information processing point to difference in processing strategies between males and females; however, many of these studies also show high variances (Chang et al., 2005; Zhou et al., 2007). Advertising research in traditional mediums indicates gender differences in reaction to advertising copy; however, this line of research has not been extensively replicated in the online mediums. Finally, there is a dearth of research investigating whether or not these gender differences in information processing result in any differences in actual behavior outcomes. In other words, just because there is a difference in the way information is processed does not necessarily imply there will be different behaviors. In fact, research in information processing decision-making between genders has shown no differences in quality or outcome (Jansen & McNeese, 2005). As such, we do not know whether or not gender is a key demographic attribute for keyword advertising.

Table 3

Major findings concerning gender influences from online shopping research.

Online shopping	Male	Female
Citrin et al. (2003) and Cho (2004)		Prefer tactile experience of shopping which environment lacks
Cyr and Bonanni (2005)	Spend more time and money online, have more trust and satisfaction in the experience	Enjoy shopping less
Rodgers and Harris (2003)	Prefer efficiency and convenience, prefer less interaction	Prefer emotional and social interaction of conventional shopping
van Slyke et al. (2002)	Have more availability and access to products they prefer	Products of interest not online
Garbarino and Strahilevitze (2004)		Perceived risks of environment are higher
Dittmar et al. (2004)	Online shopping no different than conventional shopping	Like online shopping less than conventional
Sanchez-Franco (2006)		Attitude plays stronger role in shopping
McMahan et al. (2009)	Prefer more interactive features: videos, web games, website product customization	Responded to retailers communication links, promotional info links, events links, and downloads
Hasan (2010)	Cognitive, affective, and behavioral attitudes are higher	Cognitive, affective, and behavioral attitudes are lower, more skeptical of the benefits

A literature review of demographic advertising and human information processing gender research leaves several open questions. Do gender differences affect online searching and subsequent purchasing behavior in sponsored search? If there are gender differences, what are these differences? What are the effects of these differences? Do gender differences affect advertising costs or profits? Do they affect online commercial searching behavior? These are some of the questions that motivate our research.

3. Research question

Our research question is: *Is there a significant difference in the interaction with sponsored results based on gender?*

With an understanding of the relationship between gender and consumer information searching, there is an opportunity for online businesses to optimize their search engine marketing strategies and personalize their advertising efforts. Results from such research can serve a variety of purposes such as ad creation, valuable queries, and research to support the use of more advanced targeting approaches.

However, it is unrealistic to expect online advertisers to know the gender of each individual searcher. Given the focus of sponsored search and information searching on the query, we therefore concentrate on the search phrase. Specifically, we examine the gender orientation of the key phrase that triggers a sponsored search ad. Unlike user profiles of searchers, key phrases are well within the control and data collection of online advertisers and search engines. Therefore, their use in this research has both theoretical importance and practical implications for sponsored search.

Based on our research question and prior work suggesting differences in the information processing of males and females (c.f., Meyers-Levy, 1988), our hypotheses are:

Hypothesis 01. There will be a significant difference between females and males in the number of impressions based on the gender-orientation of the query.

A key metric in any keyword advertising campaign is impressions (i.e., the number of times that a given advertisement appears on a SERP in response to a query submitted by a potential consumer). Analyzing the occurrence of impression triggered by keywords would shed light on the gender differences in terms of frequency of ecommerce searching for a retailer's given set of keyword. By categorizing keywords by gender, online businesses could tell the gender make-up of their online customers.

Hypothesis 02. There will be a significant difference between females and males in the number of clicks based on the gender-orientation of the query.

The goal of most keyword advertising campaigns is to get potential consumers to click on a given advertisement, and the click is a commonly used measure of potential interest in search engine results. The click through rate (i.e., number of clicks divided by the number of impressions) is one of the most important measures of a sponsored search advertising campaign. As such, any gender differences would shed important light on differentiation in ecommerce advertising.

Hypothesis 03. There will be a significant difference between females and males in the cost per click based on the gender-orientation of the query.

Advertisers must bid different amounts for different key phrases depending on the value that they place on those key phrases and the competition from other advertisers. One would expect the key phrases that advertisers expect to garner more clicks would be the most expensive. Therefore, higher cost-per-click for certain key phrases classified along gender lines would indicate preferences for those genders by online advertisers.

Hypothesis 04. There will be a significant difference between females and males in the average sales revenue based on the gender-orientation of the query.

Most online advertisements have the aim of generating a sale or lead (i.e., identifying a potential customer). Naturally, gender differences in terms of sales revenue would provide insight into the receptiveness of these searchers to online purchasing and the profitability of these searchers for online advertisers.

Hypothesis 05. There will be a significant difference between females and males in the number of orders based on the gender-orientation of the query.

Related to sales revenue, companies track the number of orders placed for a given set of keywords. Any differences in number of orders among gender key phrases would be an indication of differences in online shopping behaviors, such as willingness to purchase or amount willing to spend.

Hypothesis 06. There will be a significant difference between females and males in the number of items purchased based on the gender-orientation of the query.

Associated with orders, number of items purchased per order is a key metric of online sales. Consumers who purchase multiple items may be more valuable than consumers who purchase only a single item. Gender differences in this regard would be of profound importance in several areas of ecommerce.

Hypothesis 07. There will be a significant difference between females and males in the return on advertising based on the gender-orientation of the query.

While a particular key phrase or advertisement may generate sales, the final evaluation of any advertising effort, in terms of effectiveness, is how much profit the effort generates, which is referred to as the return on advertising (ROA). The ROA equals the gross sales of a particular advertising effort minus the cost of that advertising effort. If the ROA is positive, the advertising effort is effective. If the ROA is negative, the effort is ineffective, as it costs more to run the campaign than the campaign is generating in revenue. Therefore, ROA is of critical importance for advertisers.

4. Methods

4.1. Overview of sponsored search

In sponsored search campaigns, advertisers bid on key phrases that:

- relate to some product or service that they are providing and that
- they believe searchers will submit to the search engine.

These key phrases provide the link between the results provided from the advertiser and the queries submitted by potential customers, who are the searchers on the Web search engines. When searchers enter queries that match a key phrase, the corresponding set of advertising results is displayed on the SERP. Although published data is sparse, reports are that about 15% of search engine clicks occur on these keyword advertisements (Jansen & Spink, 2009).

The cost of the ad for the advertiser is determined via an online auction. The exact cost can be in constant flux, as the amount that an advertiser *must* bid to get an ad to display depends on the overall demand for that key phrase at a given time. The amount that an advertiser is *willing* to bid depends generally on the perceived possible value of the customer. Multiple advertisers are typically bidding on the same key phrases simultaneously, so the online auction and bid price can be quite dynamic. However, the search engines provide advertisers an assortment of tools to effectively manage their bids, control risk, and maximize opportunity.

