Rochester Institute of Technology

RIT Digital Institutional Repository

Theses

8-2-2013

Effects of reputation and aesthetics on the credibility of search engine results

Bryan Christiansen

Follow this and additional works at: https://repository.rit.edu/theses

Recommended Citation

Christiansen, Bryan, "Effects of reputation and aesthetics on the credibility of search engine results" (2013). Thesis. Rochester Institute of Technology. Accessed from

This Thesis is brought to you for free and open access by the RIT Libraries. For more information, please contact repository@rit.edu.

The Rochester Institute of Technology

Department of Communication

College of Liberal Arts

Effects of Reputation and Aesthetics on the Credibility of Search Engine Results

by

Bryan C. Christiansen

A Thesis submitted

in partial fulfillment of the Master of Science Degree

in Communication & Media Technologies

Degree Awarded:

August 2, 2013

The members of the Committee approve the thesis of Bryan C. Christiansen presented on August 2, 2013.

Kelly Norris Martin, Ph.D.
Assistant Professor
Department of Communication
Thesis Adviser

Ki-Young Lee, Ph.D.
Associate Professor
Department of Communication
Thesis Adviser

Rudy Pugliese, Ph.D.
Professor of Communication
Director, Communication & Media Technologies Graduate Degree Program

---

Patrick Scanlon, Ph.D.
Professor of Communication and Chairman
Department of Communication

Table of Contents

EFFECTS OF REPUTATION AND AESTHETICS ON THE CREDIBLITY OF SEARCH ENGINE RESULTS

Name: Bryan C. Christiansen
Department: Communication
College: Liberal Arts
Degree: Master of Science in Communication & Media Technologies
Term Degree Awarded: Summer Quarter 2013 (20124)

Abstract

Search engines are the primary gatekeepers of online information, but are judged differently than traditional gatekeepers due to the interactive and impersonal nature of the online search process. The researcher distributed an online survey with 141 respondents and conducted 22 observational interviews. Information credibility was tested through measures of expertise, goodwill, and trustworthiness, which were each correlated with perceived reputation and perceived aesthetics. Search engine reputation was found to have moderate correlations with expertise and trustworthiness, and a lesser, but still moderate correlation with goodwill. Aesthetics was related to the credibility measures in similar but lesser proportions. Interviews indicated search habits such as wariness towards commercial interests and the high impact of search intent on the rigor of credibility judgments.

Keywords: search engines, online credibility, expertise, goodwill, trustworthiness, reputation, aesthetics, search habits

Effects of Reputation and Aesthetics on the Credibility of Search Engine Results

Many people now retrieve needed information on any given subject through websites rather than traditional media (Pearson & Pearson, 2008; Tarafdar & Zhang, 2005a; Tarafdar & Zhang, 2005b; Vrontis, Ktoridou, & Melanthiou, 2006). The Internet is unique among media due to its ubiquity in daily life as well as consumers' ability to interact with the system itself, as well as with other users. This process generates a constant stream of new websites, messages, and applications. There is no start or end point on the Internet. It is up to users to navigate where they want to, and thus they have a greater degree of choice, precision, and time to uncover and dissect information than with other media. As this trend of self-defined user information access shows no signs of slowing, it becomes important to assess user perceptions of the credibility of these new information sources. This is particularly evident concerning search engines, the new gatekeepers of information.

Search Engines as Gatekeepers

General assessments of credibility come from judging new information against previously known information, but when searching for information that is wholly unknown, what are the cues that something is believable or not? Herman and Nicholas (2010) posit that we are in an age of "information malnutrition" brought about by the decline of professional gatekeepers, primarily on the Internet. This includes not only editors who regulate traditional media, but also librarians and other authoritative and knowledgeable professionals. Thus, they claim that everyone must now manage their information needs on their own, despite being unqualified to do so. Baildon and Damico (2011) agree, saying that it is difficult to determine author biases and purpose online. They also suggest that students are ill-equipped to contextualize their credibility

decisions (Baildon & Damico, 2011). How do Internet users create judgments in a marketplace devoid of consistent authority?

Adding to the concept that users cannot properly handle the online information overload, Herman and Nicholas (2010) suggest that the Internet is devoid of depth, and users navigate rather than consume information. During most searches, users only view one or two pages out of the thousands available (Herman & Nicholas, 2010). Herman and Nicholas (2010) also protest that the Internet lacks a professional information filter, which indicates that they do not consider search engines to be competent gatekeepers. Meyer, Marchionni, and Thorson (2010) found that people find institutions to be cold and news delivered by a human presence is more credible. In regards to the Internet, this means that an indication of human contributions, such as an author's name or picture, may increase website credibility.

Hargittai, Fullerton, Menchen-Trevino, and Thomas (2010) and Lamb (2004) argue that in the absence of traditional and trusted gatekeepers, search engines provide the initial trust necessary for users seeking content. That is, if a website does not appear on the first few search engine result pages (SERPs), its credibility shrinks (Hargittai, Fullerton, Menchen-Trevino, & Thomas, 2010). Students who participated in a study by Iding, Crosby, Auernheimer, and Klemm (2009) said that there is no difference between making judgments online versus when reading books or a newspaper and there is no way of knowing when content is faulty without previous personal knowledge, regardless of media. Online credibility now lies with the algorithms of search engines like Google, Yahoo!, and Bing. This is because the vast amount of information available online is generally unsorted, save for the efforts of search engines (Lamb, 2004).

The Internet is used for information searches because it is always available, it can answer embarrassing questions, and it may be visited for an unlimited amount of time at any moment of the day (Freeman & Spyridakis, 2004). It is, however, relatively easy for any user to create and disseminate anything they wish because there is no vetting process for posting on the Internet (Baildon & Damico, 2011). Baildon and Damico (2011) go on to say that author credentials are difficult to determine, although the absence of online credentials alone should arouse suspicion. There is a level of danger with misjudgments, which may result not only in acquiring misleading or questionable information, but also the presence of malicious sites posing as authoritative ones that can lead to phishing attacks resulting in stolen money and identification (Schwarz & Morris, 2011).

There are several other issues with the amount of trust we place in search engines to generate the results we need. Search engines do not provide information on whether or not the returned information is trustworthy except for computing the rank on the page (Nakamura et al., 2007). Page ranks, though, are determined by inbound links, keyword-rich content, and easily crawled structures (Batten, 2008). Anyone may design a site that embodies these concepts, though – for example, these are all aspects of blogging sites, which have become increasingly popular and accessible (Batten, 2008). Nakamura et al. (2007) defined trustworthiness as topic majority (the significance or quality of pages resulting from a query), topic coverage (the number of topics on the SERP that were related to the query), and locality of supporting pages (the geography of distribution of supporting pages - that is, if many other pages around the world linked to the site or if it had a very limited local network). Still, algorithms cannot be the ultimate information authority. There is a need for some human judgment, because otherwise there is no

real way to understand the information we receive (Batten, 2008). Nakamura (2008) focused more on what users want to see or what may be improved. The current study is more concerned with what people actually see and why we put so much trust in search engines when they leave out information.

Search Engine Development and Use

According to the Pew Internet project, 73% of all Americans used a search engine in February, 2012, the time the survey was conducted (Pew Internet, 2012). Out of all Internet users, 91% used a search engine in the same time span. Fifty-nine percent of adults use search engines daily, a figure that has doubled since 2004, and out of this, 54% use search engines more than once a day (Pew Internet, 2012). According to the same report, 83% of the people who were using search engines used Google (Pew Internet, 2012).