The sponsored results on the SERP are usually shown above the organic results listing (i.e., the north position), to the right of the organic results listing (i.e., the east position), and below the organic results listing (i.e., the south position). This depends on the search engine, as some only use the north and east positions. The sponsored result's rank within each listing depends on the bid price, the other bids in the auction, and a quality score (i.e., determined by several factors including click through history and landing page relevance to the ad, although this formula varies somewhat by search engine). Therefore, the sponsored search process is an interesting and complex integration of business process, information technology, and information processing, making it an interesting general system for study.

The sponsored search results are usually textual in nature and normally consist of a short headline, two diminutive lines of text describing the product or service, and a hyperlink that points to the advertiser's landing page (i.e., an advertiser designated Webpage). The predominant keyword advertising model is pay-per-click (PPC), where an advertiser only pays the search engine if a searcher actually clicks on the displayed ad hyperlink.

The entire sponsored search process can be extremely complex, and this brief overview cannot do it justice. Again, the interested reader is referred to review articles (Fain & Pedersen, 2006; Jansen & Mullen, 2008) of the sponsored search process.

4.2. Data

The data file used for this research contains daily information of a nearly 4 year a sponsored search campaign from a large nationwide retailer, with both more than 200 brick-and-mortar stores and a significant online sales presence, after decades of selling merchandise via catalog. This retailer sells a variety of novel gifts, electronic, and household products. With several hundred stores and an active online presence, the retailer offers a variety of innovative products covering a wide price range, from a few dollars to several hundred dollars. The retailer's self-advertisement promotes innovative gifts for both men and women. Over the course of many years in the marketplace, the retailer has developed a considerable brand presence. Consumers can also purchase gift cards, which are a popular item for both men and women. Given the national presence, well-known brand, both real and virtual stores, and wide range of products, we consider the retailer to be an excellent site for the study of keyword advertising.

The data is a record of the sponsored search advertising efforts by the company during a 33-month period, spanning 4 calendar years, from 30 September 2005 to 09 June 2008. The log contains a rich data set in that we have the key phrase that triggered the ad, searcher responses, and sales information.

Applicable fields in the log used for the research reported here are shown in Table 4.

The log file contains approximately 7 million records from nearly 40,000 key phrases. The data log contains a record for every day in which one of the key phrases triggered an ad. There is a unique record for each key phrase for a given day. Each record in the data log has a variety of information associated by key phrase for a given day. The record includes the key

Table 4

Fields and descriptors from data log of a sponsored search campaign.

Field	Description
Key phrase	The key phrase that triggered the advertisement
Impressions	The total number of impression for that day for the given advertisement triggered by the given key phrase
Clicks	The number of clicks on the advertisement for that day for a given key phrase
Cost	The total cost for the day for a given key phrase for a given advertisement
Sales	The revenue generated from that advertisement on that day for a given key phrase
Orders	The number of orders from the advertisement for that day for a given key phrase
Items	Number of items purchased from that advertisement on that day for a given key phrase from all orders. One order could have one or more items

phrase that triggered the ad, number of impressions for that phrase on that day, the number of clicks, the average CPC, the number of conversions (or orders), the total sales revenues, and the total number of items ordered. A query may trigger an impression but no click. If there is a click, there may not be a conversion. Given the 4 years of data collection and the relative stability of the sponsored search platforms during this period, we believe that the data provides insights into online e-commerce web searching, in addition to providing findings concerning keyword advertising.

With the exception of Ghose and Sha Yang (2008), there have been limited, published empirical studies of sponsored search campaigns using real-world data, and there are no published studies using large-scale and robust datasets. Therefore, we believe our dataset to be a rich source in which to investigate our research question and hypotheses concerning the effect of gender for sponsored search.

4.3. Gender classification

To address our research question, the first step was to classify approximately nearly 40,000 key phrases with some probability of their gender-orientation. To do this, we used the Microsoft adCenter Labs Demographics Prediction Tool (<http://adlab.microsoft.com/Demographics-Prediction/DPUI.aspx>). This application takes a given search phrase and provides the probability that the query is male or female-oriented (i.e., provides a probability for both within range of 0–1 inclusive).

As stated on the tool's website, the Demographics Prediction Tool helps advertisers learn the demographic-orientation of key phrases before bidding on these terms, with the goal of improving demographic targeting. Hypothetically, with this knowledge, advertisers could determine whether they are reaching their target audience and decide if current keyword bidding is effective. The gender-orientation probabilities are based on a 1-month MSN Live Search user online behavior log. Fig. 1 shows the Demographics Prediction tool.

Note that the gender-oriented prediction is specific to the query, not necessarily the searcher. That is, although the gender of the searcher may not be known, the query has a gender-orientation based within a certain confidence interval. The probability confidence is based on Microsoft's predictive model that speculates whether or not a particular query fits gender-specific trends of online behavior based on a log analysis.

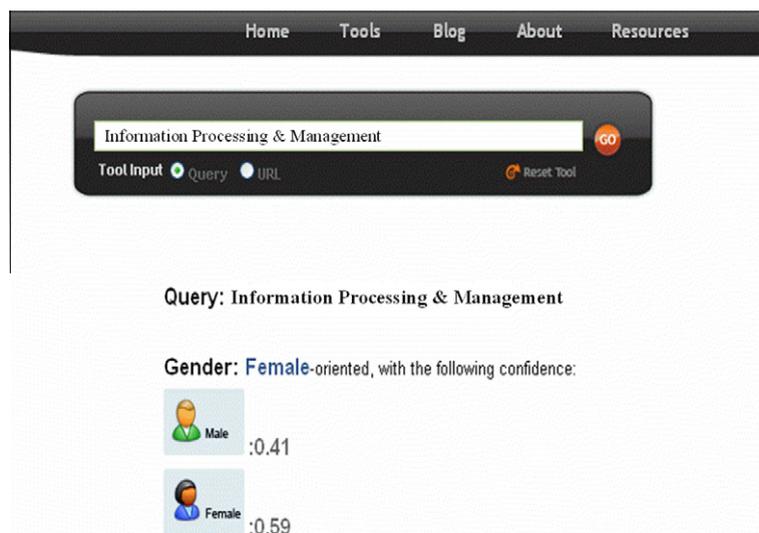


Fig. 1. MSN Demographics Prediction tool with results using the query *Information Processing and Management*.

How accurate is the demographic classification? We could locate no published manuscripts addressing this question. However, MSN adCenter Labs has access to large numbers of user profiles (e.g., Hotmail and MSN Passport), so it is a trivial matter to link queries to demographic profiles. Therefore, we expect the accuracy of the gender classification tool to be high. However, to evaluate the effectiveness of the tool, we conducted an experiment comparing the MSN Demographics Prediction with a similar tool, Yahoo! Clues.