Search engines "began as university projects that focused more on development and algorithms, and less on revenue generation. Even after transitioning into commercial entities, search engines tended to operate as a free resource to content providers and users alike" (Bhargava & Feng, 2002, p. 117). As venture capital and sweat equity dropped, search engines required new revenue streams, which caused them to invest more in paid ad placements, where companies could pay to have their products show up higher in search results through modifications of the search engine's algorithm (Bhargava & Feng, 2002). Search engines now attempt to closely understand user needs, actions, and intentions in order to improve the searching experience and create a more relevant information stream (Ashkan & Clarke, 2013). They are designed to produce more informational than commercial results (Fox, 2012). Using

past behavior to identify search intent fosters a better experience for users, but a large amount of that information is also being used to create more targeted ads (Ashkan & Clarke, 2013).

Charging content providers for priority placement serves as a viable revenue stream but also reduces credibility (Bhargava & Feng, 2002). Web search engine users interpret sponsored links as less relevant than organic links and are less likely to click them, but businesses rely on these sponsored links to promote their industries and search engines rely on this concept for their principal revenue source (Jansen, Brown, & Resnick, 2007). Bhargava and Feng (2002) used mathematical proofs to discover that there is an ideal level that exists between placement revenues and disutility. Improvements in service and marketing can counter this perceived disutility for paid placement (Bhargava & Feng, 2002). It is important to remember, though, that this study took place in 2002, and our appreciation of online ads may have changed since then. It is clear, however, that users react negatively when they find out pages are sponsored, as if they have been duped (Jansen, Brown, & Resnick, 2007). Jansen, Brown, and Resnick (2007) used two different universities to generalize their results, finding that decreased transparency negatively impacts credibility.

Search engines return their results based on a complex system of algorithms, which differ slightly across each search engine. On average, Google's algorithms change once per day (Carter, 2011). Sherman and Price (2003) suggest that this is an indicator that search engines are actively attempting to improve themselves through more advanced recognition capabilities, for instance, detecting pictures that exist without contextual clues in text (Sherman & Price, 2003). Carter (2011) suggests that this constant shift in algorithms makes it difficult for users to truly judge what returned sites are relevant (Carter, 2011). Although ranking on a SERP strongly

influences website selection, (Carter, 2011; Fallows, 2005; Pan et al. 2007), and users are less critical of websites that are ranked first by Google (Carter, 2011), this does not mean that the sites that show up in the results are actually beneficial. As Carter (2011) says, "There is no evidence that top-ranked sites are perceived as more credible or relevant than lower-ranked sites once users have viewed the sites' content." (p. 24).

Studies show that the average web searcher has "little understanding of how search engines retrieve, rank or prioritize links on the results page" (Jansen, Brown, & Resnick, 2007). Not only do few people understand how search engines develop relevant results, but users are also in the dark as to what they are not uncovering. There is such a thing as the "invisible web," which includes pages that are not easily accessible either through genuine obfuscation or because they require specialized search tools such as LexisNexis (Sherman & Price, 2003). The invisible web exists to some extent because search engines do not carry the content of every page on the Internet (Sherman & Price, 2003). The level of invisibility online ranges from simply opaque websites, to the private web, proprietary sites, and truly invisible websites that exist in file formats that are difficult to index (Sherman & Price, 2003).

A search engine's inability to detect the invisible web is not necessarily malicious - sometimes the information is ephemeral and does not need to be continuously indexed, such as current airline arrival times or weather changes (Sherman & Price, 2003). Search engines also rely on crawlers to search through links and index new sites. If a page is not easily crawled due to its nature or composition of media and file types, it will not turn up in search engine results (Sherman & Price, 2003). Since search engines primarily index pages in HyperText Markup Language (HTML), if a site exists in another language, such as one that cannot be typed into a

URL, it will become disconnected, and thus functionally invisible (Sherman & Price, 2003). The most invisible information, however, is locked in large online databases that contain far too many pages to be indexed (Sherman & Price, 2003). In addition to that, Google in particular controls the flow of information in discrete ways, such as when it quietly deleted hundreds of controversial sites from its results, such as Nazi-themed sites from its French equivalent, Google.fr (Cohen-Almagor, 2011).

Some researchers suggest that all of this makes research more difficult, not easier in the digital age. Dubicki (2010) quotes Head and Eisenberg (2009), saying students are confused by the information search process despite the "convenience, relative ease, or ubiquity of the Internet" (Dubicki, 2010 p. 361). While exploring health information websites, Hong (2006) found that there was a limited number of sites that people would check or be willing to check during a search, and users typically have a specific site in mind when searching. Dubicki (2010) goes on to suggest that students often do not critically evaluate the great amount of information that comes up from a simple search query, preferring to retrieve information quickly, although these students then admit having difficulty finding what they need. The issue, according to Dubicki (2010), is that students do not know how to research any other way. Students have access to a vast amount of information, but without an ability to sort through it, they consume information they know to be faulty.

Credibility

Credibility has been studied since Aristotle's day. Aristotle defined the qualities of a credible speaker as ethos, which consists of good sense, good moral character, and goodwill (Whitehead, 1968). Aristotle waxes long about the concept of happiness in his *Nicomachean*

Ethics, arguing that happiness is a final, self-sufficient state and should be the ultimate goal of every good citizen (Aristotle, trans. 2009). In essence, a good world is made from achieving happiness, and for that reason, we should place greater credibility on those whose perceived aim is the mutual achievement of those ends. The concept of ethos has persisted to the modern day and many researchers have attempted to further define what makes a source credible. Hovland, Janis, and Kelly (1953) defined credibility in terms of expertness, trustworthiness, and intention toward the receiver. McCroskey (1966) used Likert scales to define credibility as a combination between authoritativeness and character and then later explored its multidimensionality in scales derived from a combination of previous research and adjectives provided by college students (McCroskey & Young, 1981). In this later study, McCroskey and Young (1981) found that credibility was influenced by sociability, size, extroversion, composure, competence, time, weight, and character.

Kim (2007) separated credibility into expertise (consisting of expertness, competence, qualification, intelligence, and authoritativeness) and trustworthiness (consisting of perceived honesty, sincerity, objectivity, safety, and sagacity). Freeman and Spyridakis (2004) also considered these elements of expertise and trustworthiness, although they only sorted competence with expertise and they judged goodwill, honesty, accountability, objectivity, character, and concern for public welfare to make up trustworthiness. They also added an element of dynamism that impacts credibility (Freeman & Spyridakis, 2004).

With all of this intense study, though, there are still basic elements of credibility that align with Aristotle, namely the idea that goodwill informs credibility. Goodwill is ultimately perceived caring, which McCroskey and Teven (1999) considered to be understanding, empathy,

and responsiveness. Kim (2007) suggested that perceived caring is the source's intent toward the receiver. Within the minds of information receivers, the perceived image of the source is the most important means of persuasion (McCroskey & Teven, 1999). As McCroksey and Teven (1999) say, "Messages are interpreted and evaluated through the filter of the receiver's perceptions of the message's source" (p. 90). Thus, credibility depends more on who is speaking than the message itself.

Still, the place of goodwill within this debate is contentious. This goodwill dimension of credibility was, according to McCroskey and Tevern (1999), disproven by past researchers and dropped as a third dimension of ethos/credibility in favor of "competence" and "trustworthiness." This, however was developed through the factor analytic research of McCroskey and Young (1981), which McCroskey and Teven (1999) did not interpret as meaning that goodwill did not exist, merely that the factor analytic research did not lead to developing a good measure of the dimension. The McCroskey and Teven (1999) study was intended to prove that perceived caring/goodwill can be measured and is highly associated with other ethos/credibility measures, competence and trustworthiness.

Kim (2007) brings up the same concern as McCroskey and Teven (1999) - that goodwill has been ignored and contends that it is equal to other factors. Kim adapted the McCroskey & Teven (1999) scales to a community, rather than individual (Kim 2007). The study found that trustworthiness is associated with active cognition and goodwill is more based on levels of affection (Kim 2007).

Strong goodwill and public relations will bring a community to the side of a company, even during crisis situations (Kim 2007). This may also be considered a facet of a company's

reputation. Meyer, Marchionni, and Thorson (2010) paired a dimension of believability (the presence of factual and accurate information) with a community affiliation dimension that acknowledged that people are concerned mainly with how the source is geared towards the community's interest. Kim (2007) also echoed this high community service position. In the same vein, Meyer, Marchionni, and Thorson (2010) found that the most credible news is transparent and collaborative, meaning it should appear without bias or opinion and be conveyed as if the writer was working with other people to create it.

Hong (2006) measured the impact of expertise, goodwill, trustworthiness, depth, and fairness on online credibility and found that all but goodwill and fairness have significant relationships with intention to revisit a website. Despite this, goodwill was significantly associated with reliance on a website, which suggested that reliance alone does not predict intention to revisit (Hong, 2006). It is important to note that this research only dealt with health-information searching, although Hong (2006) used natural settings with freeform searching.

Internet credibility. While the credibility of speakers has been examined for thousands of years, the study of credibility on the Internet is relatively new. The work from Aristotle to McCroskey focused more on how audiences perceive credibility exuded from the speaker. The Internet is simultaneously visual, textual, and interactive, so more contemporary work such as Fogg (2002a), Flanagin and Metzger (2007), and Iding, Crosby, Auernheimer, and Klemm (2009) centers around the judgment of credibility that takes place in the mind of the audience. The Internet is such an interactive medium that the audience has become the focus of research of many contemporary scholars.

There are websites that contain misinformation and others that are more accurate, but every one may be trusted equally. Users form judgments about websites based on perceived trustworthiness and expertise to determine both the quality and utility of the site's information (Iding, Crosby, Auernheimer, & Klemm, 2009; Rains & Karmikel, 2009). These judgments may stem from domain names used, aesthetics, and user verification of content, stemming either from the presence of third-party sources or the user's own knowledge. Any accuracy, though, is ephemeral, and may shift dramatically at any moment without warning to users. People may navigate the Internet freely and while doing so, with minimal effort, contribute new information by publishing their own ideas on new websites or by altering information contributed by other users on existing websites. This malleable nature that allows for continuous changes in content is disastrous for people seeking consistently verified information sources. Users alone must make credibility judgments in order to filter between good and faulty information sources. How, then, do people decide what information is believable during their Internet searches?

There is a level of caution when using the Internet, part of which stems from a basic distrust. Casalo, Flavian, and Guinaliu (2007) found that a level of uncertainty exists concerning Internet transactions due to the lack of physical interaction with sellers as well as the reversed purchasing practice of giving out credit card information before the product is received. This level of suspicion on the Internet bleeds over to a distrust of sponsored links, which are perceived as demonstrating a bias on the part of search engines, generating an absence of trust in the minds of users (Jansen, Brown, & Resnick, 2007).

With search engines as gatekeepers of information, access to the Internet is "essentially controlled by a handful of companies and their advertisers" (Lamb, 2004). While search engines

are immensely helpful in sorting out information, allowing users to search with self-defined terms and pre-determined relevancy, there is little understanding towards how this relevancy is measured (Palfrey & Gasser, 2008). Additionally, there is some evidence that the technical limitations of search engines systematically give prominence to a handful of websites at the expense of others based on well-established but publically obscured backlink and PageRank systems (Introna & Nissenbaum, 2000). The Pew Internet project tested many measures of user search engine perception. They found that most adults have faith in the fairness and accuracy of search results, although this faith decreases with age (Pew Internet, 2012). This confidence, however, has increased since 2004 (Pew Internet, 2012).

Prominence-interpretation theory. There are important differences between making judgments on the Internet and the older research on ethos. Fogg (2003) developed the Stanford Website Credibility project that suggested a prominence-interpretation theory to explain users' credibility assessments of websites. Prominence-interpretation theory suggests that users must first notice something online (since they discover it on their own, it must be easily accessed and visible) and then make an interpretation of the material (Fogg, 2003). Fogg (2003) defines interpretations as a combination of the assumptions of the user, his or her existing skills and knowledge, and the context of the user's search. This means that credibility judgments can no longer be interpreted along a one-way path between the speaker and audience. Determining online credibility involves assessing a complex interaction between the audience, the information source, and the system itself.

Since Fogg, other researchers have further attempted to define these two factors of prominence and interpretation, which essentially account for the constant cycle of information

between the user and system that takes place during an Internet search. Schwarz and Morris (2011) defined Fogg's prominence-interpretation theory as an on-going online information search cycle where prominence is influenced by user involvement, user tasks, and experiences; and interpretation is made from assumptions, knowledge level, and context. Users constantly receive messages, interpret them, make judgments, and then perform another search. Flanagan and Metzger (2007) notably assessed this by researching attributes of credibility attached to the sponsor, message, and the site itself. They used the scales developed by McCroskey (1966) and adapted them to web use, finding a range of credibility between news media sites (high) and personal sites (low).

There is no system on the Internet where prominence-interpretation theory fits more than search engines. Search results are analyzed subconsciously in chunks of information, usually within two seconds after the eye hits the screen (Fox, 2012). Users then undergo a process of narrowing or refining queries based on interaction with the system (Fox, 2012). They look at what is prominent, interact with it, create interpretations, and move on. Hotchkiss (2004) studied these search behaviors and interaction with search engines and investigated what generated clickthroughs, including titles, abstracts, product information, as well as trusted sources, brand names, and URLs. He found many differences between the searching habits and characteristics of researchers versus buyers, men versus women, and quick scanners versus two step scanners (those who only skim headlines and move on vs. those who skim once then look back again for greater depth, Hotchkiss, 2004).

Freeman & Spyridakis (2003) discussed two routes to persuasion originally identified by Petty and Cacioppo (1986): central and periphery. The central route consists of a conscious

cognitive effort where users are motivated to evaluate their message processing, using personal relevance and topic knowledge (Freeman & Spyridakis, 2004). The peripheral route, however, is formed from external cues that users instinctively make based on simple judgments without evaluating the message, such as reputation or aesthetics (Freeman & Spyridakis, 2004). Peripheral judgments of credibility occur more in scenarios where users experience an overload of information and must make quick judgments (Freeman & Spyridakis, 2004). When using a search engine and skimming websites, more judgments become peripheral, although this varies based on the depth of information searched.

Presumed and earned credibility. Fogg (2002b) also defined four forms of credibility that have been accepted by Schwarz and Morris (2011), and many others. This includes presumed credibility, which is based on assumptions users make in their minds (such as the inherent trust in domains ending in .edu or .gov); surface credibility, which is based on first impressions of a site, primarily via its design and perceived professionalism; earned credibility, where trust is established over time, typically from high usability and consistently high quality information; and reputed credibility, which stems from third party opinions, including awards or certificates (Fogg 2002b). In this sense, presumed and surface credibility could be considered peripheral judgments, and earned and reputed credibility could be considered central judgments.

Presumed and earned credibility relate to goodwill. Rains and Karmikel (2009) also found that something as basic as the domain name indicates whether it is well known and believed to be in the best interest of the user. For example, people generally trust .edu sites because educational interests are viewed more positively more than commercial interests (Iding, Crosby, Auernheimer, & Klemm, 2009). If people believe that the site can actually help them -

that it is being actively run by an interested and engaged human being - they will have a greater tendency to believe the material presented.

Schwarz and Morris (2011) sought to define earned credibility through dwell time and revisits. Other apparent commercial interests can interfere with earned credibility. Banners or pop-up ads can create negative signals that the website's organization has ulterior motives in its interaction with users (Iding, Crosby, Auernheimer, & Klemm, 2009). Privacy policies (Iding, Crosby, Auernheimer, & Klemm, 2009; Rains & Karmikel, 2009), third-party endorsements, and the presence of a physical address all positively influence credibility (Rains & Karmikel, 2009). Iding, Crosby, Auernheimer, & Klemm (2009) also found that when interviewed, people will list privacy policies as important, but in practice if the visual design is captivating enough, they will not notice the absence of such policies. In terms of search engines, users generally believe that although they have positive experiences with search engines, they do not support search engines tracking personal information or targeted advertising (Pew Internet, 2012).