Yahoo! Clues is a comparable tool to the MSN Demographics tool although it suffers from a shortcoming of being less robust than its MSN counterpart. Based on experimentation, Yahoo! Clues seems to have much less demographical data compare to the MSN tool. We addressed this problem by stemming out search terms that were too descriptive, so that the Yahoo! Clues tool would be able to recognize them. We then used the resulting, smaller term set (i.e., 908 key phrases) to collect data from both the MSN Demographics and Yahoo! Clues tools comparing the results. We considered the tools in agreement if the gender probability was within ten percentage points. Inter-rater agreement based on Cohen's Kappa was 0.75 ($p < 0.01$), indicating substantial agreement (Landis & Koch, 1977). So, this finding was reassuring for us that our gender classifications were reliable.

Using the MSN Demographics Prediction tool, we classified the nearly 40,000 key phrases in the data set for gender-orientation using an automated script that submitted the query and then retrieved the gender-orientation probabilities. The script operated using an automated submission to the MSN Demographics Prediction tool via a RESTful request, which sends a GET or a PUT to a website in order to collect specific data from that site. Once a response was given, the script extracted, separated, and stored the response in a flat text file. We then used the data in this file for our statistical analysis.

For this research, we devised seven categories of gender-orientation classification strength based on the probability returned by the MSN adCenter Labs predictive model. The seven categories, a probability range, and example key phrases are given in Table 5.

From Table 5, of the seven categories, there were three male, three female, and one gender neutral. The gender specific categories represents 10–15% of the entire probability range, with Gender Neutral representing the largest range (20%) and Generally Male and Generally Female the smallest (10%) each. Other categories were 15% of the probability range. Our classification into these seven categories was designed to both tease apart the levels of gender-orientation personalization and make the results usable for practitioners in the field. Our selection of exact ranges for each category was somewhat arbitrary; however, it appears that people are best able to perceive probabilities in five and ten scale increments (Ayres, Ayres-Brown, & Ayres-Brown, 2007).

The occurrences for each category for the nearly 40,000 unique phrases are shown in Table 6. We see from Table 6 that 39% of the key phrases were oriented male (i.e., Really Male, Generally Male, and Leaning Male). There were 36% of the key phrases that were oriented female (i.e., Really Female, Generally Female, and Leaning Female). So, the gender-orientation was roughly in balance.

4.4. Data analysis

Once each unique key phrase was classified, we used this set of queries to automatically classify the nearly 7 million records in the complete dataset. Once this was accomplished, we could then link each key phrase, now classified into one of the gender-orientation categories, to the associated user behavioral and sales data. Once done, we could then examine each key phrase occurrence based on our research question and hypotheses. A snippet of records from the data log showing applicable fields is shown in Table 7.

Once we processed all the information, we imported the collected data into SPSS, which we used to run the ANOVA tests to investigate differences among the means among the gender categories. However, our data is not multivariate normal; instead, it has a power law distribution. We transformed the data via the Box–Cox power transformation (Box & Cox, 1964) by using $\log(\text{variable} + 1)$. After employing the Box–Cox power transformation, we plotted our data to check for normalization. The data were successfully normalized, although the distributions were skewed to the left (i.e., weighted toward lower cost click, lower sales, lower number of items ordered, etc.), which would be reasonable given the type of data. Although skewed, several prior works have noted that the ANOVA method is remarkably robust to such deviations from normality (c.f., Box &

Table 5
Gender classification categories with probability ranges and example key phrases.

Gender classification	Probability range	Example of key phrases in category
Really Male	1.00 to >0.85 (Male)	Electronic chess blue tooth wireless headsets
Generally Male	0.85 to \geq 0.75 (Male)	Cool automotive lights self-powered radios
Leaning Male	0.75 to \geq 0.60 (Male)	Blue mini ipod shake flash light
Gender Neutral ^a	0.60 > (Male) and <0.60 (Female)	Clock radio silent air purifier
Leaning Female	0.75 to \geq 0.60 (Female)	Quality bathroom lights bed bath devices
Generally Female	0.85 to \geq 0.75 (Female)	Special kitty bowls presents for dads day
Really Female	1.00 to \geq 0.85 (Female)	Talking photo album unique christmas presents

^a Note that the Gender Neutral category range is twice as large.

Table 6
Gender classification of key phrases occurrences and probabilities.

Gender classification	Occurrences	%
Really Male	4322	10.9
Generally Male	3738	9.4
Leaning Male	7601	19.2
Gender Neutral	9517	24.1
Leaning Female	8133	20.6
Generally Female	4515	11.4
Really Female	1735	4.4
	39,561	100.0

Note: Highest percentage is **bolded**.

Table 7
Snippet from data log showing applicable fields and sample data values.

Gender classification	Impressions	Clicks	CPC	Sales revenue	Orders	Items ordered
Generally Male	5	1	0.20	49.95	1	1
Gender Neutral	37	3	0.45	29.95	1	3
Really Female	10	5	0.26	9.96	3	6
Gender Neutral	2	1	0.10	56.01	1	4

Anderson, 1955; Hull, 1993; Lindman, 1974). The use of the power transformation, along with other measures (discussed later), ensured our statistical approach was valid.

5. Results

Prior to investigating our specific hypotheses, we provide some aggregate results from our data analysis. We first present overall statistics for the data set of 6,871,461, as shown in Table 8.

From Table 8, we see that this was a substantial keyword advertising effort generating more than \$56 million in sales and moving nearly seven hundred thousand items. Table 8 also presents the average figures per day and the standard deviations. The standard deviations are high due to the nature of retailing, when there are substantial sales during the holiday buying season, typically October through early January.

5.1. Gender-orientation classification

Using the nearly 40,000 queries labeled in gender-orientation categories, we automatically categorized the entire approximately seven million query dataset, with findings presented in Table 9.

Comparing Table 6 (i.e., unique key phrases) and Table 9 (i.e., key phrases occurring in total dataset), we see that the percentages are generally similar. The only difference in ranking between the two data samples is a swap between Leaning Female (the second most frequent key phrase) and Leaning Male (the third most frequent key phrase), based on occurrences in the entire dataset. A Spearman's rank correlation coefficient test found no significant difference between the two rankings, indicating that the key phrases for the search engine marketing campaign are in line with the expressed interests of the on-line consumers.

Examining the comparison graphically, as is shown in Fig. 2, the female-oriented queries are generally over represented in the set of unique key phrases relative to the overall dataset. Since the sample sizes are uneven, this could confound our analysis. Therefore, we use the weighted means for the samples in the data analysis (Clark-Carter, 2009, p. 249), which addresses the confounding issue in our ANOVA methods. Additionally, we employed the Welch equality of means (Welch, 1947), which does not assume equal sample sizes or equal variance, to verify our results.

Table 8
Aggregate statistics from the dataset showing values of standard keyword advertising metrics.