Reputed and surface credibility. While presumed and earned credibility judgments are user-centric, to some degree websites may control their own reputed and surface credibility. Reputation refers to the brand equity or the customer-perceived credibility of the organization (Casalo, Flavian, and Guinaliu, 2007). This construct is formed from customer-perceived product quality compared to alternatives, as well as the gap between what a company promises and what customers believe it can deliver (Casalo, Flavian, and Guinaliu, 2007). This variance causes reputation to be very unstable - it can change many times within a company's lifespan (Casalo, Flavian, and Guinaliu, 2007). Negative actions have a greater impact than positive ones, and as a result, good reputations are difficult to achieve and sustain (Casalo, Flavian, and Guinaliu, 2007).

Schwarz and Morris (2011) identified reputed credibility elements such as awards, PageRank, and the frequency of shares (Schwarz & Morris, 2011). Baildon and Damico (2011) also found that users rely on other elements of the system to corroborate information. This may lead to selective exposure where researchers lean towards sources that already align with pre-existing beliefs (Baildon & Damico, 2011). They then become trapped in a constant cycle, validating sites based on their validations of other sites (Baildon & Damico, 2011). Freeman and Spryridakis (2003) determined that familiarity or interest could influence credibility. In the case of online credibility this could entail familiarity with the subject matter or familiarity using the Internet itself, which results in users judging information less critically (Freeman & Spyridakis, 2004). Casalo, Flavian, and Guinaliu (2007) separated reputation into competence (how much a consumer believes the seller has the knowledge and skills to satisfy their needs), honesty (the consumer belief that the other party "will keep their word, fulfill promises, and be sincere" [p.587]), and benevolence (consumer belief that one party is looking out for the best interest of the other). Many of these constructs share similarities with goodwill.

Surface credibility deals mainly with aesthetics, although there are other components as well. Sites that use modern conventions in layout and newer designs have been found to be more credible (Flanagin & Metzger, 2007). How sites structure their information also influences user judgments of credibility. Coherent layouts allow users to understand what the website's focus is and how deeply the topics are covered (Iding, Crosby, Auernheimer, & Klemm, 2009). Structural elements such as the presence of navigation menus and links to external websites help create this coherent structure (Rains & Karmikel, 2009).

In addition to these aesthetic elements, Schwarz and Morris (2011) identified on-page surface credibility elements such as the domain type, the presence of advertising, and possible spelling errors. Adding to this, Freeman and Spyridakis (2003) used actual websites, not mock-ups, to find that users typically only judge what is on a single page to determine the page's credibility, without investigating an additional "about us" section that may further explain an organization or user's credentials. Accurate and current information also assures customers that the website is frequently updated, which builds confidence, trust, and loyalty in customers (Vrontis, Ktoridou, & Melanthiou, 2006).

Earlier studies by Fogg (2002a) found that everything from the presence of accuracy verification tools, physical addresses, easily accessed contact information, updated content, and a lack of errors can affect surface credibility (Fogg, 2002a). Rains and Karmikel (2009) again used a Likert scale to assess the perceptions of student participants after searching websites for answers to health care questions. Their results focused on structural elements of the website, such as the presence of navigation menus, links to external websites, as well as content elements such as facts, quotes, and proof of authorship that increase credibility (Rains & Karmikel, 2009). Iding, Crosby, Auernheimer, and Klemm (2009) also found that people are influenced by how information is structured on the site, what the focus is on, and how extensive it appears to be. Martin and Johnson (2010) performed a content analysis of public relations blogs, using comments posted online under each entry to gage the responses of users. They found that posted author credentials, interactive features such as an ability to subscribe to future posts, and dynamic visual elements such as photos and graphics positively impacted visual credibility (Martin & Johnson, 2010).

Fogg's credibility concepts may be applied to many other research studies that generally separate into two tracks: those concerned with what the audience believes going into the search (presumed credibility), and those concerned with what messages the Internet offers them (reputed credibility). Iding, Crosby, Auernheimer, & Klemm (2009) conducted two studies designed around investigating audience credibility perceptions. One group of subjects was provided sites by the researcher to critique and another group was allowed to find sites on their own that they believed to contain accurate and inaccurate information. The researchers found that users who do not consider themselves experts on a given subject check multiple websites when searching for the same content in order to verify the information they need (Iding, Crosby, Auernheimer, & Klemm, 2009). Information from a single website that contradicts a user's previously believed knowledge detracts from its credibility. Often this can be remedied with the presence of an unbiased third-party source or additional facts to back up information (Iding, Crosby, Auernheimer, & Klemm, 2009). The ultimate judgment of credibility lies not with the distributors of new media content but with its consumers (Flanagin & Metzger, 2007). The perceived expertise of a user coming into a website greatly affects credibility (Flanagin & Metzger, 2007; Iding, Crosby, Auernheimer, & Klemm, 2009).

Schwarz and Morris (2011) also found elements that impact credibility on websites such as general popularity, geographic reach, and expert opinions. They accomplished this through ranking 1,000 websites and then gathering user credibility ratings from participants in a laboratory experiment (Schwarz & Morris, 2011). This research has many implications for search engines, most notably the presence of advertising and PageRank in credibility judgments. Hargittai (2010) found that users assign high credibility to those sites that show up first in search

engine results pages, even though that has more to do with keywords and SEO and not factual accuracy. This underlies a fundamental misunderstanding of how search engines work – that many people ignore the fact that search engines may return biased or exclusionary results (Introna & Nissenbaum, 2000), and the most "relevant" results are often merely the ones that are most easily indexed by search engines, regardless of the actual content (Batten, 2008). Schwarz and Morris (2011) found that visualizations of this kind of data on the search engine page itself can actually improve credibility judgments in users while simultaneously adding value to search engines.

In summation, credibility factors regarding the content of online material is well defined despite its relatively brief history. In an age where many people are transitioning their primary means of obtaining information to online sources it is important to understand how this information is determined to be true. People can become better judges of information authenticity by understanding what aspects of websites render their content believable. In addition to consumers, though, web content creators need to be able to present themselves in credible ways. Search engine creators also need to be able to construct their designs and reputations around believability. The research in this study investigated how search engines' reputations may interact with aesthetic factors to enhance credibility. In essence, this is a distillation of presumed, earned, reputed, and surface credibility variables. This project suggests a framework for the balance of these variables and their impact on comprehending and believing information on the Internet.

Due to the now ubiquitous use of search engines, their role as information gatekeepers, and the lack of proven measures to attest how users make credibility judgments of search engine

results, this project will attempt to fill a gap in research by asking the following three research questions:

RQ1: What differences are there in user perception of credibility between information obtained through search engines with established reputations and the credibility of information obtained through search engines without established reputations?

RQ2: What differences are there in user perception of credibility between information obtained through search engines with good aesthetics and the credibility of information obtained through search engines with poor aesthetics?

RQ3: What do people think of the goodwill of search engines?

Based on the literature and the research questions presented, the researcher investigated the following hypotheses:

H1: The perceived reputation of a search engine is related to the credibility of its results.

H1-1: The perceived reputation of a search engine is related to the perceived expertise of its results.

H1-2: The perceived reputation of a search engine is related to the perceived goodwill of its results.

H1-3: The perceived reputation of a search engine is related to the perceived trustworthiness of its results.

H2: The perceived aesthetics of a search engine is related to the credibility of its results.

H2-1: The perceived aesthetics of a search engine is related to the perceived expertise of its results.

H2-2: The perceived aesthetics of a search engine is related to the perceived goodwill of its results.

H2-3: The perceived aesthetics of a search engine is related to the perceived trustworthiness of its results.

H3: There is a relationship between the interaction between perceived reputation and aesthetics of a search engine and the perceived credibility of its results.

Method

This study used methods similar to Dubicki (2010), which included a combination of qualitative and quantitative measures via a survey distributed online as well as observational sessions with undergraduate communications students. This study sought to improve on Dubicki's (2010) design through the inclusion of a ranking system for each item. The purpose of the mixed methods design was to add to the complexity of the findings. Qualitative data alone is limited by interviewer bias and a focus on the individual, not a whole culture (Kawulich, 2005). Quantitative data is limited by a lack of depth, exploration, or flexibility that cannot address context (Tewksbury, 2009). Using a mixed methods approach, the researcher compensated for the weaknesses of both methods and by detecting similarities across both methods, depth and context was added to a broad and unbiased sample. It helped portray a wider perspective and illuminate intangible elements in the search process. The research was approved to involve human subjects by the Institutional Review Board of the Rochester Institute of Technology. Participants of both the online survey and the qualitative interviews were provided consent forms to participate and be observed, respectively. The researcher operationalized the variables, credibility, reputation, and aesthetics through the following:

Credibility

Judging credibility is primarily an analysis of the believability of content. This is achieved through an internal comparison of previously known facts with new information. This study used the four factors developed from Johnson, Kaye, Birchard, and Wong (2007): fairness, believability, accuracy, and depth of information. That study, however, did not account for the fact that these factors have varying levels of importance to different users. Therefore, a measure was added through additional questions asking participants to rank which elements were most important to the subject.

Added to this credibility definition was goodwill, which has been a contentious element of credibility. This study sought in particular to define the role of goodwill. McCroskey and Teven (1999) defined goodwill as understanding, empathy, and responsiveness. They further developed a perceived caring model based on Koehn and Crowell (1996) and Teven and McCroskey (1997) that used a Likert-type scale to measure competence, goodwill, and trustworthiness. These scales were adapted by Kim (2007) in a study of goodwill and trustworthiness in corporate blogs designed for public relations. The current study mimicked this adaptation for its own purposes. It integrated items adapted from Flanagin and Metzger (2007) such as trustworthy, believable, reliable, authoritative, honest, safe, accurate, valuable, attractive, pleasant, colorful, aggressive, bold, interactive, biased, and organized. The three elements of credibility the study focused on were expertise, goodwill, and trustworthiness.

Reputation

There are a handful of other studies that have found means to determine how subjects perceive reputation online. This current study adapted questions used by Casalo, Flavian, and

Guinaliu (2007) that explored usability and reputation affecting online banking. Their initial item set was proposed from a literature review, qualified by a panel of experts, and then validated for reliability and dimensionality via Cronbach's alpha indicators, principal components analysis, confirmatory factor analysis, composite reliability, and construct validity. Some of the questions would not make sense in the context of this study such as "I think that this web site would not do anything intentional that would prejudice the user," and "This web site does not make false statements." Many were modified to make sense when referring to search engines, and some such as these were dropped entirely. The entire survey may be seen in Appendix B.

Aesthetics

The method operationalized aesthetics through measures used in two previous studies. Pollach (2005) defined aesthetics through representations within the interface such as maps, menus, or icons that allowed for efficient navigation. Nathan and Yeow (2009) presented a more direct array of items: graphic, font, and color use. Thus, the researcher measured aesthetics through both a general arrangement of visual elements and specific graphics, font choices, and the interplay of colors. This measure also included questions about the visualization of ads. Every question may be seen in Appendix C.

Procedure

Setting and participants. An online survey was offered to every RIT student in the College of Liberal Arts and every graduate student. The survey was also distributed through listservs, social networking sites, and other e-mail lists. It measured differences in participants' views of the credibility, reputation, and aesthetics of search engines. This was a convenience sample with participants between 18 and 65 years old. Random assignment improved the

generalizability of the results. The demographics varied, but tended towards people who were female (60%), Caucasian (67%), between 18 and 29 years old (30%), and who had graduate degrees (45%). Every major race and ethnicity was represented in close proportion to Monroe County demographics (U.S. Census Bureau, 2013). Respondents were 6% Asian, 13% Black/African-American, 7% Hispanic or Latino, 1% Native Hawaiian or American Indian, and 67% White. Considering that the survey was primarily distributed in a university setting, more specific ages were determined for participants whose ages fell between 18-21 (17%), 22-25 (19%), 26-29 (11%), and 30-34 (5%). After that, ages were determined in ten-year brackets, with 41% over the age of 35. The education level of respondents also reflected the university setting. While 45% held graduate degrees, 17% were current graduate students, 13% had completed undergraduate degrees, and 23% were current undergraduate students. More demographic information may be found in Figures A5 - 8.

The qualitative data was collected primarily from undergraduate and graduate students pulled from communication classes at RIT. They were invited through e-mail, social media, or through professors to participate in one-on-one observation sessions in a computer lab in the Liberal Arts Building. No population was intentionally excluded. As an incentive, each participant was entered into a drawing for a $100 VISA gift card. Students logged in to the computers in the Liberal Arts Building PC lab one at a time using their RIT user names. The Internet history for each browser was wiped clean before use, so students would not be prompted by earlier clicks or search terms.

The researcher encouraged the students to use the search engines in order to discover more about a topic of their own choosing, from academic research to any topic of interest. This

achieved a natural searching environment and gave the students incentive to participate at minimal inconvenience. It also provided the students with topics that they had working knowledge of, so they could accurately discern between credible and suspicious results. The students were technologically competent, meaning they quickly adapted to the session without the need for lengthy explanations of how search engines work and how to use them.

Data sources. Interviews were conducted with 22 students for ten minutes apiece. The researcher devised an ethnographic interview framework that followed the techniques of Spradley (1979). The researcher composed a list of 12 guiding questions during the search process in order to create a dialogue with the subject. Based on the natural flow of the interviews, which allowed the participants to react and ask questions during the search process, not every student answered every question. The researcher followed methods for good observational sessions proposed by Kawulich (2005), such as establishing a rapport through light jokes, dressing neutrally, being familiar with the setting, paying attention to and writing down keywords, shifting perspectives, and being tolerant. Seventy-four total thematic categories were created based on answers to the guiding questions. The goal of the questions was to find out where students were looking on search engines, how they evaluated results, and what caused them to click where they did. Other questions attested basic habits such as whether students would be more likely to move on to new pages or modify their queries if they did not discover what they were originally looking for.

Audio from each session was digitally recorded and then logged for analysis in an Excel spreadsheet under a random six-digit number generated from the same spreadsheet. This made for 22 ten minute recordings and 22 spreadsheets. To ensure anonymity and confidentiality there

were no other identifying characteristics besides the random number assigned on the spreadsheet and the recordings. In order to diversify the subjects' experiences, the search engines the students were instructed to use were also determined from this random number (numbers ending in 1-3: Google, numbers ending in 4-6: Yahoo, numbers ending in 7-9: Bing, Numbers ending in 0: re-generated randomly).

Students were not instructed to pay attention to any of the variables from the outset in order to generate a genuine reaction to the search process and to attest whether each variable is prominent over any other. After each student performed their searches, they were invited to take the same online survey link as the quantitative participants.