	Total	Average (by day)	Standard deviation
Impressions	423,211,204	61.59	809.61
Clicks	13,286,944	1.93	43.95
Advertising Cost	\$8,484,855	\$1.24	\$19.68
Sales	\$56,232,819	\$8.18	\$377.02
Orders	372,445	0.05	2.61
Items	691,023	0.10	5.17

Table 9

Gender classification of key phrases in dataset along with occurrences and percentages.

Gender Classification	Occurrences	%
Really Male	804,896	11.7
Generally Male	824,850	12.0
Leaning Male	1,539,031	22.4
Gender Neutral	1,922,106	28.0
Leaning Female	1,141,424	16.6
Generally Female	446,134	6.5
Really Female	185,740	2.7
	6,864,181	100.0

Note: Highest percentage is **bolded**.

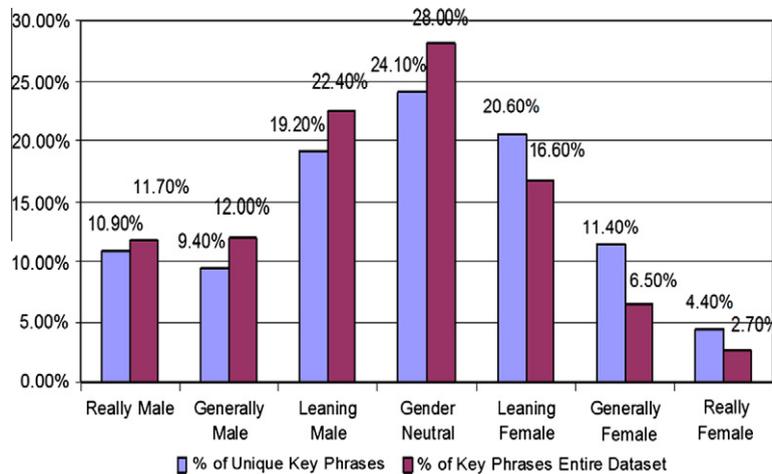


Fig. 2. Comparison of key phrases occurrences.

5.2. Hypotheses testing

In evaluating our seven hypotheses, we used a one-way ANOVA statistical analysis to compare means and variance between the groups. The one-way ANOVA tests whether two or more groups are significantly different. For all ANOVA tests presented, the critical value of $F = 3.78$ and the critical value of P was 0.05. We also used Tamhane's T2 Test as the post hoc analysis to uncover the exact differences among the groups, which does not assume equal variance among samples.

Hypothesis 01. There will be a significant difference between male and female key phrases in the number of impressions based on the gender-orientation of the query.

The results indicate that there is a significant difference among the categories ($F(6) = 31967.39, p < 0.01$). The Tamhane's T2 Test indicated a significant difference among all seven gender classifications. Therefore, Hypothesis 01 is fully supported.

From Table 10, we see that the mean number of impressions was just more than 61 for all gender-classifications. However, the mean is skewed due to the extremely high mean number of impressions for key phrases with a gender-neutral ori-

Table 10

Mean impressions by gender classification.

Gender classification	Mean	St. dev.
Really Male	29.37	549.73
Generally Male	32.09	203.36
Leaning Male	49.57	409.80
Gender Neutral	107.75	1,262.83
Leaning Female	54.21	422.29
Generally Female	38.11	366.43
Really Female	36.40	254.44
All Categories	61.06	752.11

Note: Highest mean value is **bolded**. $p < 0.05$ for all Tamhane's T2 Test.

entation. All other categories were below the overall mean. We also note that female-oriented key phrases generated more impressions than the corresponding male-oriented categories.

This would indicate that prior reports of females being less inclined to shop online (c.f., Rodgers & Harris, 2003) may no longer be valid. At least for this retailer's set of keyword, there are as many or more female-oriented queries as male-oriented ones even those the retailer had bid on many more male oriented key phrases than female ones.

Hypothesis 02. There will be a significant difference in the number of clicks based on the gender-orientation of the query.

The results indicate that there is a significant difference among the categories ($F(6) = 35638.70$, $p < 0.01$). The Tamhane's T2 Test indicated a significant difference among all seven gender classifications. Therefore, Hypothesis 02 is fully supported.

From Table 11, we see that the mean number of clicks overall is 1.93 clicks per occurrence of a key phrase. However, the mean is again skewed due to the Gender Neutral, with a mean number of clicks of 4.17. All other categories were below the overall mean. Combined with the higher rate of impression, this would indicate that not only are females conducting online ecommerce searching, they are willing to engage in online advertising.

Hypothesis 03. There will be a significant difference in the cost per click based on the gender-orientation of the query.

The results indicate that there is a significant difference among the categories ($F(6) = 8071.72$, $p < 0.01$). The Tamhane's T2 Test indicated a significant difference among all seven gender classifications. Therefore, Hypothesis 03 is fully supported.

From Table 12, our pattern that we saw with impressions and clicks does not hold with CPC. The mean CPC for both Gender Neutral (\$0.70) and Leaning Male (\$0.85) oriented key phrases were the highest mean CPC, and the other five categories lower than the overall mean. Although, we saw earlier that the female-oriented key phrases generated more impression and clicks, the CPC is higher for two of the male oriented categories relatively to the corresponding female-oriented categories.

Advertisers have an incentive to bid higher on the key phrases that convert more customers. So, these higher CPC would indicate that, regardless of the number of impressions or clicks, the advertisers consider these potential customers (i.e., those searching with key phrases of Gender Neutral, Leaning Male, and Generally Male orientation) of higher value, which is reflected in the higher CPC for these gender-oriented categories.

Hypothesis 04. There will be a significant difference in the average sales revenue based on the gender-orientation of the query.

The results indicate that there is a significant difference among the categories ($F(6) = 14495.96$, $p < 0.01$). The Tamhane's T2 Test indicated a significant difference among all seven gender classifications. Therefore, Hypothesis 04 is fully supported.

We see from Table 13 that the Gender Neutral oriented key phrases generated far more average sales revenue than any other category. We also see that the female-oriented key phrases generate more average revenue than the corresponding male-oriented categories.

Table 11
Mean clicks by gender classification.

Gender classification	Mean	St. dev.
Really Male	0.51	3.82
Generally Male	0.72	4.07
Leaning Male	1.14	12.14
Gender Neutral	4.17	80.58
Leaning Female	1.59	12.19
Generally Female	0.97	6.84
Really Female	1.16	15.74
All Categories	1.93	43.49

Note: Highest mean value is **bolded**. $p < 0.05$ for all Tamhane's T2 Test.

Table 12
Mean CPC by gender classification.

Gender Classification	Mean (\$)	St. dev. (\$)
Really Male	0.43	0.52
Generally Male	0.59	0.83
Leaning Male	0.85	1.08
Gender Neutral	0.70	0.91
Leaning Female	0.53	0.57
Generally Female	0.51	0.48
Really Female	0.51	0.52
All Categories	0.65	0.85

Note: Highest mean value is **bolded**. $p < 0.05$ for all Tamhane's T2 Test.

Table 13
Mean sales revenue by gender classification.

Gender classification	Mean (\$)	St. dev. (\$)
Really Male	0.80	70.95
Generally Male	0.97	20.30
Leaning Male	2.34	46.51
Gender Neutral	23.31	705.51
Leaning Female	4.69	94.93
Generally Female	1.64	46.48
Really Female	1.44	31.21
All Categories	8.19	377.22

Note: Highest mean value is **bolded**. $p < 0.05$ for all Tamhane's T2 Test.

This would indicate that, based on sales revenue, the higher CPCs for the male-oriented key phrases are not justified, as the female-oriented phrases generate higher sales revenue.

Hypothesis 05. There will be a significant difference in the number of orders based on the gender-orientation of the query.

The results indicate that there is a significant difference among the stages ($F(6) = 14715.71$, $p < 0.01$). The Tamhane's T2 Test indicated a significant difference among all seven gender classifications. Therefore, Hypothesis 05 is fully supported.

From Table 14, the highest mean was, again, the Gender Neutral oriented queries, with a mean order of 0.15 per query. The male and female oriented queries were balanced at the Generally Female and Generally Male.

Hypothesis 06. There will be a significant difference in the number of items purchased based on the gender-orientation of the query.

The results indicate that there is a significant difference among the stages ($F(6) = 14511.59$, $p < 0.01$). The Tamhane's T2 Test indicated a significant difference among all seven gender classifications. Therefore, Hypothesis 06 is fully supported.

From Table 15, we see that the Gender Neutral oriented key phrases had the highest mean number of items sold (0.28). The female-oriented key phrases had a higher mean number of items sold than the corresponding male-oriented phrases.

Hypothesis 07. There will be a significant difference between females and males in the return on advertising based on the gender-orientation of the query.

The results indicate that there is a significant difference among the stages ($F(6) = 15874.77$, $p < 0.01$). The Tamhane's T2 Test indicated a significant difference among all seven gender classifications, except between the Really Male and Generally Male categories. Therefore, Hypothesis 07 is supported.

Table 14
Mean orders by gender classification.

Gender classification	Mean	St. dev.
Really Male	~0.00	0.15
Generally Male	0.01	0.12
Leaning Male	0.02	0.232
Gender Neutral	0.15	4.92
Leaning Female	0.04	0.65
Generally Female	0.01	0.14
Really Female	0.02	0.31
All Categories	0.05	2.62

Note: Highest mean value is **bolded**. $p < 0.05$ for all Tamhane's T2 Test.

Table 15
Mean items sold by gender classification.

Gender classification	Mean	St. dev.
Really Male	0.01	0.27
Generally Male	0.01	0.24
Leaning Male	0.03	0.44
Gender Neutral	0.28	9.72
Leaning Female	0.07	1.30
Generally Female	0.02	0.34
Really Female	0.03	0.64
All Categories	0.10	5.18

Note: Highest mean value is **bolded**. $p < 0.05$ for all Tamhane's T2 Test.

Table 16
Mean ROA by gender classification.

Gender classification	Mean (\$)	St. dev. (\$)
Really Male	0.48	67.44
Generally Male	0.57	19.92
Leaning Male	1.35	41.39
Gender Neutral	20.75	694.70
Leaning Female	3.80	93.55
Generally Female	0.93	44.04
Really Female	0.82	25.65
All Categories	6.95	371.18

From Table 16, we see that Gender Neutral key phrases had by far the highest ROA (\$20.75). This would indicate that, in conjunction with other findings, that the use of gender neutral key phrases are the most profitable for the advertiser, factoring in both revenue and costs.

5.3. Effect size

Although there were statistically significant differences in our brand groupings, a portion of this difference can be attributed to the large sample sizes. We performed an additional analysis to ensure that the observed differences among the brand groupings were meaningful. We used Cohen's *d* tests to determine the effect size of each between group (Cohen, 1988). Table 17 shows the obtained *d*-values.

From Table 17, almost all of the comparisons show a small effect, meaning that although the results were statistically significant, and the outcome on keyword advertising metrics are of some but limited significance. Also, we see in Table 14 that in nearly all cases the Gender Neutral had the highest practical result. From an analysis of the Cohen's *d* results, it is clear that gender demographic categories are of limited practical guidance. There appears to be no advantage overall for demographically targeting by gender by itself. However, some of the small effect size is due to the high standard deviations. Regardless, the overall trend of the results is clear – the gender neutral queries (i.e., those without substantial gender target performed the best. In the next section, we discuss this overall trend, along with implications for keyword advertising.

6. Discussion and implications

6.1. Discussion of results

In this research, we investigated the gender-orientation of key phrases as an aspect of human information processing and personalization. Traditional wisdom would hold that increased personalization will lead to a better online experience for the searcher, thereby leading to increased revenue for the advertiser. Survey results clearly indicate that marketers are willing to pay more for demographically targeted advertising (SEMPO, 2011). However, the research results from our study indicate that increased gender-oriented key phrases are more expensive and do not generate increase sales revenue. In fact, by all metrics examined, the key phrases that were gender neutral performed the best overall. In other words, the key phrases that were less personalized performed better.

The gender-neutral key phrases generated nearly twice as many impressions as any other category of key phrases and nearly two and half times as many clicks. So, not only are these gender-neutral phrase the most searched for (i.e., more impressions), they are the most clicked on (i.e., have the most consumer interest).

Although, the gender-neutral key phrases are the second most expensive (trailing Leaning Male by 21%), they generated far and away the most sales revenue, beating the nearest other category (Leaning Female) by a factor of nearly five. Continuing, these gender-neutral queries generate four times the orders of the nearest other category (again Leaning Female) and also four times as many items ordered. All in all, gender neutral-oriented queries appear to be the best performing by any of the standard keyword advertising metrics.

However, there were also some interesting findings when comparing the male and female oriented key phrases. The female-oriented key phrases generated about 10–20% more impressions and approximately 40–127% more clicks than the male oriented key phrases. Obviously, the female-oriented key phrases are triggering more ads, and the searchers submitting these female-oriented queries are engaging more with the resulting ads than searchers submitting male-oriented queries.

Additionally, these female-oriented key phrases generated 70–100% more mean sales revenue than the analogous male-oriented key phrases, with nearly the same mean number of orders (although Leaning Female key phrases are higher than the Leaning Male phrase). This would imply that the searchers with female-oriented queries are making more expensive orders. This is also borne out in the number of items ordered, with the average number of items ordered for female-oriented key phrase 100–200% higher than male-oriented phrase. These factors would indicate that, aside from gender neutral ones, female-oriented key phrases are the most valuable, given they generate the most sales. However, this is not reflected in the mean CPC for these key phrases, as illustrated in Fig. 3.