Data analysis. The researcher took notes during the live interviews along with the audio recordings to fill in gaps on a multi-page spreadsheet. Each spreadsheet was then cross-tabulated to find significant similarities and differences between subjects. The data was then analyzed for recurring patterns and underlying themes in line with Kawulich (2005) and Spradley (1979) to form categories of information in regards to each guiding question. These categories were transcribed on a separate spreadsheet that aggregated the data. Similar responses were coded together with a tally mark. If a response differed, it was given a new column. After all the data were assessed, the tally marks were replaced with numerals to indicate how often they appeared. After this, the researcher looked for trends and correlations in an attempt to rationalize the results. There were a total of 204 data units within the 74 distinct categories. This data is displayed in Table A7. An intercoder reliability analysis was also completed. After being trained by the researcher, a second coder analyzed 11% of the data units, which resulted in a 91% reliability rate.

Results

Quantitative Results

A total of 192 participants took an online survey consisting of 58 items, which resulted in 11,136 possible responses. Of these, 742 responses were answered with "not applicable" statements (6.7%) and 279 responses were left completely blank (2.5%), leaving 9.2% of responses missing. The researcher deleted the participants' entire response when the participant was missing over 10% of data (over 6 responses that were either blank or "not applicable"). This resulted in 141 usable responses. This was in order to more accurately impute missing data from each respondent (Roth & Swtizer, 1999; Rubin, 1996; Schafer, 1999).

Imputations were appropriate, because the data were missing at random (Shafer, 1999) and only 9.95% of the data were missing, which is under acceptable limitations (Roth & Switzer, 1999). The imputations were created by averaging each individual participant's responses for each section so that each section's overall average did not change, but the missing response could still be measurable. The usable data was then analyzed using SPSS for internal reliability, Pearson's *r* correlations, and linear regressions. Analyses were also run without imputations, which generally found similar correlations, but with a much smaller sample size. Trustworthiness in particular was found to be strongly and significantly correlated with reputation, but this was from only 52 usable respondents.

Expertise, goodwill, and trustworthiness were each composed of eight items. To determine internal reliability, Cronbach's Alpha was found for each. Cronbach's Alpha was also found for the 21 items that made up the reputation measure and the 13 items within aesthetics. These results may be found in Table A1.

Pearson's *r* correlations were run between expertise and reputation, expertise and aesthetics, goodwill and reputation, goodwill and aesthetics, trustworthiness and reputation, and trustworthiness and aesthetics. The strongest correlation between an individual item and reputation was trustworthiness at .572. Reputation and aesthetics had a marginally higher correlation with each other, but their relationship was not under investigation in this research project. More in-depth results can be seen in Table A2.

We thus see support for expertise measured against reputation (*H1-1*). There is also some support for goodwill against reputation (*H1-2*), and trustworthiness against reputation (*H1-3*). Each relationship was moderately positive, although *H1-1* and *H1-3* were stronger than *H1-2*, which can closely be considered moderate. There is also support for each measure of credibility against aesthetics (*H2-1*, *H2-2*, and *H2-3*). Each relationship was again positive and may be considered a moderate correlation, but the correlation was lesser than for reputation.

Linear regressions were run for the averages of expertise, goodwill, and trustworthiness against three models. These models were reputation alone, reputation with aesthetics, and finally, reputation, aesthetics, and a reputation/aesthetics interaction term. The interaction term was devised through multiplying the average results for reputation with the average results for aesthetics. These models may be found in Tables A3 - 5. Thus we have some support for *H3*, although the regression was strongest for expertise.

The study also sought to quantify in some way not only what users perceived of each element of credibility, but also how important they thought each element was. In terms of search engines, users believed that the most important expertise item was accuracy; in terms of

goodwill, search engines should be ethical, and for trustworthiness, safety was the top priority. The complete results of where each user ranked each item may be seen in Figures A1 - 3.

Qualitative Results

The live observational interviews uncovered many search habits in addition to the research questions. On the basic need for search engine use, one student commented, "If I don't know where to go, that's where I'll go." At the heart of the qualitative study was a search for how students choose which websites to click on based on their initial search query. Students looked at many elements on the SERP to make this judgment. Many looked at the headlines of pages first, in particular if they matched keywords. Second, they looked at URLs, although only one student cited the danger of clicking nefarious URLs that could be full of strange HTML characters.

Dates and pictures also had large influences, which built on the effort by Vrontis, Ktoridou, and Melanthiou (2006) that suggested that the appearance of more current information builds trust. As one student said, "I know that a good website contains meta data, such as a date." Students cited reading abstracts least of all on the SERP, although some noted that if the information search is important, reading abstracts is essential:

> If I were searching for more academic-type things I would probably use the abstract first because I would get more of a feel for specific information.
>
> I read the abstract if I'm looking for specific terms, like if I'm trying to cheat on my homework. It depends on the context.

Others noted that there is a subconscious need for pages to still have abstracts, because it immediately shows that the page actually contains content:

> I think it would change things a bit [if there was no abstract], definitely, because I do always read a little bit to see what I am going to be reading.
>
> It almost gives an illusion of more material or content, even if it's not actually read. You really don't think about it, but it's true.

These students commented on how it would be strange if the abstracts were absent but one student commented that he is getting used to searching on his phone, so the absence is less jarring. "I use my phone a lot for Google searches and I'm pretty sure on the phone it just shows this part [the headline] and not even the website, and I still find what I'm looking for."

Most users make their judgments very quickly, after a quick skim, which aligns with the work of Fox (2012). They examine the top results first and rarely make it beyond the third page. In fact, only one respondent claimed to typically go all the way to the fourth page. In their own words:

> I've never not found what I've needed on the first few pages.
>
> I would rather do another search than go through all these pages.

Another habit cited by numerous subjects was opening up many new tabs to look at varying web pages in greater detail later, or to seek more pages if more questions arose. As a student commented:

> I generally, whenever I browse for something, I just middle-click on a bunch of stuff and I just read the headlines. I do the center click and open a bunch of tabs and go into the tabs and if I see something I might be interested in, I leave it open and go through it. If not, I close it.

During the search process, students are constantly attempting to match the image in their head with that on the page. As noted, they were much more likely to modify their query rather than move on to the next site in order to do this. This is essentially the heart of prominence-interpretation theory (Fogg, 2003).

RQ1: What differences are there in user perception of credibility between information obtained through search engines with established reputations and the credibility of information obtained through search engines without established reputations? Corresponding Categories: Reputation, Brand identity, Total trust, Familiar URLs, Previous Experiences

In order to make their belief judgments, students highly cited previous knowledge and third party sources, along with minor factors such as brand identities, consistency, and recency. This confirms previous findings by Iding, Crosby, Auernheimer, and Klemm, (2009) as well as Vrontis, Ktoridou, and Melanthiou (2006). The highest factor for determining quality information, though, was reputation. As the students described their processes:

> I have no idea what that is. I'm not clicking it.
>
> I don't want to waste my time clicking a link if I know I'm not going to trust it.
>
> If it's something that I recognize and find trustworthy, it's good.
>
> I would definitely look at the URL because I have never heard of this company. Brand identity, brands that I know are good. This is getting towards the bottom of the page, though, so I am not likely to click on it.

This trust of search engines is built upon both previous successful experiences and convenience.

Participants cited many other sources of information retrieval such as using peers, libraries, databases, academic institutions, online forums, online databases, social networks,

going to websites directly (bypassing search engines), and *Wikipedia*. A relatively small number of participants used Google Scholar, but many used Google itself as an information search for academic projects, although it quickly became clear that the kind of search being performed had a high impact on how users analyze results. For instance, when seeking cursory knowledge, there is a level of acceptable error for searches of low importance. Many students cited a need for deeper searches for topics that required a higher level of rigor, such as academic research.

Another student cited how she forms her trust judgments very succinctly, "Well, Google. If it's Google, yes." As long as the source was Google, she believed in the results. This student fell in line with Bhargava and Feng's (2002) finding that perceived high service on the part of a search engine can compensate for the inconvenience of product placement. Based solely on Google's reputation and brand she bought into the information presented.

RQ2: What differences are there in user perception of credibility between information obtained through search engines with good aesthetics and the credibility of information obtained through search engines with poor aesthetics? Corresponding Categories: Aesthetics, Pictures

Aesthetics were notably absent in many of the responses, but one student did comment that the look of a website had an effect on her credibility judgment. "Some of the formats show how old it is. If it is a very 'plain Jane' website, obviously the person who made it is not that proficient, and I won't give them as much credit." Another commented that she makes these judgments very quickly; as soon as she clicks on a site she "…can make split-second judgments." Without prompts, others did not cite aesthetics as having an impact on which search results they click on. This was possibly because Google, Yahoo, and Bing all currently use

aesthetically similar formats and there is effectively no need to differentiate between high and low aesthetics.

Many students did, however, mention that they would be more likely to click a link or search result that contained a compelling picture:

> Pictures, I don't know, I could give or take. This one isn't very helpful. I'm not sure who that is. From time to time I can point something out, though. [sees a picture of a building]. Oh that's a nice building.
>
> I always go for visuals.

RQ3: What do people think of the goodwill of search engines? Corresponding Categories: Garbled text, It knows what you want, Skeptical of Top Spots, Convenience, First Page/Top More Trustworthy, New Info – Self Doubt, Why not trust?, Wary of Ads, Numb to Ads, Acceptable error

Many students said that the results near the top of a SERP were the most trustworthy because they had faith in search engines. A nearly equal number, though, claimed to be skeptical of top spots because they believed companies paid for those positions:

> I think about what they're doing on the business side - if larger websites that are being driven by more powerful and financially stable companies are paying off search engines for the top spot in the search, so I always have that in the back of my mind but from a convenience standpoint, If I'm on the run I'm not really going to debate if I'm being spoonfed.
>
> Ads can be helpful if they are relevant, but I don't like how the first thing Google shows is an ad, it should be the first thing you're looking for, not something disguised.

I don't think a sponsored ad has anything to do with what I'm searching for. It makes me think 'What are they trying to sell me?' I would just skip over that.

A commercial can be informative, but because I don't know and I don't want to buy anything, I'm not going to click on it. This one is an edu so I trust it. I'd go to Wikipedia first though. It's easy English to understand, I get what they're saying. They have pictures.

There were also similar numbers of students who were wary of ads as there were students who said they were numb to ads, or did not pay attention to them at all:

They try to push ads on your rather than content, so I don't follow those.

I've gotten used to knowing there will be ads.

Even when search engines had heavy ad presences, it had little effect on how students made their information judgments. Even though Jansen, Brown, and Resnick (2007) found that a high number of ads created a lack of trust on the part of users, the subjects in this study either ignored the ads or merely found them irritating, without detracting from their actual credibility judgments.

Students cited that the search engine in some cases can provide complete thoughts in the search bar before they even know what they want. In one case, it caused a user to doubt herself. The student was searching for more information on a building collapse in India she saw on television, but could not find the exact story that she knew fit all the bits and pieces of information she had. She finally realized the collapse was actually in Bangladesh, but she then thought that “Maybe when I was viewing information earlier I was incorrect.” She believed that the search engine was looking out for her.

Another subject simply stated that there is no reason to not trust a search engine, saying that there is no reason for it to be malicious:

> At the end of the day, the stakes are pretty low. Why not trust it? What do I have to lose? What's going to happen, I'm just going to be wrong? It's such a marginal amount of effort, I have never thought of it from a trust point of view.

At the same time, though, students were wary of misdirection and ads. For instance, one student was searching for shoes and instead of finding more information or reviews, the SERP displayed ads to buy shoes. There is a tremendous amount of difference between information search intent and purchasing intent and the search engine continually reverts to the latter. This recalls Jansen, Brown, and Resnick's (2007) finding that users are less likely to click links that are not organic.

Discussion

The results of the multiple linear regression revealed limited relationships between credibility, reputation, and aesthetics. Reputation was positively related to expertise and trustworthiness, while controlling for aesthetics. According to Model 2 of Table A2, though, this relationship is not shared by aesthetics due to a non-significant coefficient. The same is reflected in Model 2 of Table A4. Because the *R* square values remain virtually unchanged throughout each model, the null finding is also confirmed. As for the interaction term, no coefficient is significant, meaning that the effects of reputation on expertise and trustworthiness were not dependent on aesthetics. The Pearson's *r* correlations of both expertise and trustworthiness were also similar and indicated a moderate relationship.

The concept of goodwill remains contentious. Although a significant coefficient was found for the linear regression equation relating to reputation, no significant coefficients

occurred when controlling for aesthetics or when implementing the interaction term. It also possessed the weakest Pearson's *r* correlation between reputation and aesthetics out of all of the credibility factors (although still considered moderate). From the qualitative data, it is clear that interviewees made their judgments based much more on perceived expertise and trust rather than whether or not their information provider was looking out for their best interests. Although there were repeated categories such as being "wary of ads" and being "skeptical of top spots because of companies paying," there were many more instances of forming trust from "third-party sources," "previous knowledge," and "reputation."

Aesthetics yielded similar but smaller Pearson's *r* correlation results in the same proportion as reputation: goodwill was low, followed by expertise, then trustworthiness, which was higher but closer to expertise. Aesthetics in general had a lower mean than reputation on a 5-point scale (3.73 vs. 3.98). This may be because search engines do not require good aesthetics as much as they need functionality for quick operation and reaction to user needs and desires. Search engines still require high reputations, though, to yield credible results.

The Pearson's *r* correlation was 32% lower for expertise and aesthetics than it was for expertise and reputation. Trustworthiness experienced a similar significant drop between reputation and aesthetics (-29%). For goodwill, though, this drop was less pronounced (-9%). This means that despite the smaller correlation, aesthetics and reputation have a more similar effect on goodwill than they have with either expertise or trustworthiness.

It is also notable that the mean for responses for goodwill was found to be much lower (4.53) than that of either expertise (5.40) or trustworthiness (5.51) on the 7-point scale. This may simply indicate that users do not believe that search engines contain as much goodwill as they do

expertise or trustworthiness. The fact that this does not affect reputation or aesthetics as much as the other two factors aligns with previous research by Kim (2007), Freeman and Spyridakis (2004), Iding, Crosby, Auernheimer, and Klemm (2009), and Rains and Karmikel (2009) that considered these measures to be the strongest credibility elements.

In regards to item rankings, it is clear that in this study, participants found it most important that search engines be accurate, reliable, and organized. Perhaps the fact that users placed less emphasis on information being unbiased, valuable, or believable is an indicator as to why they considered goodwill to be less important. It was beyond the scope of this study to rank all 24 items of credibility, but judging from the relatively higher number of "not applicable" answers for goodwill and trustworthiness items, it was clear that high expertise was vital for a larger number of search engine users, and within this, accuracy was paramount. Low accuracy also created issues with user experiences during the qualitative interviews. At times the results page was cluttered with inappropriate ads or a few rounds of mistaken queries before the correct information was found, which resulted in frustration.

For this study, users ranked the most important aspects of goodwill to be "ethical" and "moral." These are more abstract than the other items, which included "cares about me" and "concerned with me." Some participants in the qualitative portion were actually confused as to how these two items were distinguished, which may be worth some investigation. It is clear that many people were simply looking for baseline ethicality and were not concerned that their information provider actively looked out for their best interests.

Concerning trustworthiness, the highest-ranked item in this study was "safety," followed closely by "honesty." Items like "genuine," "bold," and "honorable" were deemed less

important. As one user in the qualitative portion put it when he found some results that appeared cluttered with keywords and garbled HTML text that lacked content, "That weird expression, the gobbledy-gook is a pitfall. It is probably a phishing attempt." Thus, trustworthiness was captured more by the idea that the search results would either not infect a user's computer or yield inappropriate or X-rated results. This echoes some of the work done by Schwarz and Morris (2011), which found that credibility misjudgments were a significant danger online because of malicious websites. Examining URLs that seem suspicious and actively avoiding those sites also supports the presumed credibility concept of Fogg (2003).