Table 17
Cohen's *d* results to account for effect sizes.

Gender categories	1	2	3	4	5	6	7
<i>Impressions</i>							
Really Male (1)							
Generally Male (2)	0.10						
Leaning Male (3)	0.08	0.08					
Gender Neutral (4)	0.12	0.12	0.12				
Leaning Female (5)	0.08	0.08	0.08	0.08			
Generally Female (6)	0.13	0.13	0.13	0.13	0.13		
Really Female (7)	0.05	0.05	0.05	0.05	0.05	0.05	
<i>Clicks</i>							
Really Male (1)							
Generally Male (2)	0.10						
Leaning Male (3)	0.07	0.07					
Gender Neutral (4)	0.16	0.16	0.16				
Leaning Female (5)	0.09	0.09	0.09	0.09			
Generally Female (6)	0.11	0.11	0.11	0.11	0.11		
Really Female (7)	0.05	0.05	0.05	0.05	0.05	0.05	0.05
<i>CPC</i>							
Really Male (1)							
Generally Male (2)	0.14						
Leaning Male (3)	0.25*	0.25*					
Gender Neutral (4)	0.14	0.14	0.14				
Leaning Female (5)	0.09	0.09	0.09	0.09			
Generally Female (6)	0.06	0.06	0.06	0.06	0.06		
Really Female (7)	0.04	0.04	0.04	0.04	0.04	0.04	0.04
<i>Sales</i>							
Really Male (1)							
Generally Male (2)	0.03						
Leaning Male (3)	0.06	0.06					
Gender Neutral (4)	0.10	0.10	0.10				
Leaning Female (5)	0.06	0.06	0.06	0.06			
Generally Female (6)	0.07	0.07	0.07	0.07	0.07		
Really Female (7)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<i>Orders</i>							
Really Male (1)							
Generally Male (2)	0.03						
Leaning Male (3)	0.06	0.06					
Gender Neutral (4)	0.10	0.10	0.10				
Leaning Female (5)	0.07	0.07	0.07	0.07			
Generally Female (6)	0.07	0.07	0.07	0.07	0.07		
Really Female (7)	0.02	0.02	0.02	0.02	0.02	0.02	0.02
<i>Items</i>							
Really Male (1)							
Generally Male (2)	0.03						
Leaning Male (3)	0.06	0.06					
Gender Neutral (4)	0.10	0.10	0.10				
Leaning Female (5)	0.07	0.07	0.07	0.07			
Generally Female (6)	0.07	0.07	0.07	0.07	0.07		
Really Female (7)	0.02	0.02	0.02	0.02	0.02	0.02	0.02
<i>ROA</i>							
Really Male (1)							
Generally Male (2)	0.04						
Leaning Male (3)	0.06	0.06					
Gender Neutral (4)	0.12	0.12	0.12				
Leaning Female (5)	0.08	0.08	0.08	0.08			
Generally Female (6)	0.09	0.09	0.09	0.09	0.09		
Really Female (7)	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: * – Small effect.

From Fig. 3, the category with the highest mean CPC is Leaning Male key phrases, even though this category is the only ranked third in terms of sales revenue generation (trailing Gender Neutral by a factor of ten and Leaning Female by a factor of

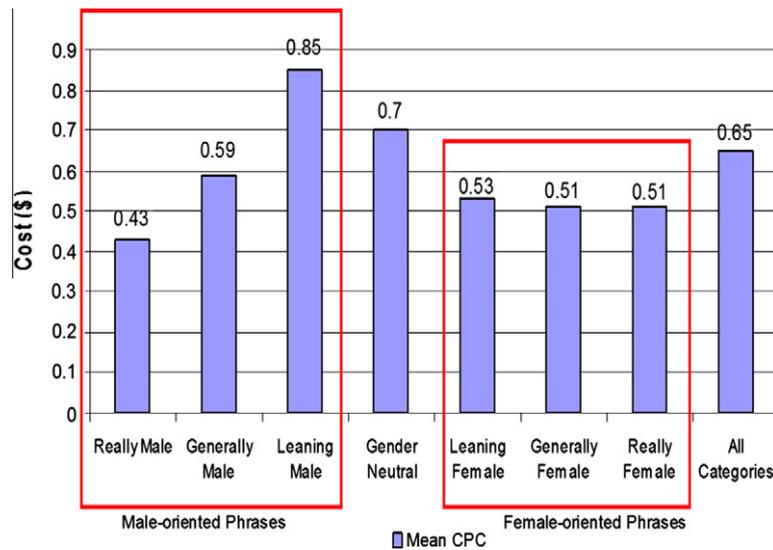


Fig. 3. Comparison of mean CPC by gender classification showing groupings of male and female relative to gender neutral and overall.

two). Also, the Generally Male key phrases category is the third most expensive. Even then, it trails all female-oriented key phrase categories in sales revenue. Obviously, from a gender-orientation perspective, the higher mean CPCs for the Leaning Male and Generally Male are not justified relative to their sale generation.

Through the lens of gender classification, these research findings point to an inconsistency in spending on keyword advertising phrase in relation to the measurable results that these phrases are generating.

6.2. Theoretical implications

In terms of human information processing, it appears that our findings do support prior work indicating differences in the way that males and females gather and process information. With SERPs containing both organic and sponsored results, there is a choice of information types available. Females appear more willing to click on sponsored results, as indicated by the higher click through rates on female-oriented phrases. This conforms to earlier information processing research indicating that females are more willing to consider a range of information sources. While men's decision-making tends to be systematic, women's decision-making is less linear and puts more factors into consideration (Hanson, 2009). So, if an advertiser considers gender targeting critical to their online business, it is necessary to incorporate these differences in information processing into their advertising.

The theory of social categories (Deaux, 1984), the basis of demographic targeting in advertising, would indicate the gender-oriented phrase should have performed better. What could account for this perhaps unexpected outcome? An answer might lie in the theory of individual differences (Motowidlo, Borman, & Schmit, 1997), which posits that an individual responds to information according to unique cognitive, affect, psychological, and contextual needs. In such a complex interplay, such as online commerce, gender may just not be a dominant factor in predicting behavior. Trauth has noted similar findings with gender and information technology (Trauth, 2002). Constantinides, Lorenzo-Romero, and Gómez (2010) report that online experience is the primary predictor of ecommerce behaviors.

6.3. Practical implications

Even though our findings are statistically significant, one must ask, "are these differences of practical significance?". Based on the findings from this research study, the gender-neutral key phrases are the most successful in snatching interested searchers (i.e., impressions) and potential customers (i.e., those searchers who click on a sponsored result). The gender-neutral key phrases are also the most profitable, generating the most sales revenue, orders, and items ordered. While this may seem somewhat counter intuitive (i.e., what if one is selling a gender-specific product?), beyond a few specialty items (e.g., personal hygiene, etc.), so many items can be 'gifted' that demographic targeting can be counterproductive (i.e., more expensive and less profitable). Additionally, given individual differences, gender may just be too coarse of a discriminatory factor for keyword advertising when used in isolation.

Of the key phrases that are gender-oriented, the female-oriented ones appear to be the most fruitful area to pursue. The female-oriented key phrases generate more impressions and clicks, relative to the male-oriented phrases. The female-oriented phrases also generate significantly more sales revenue, orders, and items sold. Also, the female phrases generate this advertising success while being generally cheaper than the corresponding male-oriented phrases. So, if an advertiser was to

gender target key phrases, these findings would indicate that female-oriented phrases might be the more fruitful area for demographic targeting.

Correspondingly, the male-oriented key phrases include some of the most expensive average CPC, significantly more expensive than even the gender-neutral phrases. Given their high CPC and relatively lower performance, demographic targeting male-oriented key phrases might not be a fertile endeavor. However, we must caveat this by saying that sponsored search is a multi-variable process. These male-oriented phrases may be associated with some other advertising goal, such as brand awareness or image. However, in terms of sales, they are not generating revenue in line with their relative CPCs.

These results also point to possible design implications for sponsored search technology platforms and services. Many of these permit demographic targeting of advertisements that display on Websites (a.k.a., known as content advertising). However, none of these systems, that the authors could determine, permitted gender targeting of keyword advertisements. Such a feature, triggered by a method similar to that used in this research, would allow advertisers to rotate ads based on the gender-focus of the query, rather than just by key phrases.

6.4. Limitations and strengths

As with any research, there are limitations to our study. First, the data set is from one retail company. Although the dataset is quite large both in terms of number of records and temporal span, one will need additional research using data from other companies and other industry sectors to ensure generalizable results to these other areas. Nevertheless, we believe that the research reported in this study is an important step in the empirical research of demographic targeting in the sponsored search area. Second, the dataset used in this research does not contain the offline customer behaviors of the searchers. Customers may be using Web search engines for some portions of the purchasing process and then using alternative information systems sources for other portions and this behavior may vary by gender. However, at least for the behaviors on the search engine and the landing page, the findings reported here seem to support gender-oriented differences. Finally, the accuracy of the MSN gender tool has not been independently verified. However, the distribution of gender-neutral to gender-specific keywords was what one would expect and the use of a similar tool by another source provided similar findings, so the application's results are at least within expectations.

This research also has several strengths. Based on the significance of the findings, we believe that the research results discussed here make valuable contributions to growing research in the important sponsored search area, with implications for information seeking, marketing, and advertising. Given the substantial impact that sponsored search technology and its related business process has had in the development and growth of the Web and online commerce, it is an area that deserves extensive investigation from a variety of academic disciplines.

As an additional strength of this research, the data set was quite substantial, with nearly seven million records and of a lengthy data collection period (i.e., 33 months spanning four calendar years). The dataset also includes a varied set of search and consumer behavior and interactions (e.g., impressions, clicks, orders place, items ordered, and money spent). Therefore, the research findings provide important insights into behavior in the real world, online commercial domain.

7. Conclusion

The results of this research conclude that gender-oriented key phrases do not generate more sales but can cost more for advertisers relative to gender-neutral queries. Generally, it might be beneficial for online advertisers to devote resources targeting these gender-neutral key phrases. For future work, investigating the gender perception of online advertisements over time could lead, perhaps, to better advertisement creation. This could increase the receptiveness of the ads and in turn enhance the overall effectiveness of a campaign.

Acknowledgments

Our thanks to The Rimm-Kaufman Group, particularly George Michie, for providing data for this study. Dr. Jansen acknowledges a gift from Google and a grant from the Air Force Office of Scientific Research in support of portions of this research. We also acknowledge and thank Alan Rimm-Kaufman, who departed this world on 18 July 2009.

References

- Abraham, L., Morn, M., & Vollman, A. (2010). Women on the web: How women are shaping the internet. <http://www.comscore.com/Press_Events/Presentations_Whitepapers/2010/Women_on_the_Web_How_Women_are_Shaping_the_Internet> Accessed 14.03.11.
- Ayres, I., Ayres-Brown, A., & Ayres-Brown, H. (2007). Seeing significance: Is the 95% probability range easier to perceive? *Chance*, 20(11), 11–17.
- Bae, S., & Lee, T. (2011). Gender differences in consumers' perception of online consumer reviews. *Electronic Commerce Research*, 11(2), 201–214.
- Bezjian-Avery Alexa, B. C., & Iacobucci, D. (1998). New media interactive advertising vs. traditional advertising. *Journal of Advertising Research*, 38(4), 23–32.
- Box, G. E. P., & Anderson, S. L. (1955). Permutation theory in the derivation of robust criteria and the study of departures from assumptions. *Journal of the Royal Statistical Society*, 17(1), 1–34.
- Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, 26(2), 211–252.
- Carpenter, C. J., & Huston-Stein, A. (1980). Activity structure and sex-typed behavior in preschool children. *Child Development*, 51(3), 862–872.
- Carisky, H. L., & Zuckerman, M. E. (1991). Search of gender differences in marketing communication: An historical/contemporary analysis. In J. Costa (Ed.), *Gender and consumer behavior* (pp. 43–52). Salt Lake City, UT: Association for Consumer Research.

- Chang, M. K., Cheung, W., & Lai, V. S. (2005). Literature derived reference models for the adoption of online shopping. *Information and Management*, 42(4), 543–559.
- Cho, J. (2004). Likelihood to abort an online transaction: Influences from cognitive evaluations, attitudes, and behavioral variables. *Information and Management*, 41(7), 827–838.
- Citrin, A. V., Stem, D. E., Jr., Spangenberg, E. R., & Clark, M. J. (2003). Consumer need for tactile input: An internet retailing challenge. *Journal of Business Research*, 56(11), 915–923.
- Clark-Carter, D. (2009). *Quantitative psychological research: The complete student's companion*. New York, NY: Psychology Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Routledge.
- Constantinides, E., Lorenzo-Romero, C., & Gómez, M. A. (2010). Effects of web experience on consumer choice: A multicultural approach. *Internet Research*, 20(2), 188–209.
- Cyr, D., & Bonanni, C. (2005). Gender and website design in e-business. *International Journal of Electronic Business*, 3(6), 565–582.
- Darley, W. K., & Smith, R. E. (1995). Gender differences in information processing strategies: An empirical test of the selectivity model in advertising response. *Journal of Advertising*, 24(1), 41–56.
- Deaux, K. (1984). From individual differences to social categories. Analysis of a decade's research on gender. *American Psychologist*, 39(2), 105–116.
- Dittmar, H., Long, K., & Meek, R. (2004). Buying on the internet: Gender differences in on-line and conventional buying motivation. *Sex Roles*, 50(5–6), 423–444.
- Fain, D. C., & Pedersen, J. O. (2006). Sponsored search: A brief history. *Bulletin of the American Society for Information Science and Technology*, 32(2), 12–13.
- Garbarino, E., & Strahilevitz, M. (2004). Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation. *Journal of Business Research*, 57(7), 768–775.
- Ghose, A., & Yang, S. (2008). An empirical analysis of sponsored search performance in search engine advertising. In M. Najork (Ed.), *First ACM international conference on web search and data mining (WSDM 2008)*, Palo Alto, California (pp. 241–250).
- Hanson, S. (2009). Beyond pink. <http://www.director.co.uk/magazine/2009/1%20jan/marketing_women_62_6.html> Accessed 03.09.09.
- Hasan, B. (2010). Exploring gender differences in online shopping attitude. *Computers in Human Behavior*, 26(4), 597–601.
- Horrigan, J.B. (2008). *Online shopping* (p. 32). Washington, DC: Pew Internet & American Life Project.
- Hull, D. (1993). Using statistical testing in the evaluation of retrieval experiments. In *16th Annual international ACM SIGIR conference on research and development in information retrieval*, Pittsburgh, Pennsylvania (pp. 329–338).
- Jackson, L. A., Ervin, K. S., Gardner, P. D., & Schmitt, N. (2001). Gender and the Internet: Women communicating and men searching. *Sex Roles*, 44(5/6), 363–379.
- Jansen, B. J., & McNeese, M. D. (2005). Evaluating the effectiveness of and patterns of interactions with automated searching assistance. *Journal of the American Society for Information Science and Technology*, 56(14), 1480–1503.
- Jansen, B. J., & Mullen, T. (2008). Sponsored search: An overview of the concept, history, and technology. *International Journal of Electronic Business*, 6(2), 114–131.
- Jansen, B. J., & Spink, A. (2009). Investigating customer click through behaviour with integrated sponsored and non-sponsored results. *International Journal of Internet Marketing and Advertising*, 5(1/2), 74–94.
- Kempf, D., Laczniak, R., & Smith, R. (2006). The effect of gender on processing advertising and product trial information. *Market Letters*, 17(1), 5–16.
- Kirouac, G., & Dore, F. Y. (1983). Accuracy and latency of judgment of facial expressions of emotions. *Perceptual and Motor Skills*, 57(1), 683–686.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174.
- Large, A., Beheshti, J., & Rahman, T. (2002). Gender differences in collaborative web searching behavior: An elementary school study. *Information Processing and Management*, 38(3), 427–443.
- Lindman, H. R. (1974). *Analysis of variance in complex experimental designs*. San Francisco: W.H. Freeman & Co.
- Lorigo, L., Pan, B., Hembrooke, H., Joachims, T., Granka, L., & Gay, G. (2006). The influence of task and gender on search and evaluation behavior using Google. *Information Processing and Management*, 42(4), 1123–1131.
- McMahan, C., Hovland, R., & McMillan, S. (2009). Online marketing communications: Exploring online consumer behavior by examining gender differences and interactivity within internet advertising. *Journal of Interactive Advertising*, 10(1), 61–76.
- Meyers-Levy, J. (1988). Influence of sex roles on judgement. *Journal of Consumer Research*, 14(4), 522–530.
- Meyers-Levy, J. (1989). Gender differences in information processing: A selectivity interpretation. In P. Cafferata & A. Tybout (Eds.), *Cognitive and affective responses to advertising*. Lexington, MA: Lexington Books.
- Meyers-Levy, J., & Maheswaran, D. (1991). Exploring differences in males' and females' processing strategies. *Journal of Consumer Research*, 18(1), 63–70.
- Meyers-Levy, J., & Sternthal, B. (1991). Gender differences in the use of message cues and judgments. *Journal of Marketing Research*, 28(1), 84–96.
- Morrisson, M., & Shaffer, D. (2003). Gender-role congruence and self-referencing as determinants of advertising effectiveness. *Sex Roles*, 49(5/6), 265–275.
- Motowidlo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, 10(1), 71–83.
- Palanisamy, R. (2005). Impact of gender differences on online consumer characteristics on web-based banner advertising effectiveness. *Journal of Services Research*, 4(2), 45–74.
- Prakash, V., & Flores, R. C. (1985). A study of psychological gender differences: Applications for advertising format. *Advances in Consumer Research*, 12(1), 231–237.
- Putrevu, S. (2002). Exploring the origins and information processing differences between men and women: Implications for advertisers. *Academy of Marketing Science Review* 10(1), Article 1.
- Rodgers, S., & Harris, M. A. (2003). Gender and e-commerce: An exploratory study. *Journal of Advertising Research*, 43(1), 322–329.
- Roy, M., & Chi, M. T. H. (2003). Gender differences in patterns of searching the web. *Journal of Educational Computing Research*, 29(3), 335–348.
- Sanchez-Franco, M. J. (2006). Exploring the influence of gender on the web usage via partial least squares. *Behaviour and Information Technology*, 25(1), 19–36.
- SEMPO (2011). *The state of search engine marketing 2010*. In *Survey of advertisers and agencies search engine marketing professional organization (SEMPO)*, February, 2011. Search Engine Marketing Professional Organization.
- Trauth, E. M. (2002). Odd girl out: An individual differences perspective on women in the IT profession. *Information Technology and People*, 15(2), 98–118.
- Ulbrich, F., Christenson, T., & Stankhus, L. (2011). Gender-specific on-line shopping preferences. *Electronic Commerce Research*, 11(2), 181–199.
- van Slyke, C., Comunale, C. L., & Belanger, F. (2002). Gender differences in perceptions of web-based shopping. *Communications of the ACM*, 45(8), 82–86.
- Welch, B. L. (1947). The generalisation of students problem when several different population variances are involved. *Biometrika*, 34(1), 23–35.
- Wilson, T. D. (2000). Human information behavior. *Informing Science*, 3(2), 49–55.
- Wolin, L., & Korgaonkar, P. (2003). Web advertising: Gender differences in beliefs, attitudes and behavior. *Internet Research: Electronic Networking Applications and Policy*, 13(5), 375–385.
- Yeh, J.-C., Hsiao, K.-L., & Yang, W.-N. (2012). A study of purchasing behavior in Taiwan's online auction websites: Effects of uncertainty and gender differences. *Internet Research*, 22(1), 98–115.
- Zhou, L., Dai, L., & Zhang, D. (2007). Online shopping acceptance model – A critical survey of consumer factors in online shopping. *Journal of Electronic Commerce Research*, 8(1), 41–62.