Regarding more general searching behavior, much of the qualitative data reflects previous studies. Students used search engines to try to find specific websites in accordance with Hong (2006): "I use this [searches for a specific website from search engine] because I am used to Google Chrome." Others, however, disagreed, but this could also be due to the nature of the observation, where participants were forced to search without necessarily having strict intentions: "I'm exploring, I don't have a preconceived idea of what I want."

Fox (2012) stated a subconscious discovery of relevancy clues such as a headline and abstract, along with the simple fact that we judge a website by its aesthetics within 50 milliseconds. This was reflected in much of the qualitative data when analyzing search processes. To add more depth to this notion though, this research project sought to find out if it matters whether or not those aesthetics judgments were positive or negative. Could users still make positive credibility judgments from search engines with poor aesthetics? The moderate correlation indicates that this is somewhat the case, but more research would be valuable.

Limitations and Future Research

There is some danger with imputation in that it does not reflect the true feelings of the respondents. There were higher numbers of "not applicable" answers for the items "Moral" (7.07%), "Aggressive" (5.43%), and "Text/Graphic Arrangement" (5.43%). Even though goodwill had the highest number of "not applicable" (4.5%) and missing answers (2.4%), trustworthiness was the measure that had the highest number of separate participants who chose at least one "not applicable" answer or left a response blank (70.1%). In many cases, though, this was scattered at random, representing only one out of 58 total responses. If this study were to be reproduced, however, those items in particular need to be reassessed, and the ability to choose a "not applicable" response should possibly be dropped. Conversely, though, it would be worthwhile to investigate the reasons participants selected this answer further. A larger sample size, which would then be divided into response sets, would prove to have interesting results. More qualitative studies that asked why users selected "not applicable" would also be beneficial.

According to the demographic data, the "not applicable" answers were highly distributed, with no real indicator as to why. A more precise study could actually attempt to ascertain which groups of people may think this way towards search engines - that is, actually test who would believe a search engine's goodwill or trustworthiness was "not applicable." Considering the lack of any demographic indicator, the researcher suggests that a deeper psychographic analysis may be appropriate to understand who is feeling this way about their information searches. In all, this contradicts to some extent the measurability of goodwill proposed by McCroskey and Teven (1999), although this may improve with refinement of the measure.

There are other issues inherent to conducting an online survey, including a lack of control over who is actually taking the survey. According to the reported demographics, the age range was widely distributed. Future studies on the subject could have a narrower target. This study also leaned towards women, Caucasians, and those with advanced education. Future research could be more precise. Nothing indicated that a specific group of people left answers blank. While some did stop filling out the survey entirely, others who left large portions blank continued to answer every demographic question without issue. There was no pattern that indicated that a specific group of people left items blank. The surveys were also logged with times they were taken, which did not elicit any discernible pattern of absence based on time of day or day of the week.

The entire search engine process deserves much more study. From the qualitative interviews it is clear that search intention has a tremendous impact on the rigor of a search as well as the intensity of information judgments. Whether users are shopping, conducting academic research, or just scanning for news has a critical impact on how they interpret information. It would be beneficial for future studies to investigate one of these behaviors instead of searching in general. Some of the discrepancies may also be a result of the fact that many users still have little understanding how search engines actually work, which falls in line with the study of Jansen, Brown, and Resnick (2007):

> No, I have no idea how a search engine really works.
>
> I don’t really know how they work.
>
> I just imagine this giant building where a server rests.
>
> I don’t know if that’s what they do.

In support of this behavior-specific research, there were other demographic questions that yielded interesting results that did not directly impact the study, but are worth mentioning. In the week before they took the survey, 98.90% of participants indicated they had used Google. This is actually much higher than the 83% that the Pew Internet Project (2012) found. Other search engines that were used may be seen on Figure A4. While the Pew Internet Project (2012) also found that faith in search engines decreases with age that was not the case with the participants of this study. This study also found no difference between genders, contradicting Hotchkiss (2004).

The researcher also thought it worthwhile to discover what users actually used search engines for, as this naturally impacts the rigor of the search as well as how critical users are of the information they find. Fox (2012) defined five query types: navigational, commercial, informational, prepurchase research, and action. The most cited reasons for search in the current research project were "general subject knowledge," "images," and "geographic locations," in that order. The least cited reasons were "sports information," and "personal information about other people," in that order. The full array of responses may be seen in Figure A9.

Future studies may explore each of these reasons in further detail to discover how people make their judgments based on the credibility of each source. For instance, would users be more skeptical of sports information found through search engines than geographic locations? Are there other reasons that sports information was ranked lower - for instance, are there more outlets to obtain that information? Or are there just a lower number of people interested in the subject than geography, which affects everyone? Answers to these questions are beyond the scope of the current study, but would be interesting to explore.

The prevalence of *Wikipedia* also often came up in the qualitative interviews. User judgments of *Wikipedia* credibility would also be a worthwhile subject to tackle, considering how prevalent it has become in any kind of research. On a similar subject, with the recent advent of Facebook's Graph Search there are entirely new credibility issues. Graph Search depends wholly on peer credibility not sponsor-credibility. It works under the assumption that users will search their friends' profiles for reviews and information and base credibility judgments on peer opinions. It is one possible direction for the future of online credibility study.

References

Ashkan, A. & Clarke, C.L.A. (2013). Impact of query intent and search context on clickthrough behavior in sponsored search. *Knowledge Information Systems*, 34, 425-452.

Baildon, M. & Damico, J. (2011). Internet sources: developing critical and reflexive readers of complex digital texts. *Social Education*, 75(5), 269-273.

Batten, J. (2008, March 31). Search engine results, blogs, and credibility. Retrieved July 13th, 2013 from http://www.clickz.com/clickz/column/1715803/search-engine-results-blogs-credibility

Bhargava, H. K., & Feng, J. (2002). Paid placement strategies for internet search engines. In *Proceedings of the 11th international conference on World Wide Web*, 117-123. ACM.

Carter, K.J. (2011). *Credibility by Google: do search engine cues influence website credibility and relevance assessments?* (Unpublished master's thesis). The Ohio State University, Columbus, OH.

Casalo, L.V., Flavian, C. & Guinaliu, M. (2007). The role of security, privacy, usability and reputation in the development of online banking. *Online Information Review*, 31(5), 583-603.

Cohen-Almagor, R. (2011). Freedom of expression, internet responsibility, and business ethics: the Yahoo! saga and its implications. *Journal of Business Ethics*, 106, 353-365.

Dubicki, E. (2010). Research behavior patterns of business students. *Reference Services Review*, 38(3), 360-384.

Fallows, D. (2005). Search engine users. Pew Internet & American Life Project. Retrieved July 1st, 2013 from http://www.pewinternet.org/~/media/Files/Reports/2005/PIP_Searchengine_users.pdf.pdf

Flanagin, A.J. & Metzger, M.J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, 9(2), 319-342.

Fogg, B.J. (2002a). *Stanford web credibility research*. Retrieved November 6th, 2012 from http://credibility.stanford.edu/guidelines/index.html

Fogg, B.J. (2002b) Persuasive Technology: Using Computers to Change What We Think and Do. *Morgan Kaufmann*, 2002.

Fogg, B. J. (2003). Prominence-interpretation theory: Explaining how people assess credibility online. In *CHI' 03 extended abstracts on Human factors in computing systems*, 722-723.

Fox, V. (2012). *Marketing in the age of Google*. Hoboken, NJ: John Wiley & Sons, Inc.

Freeman, K.S. & Spyridakis, J.H. (2004). An examination of factors that affect the credibility of online health information. *Technical Communication*, 51(2), 239-263.

Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K.Y. (2010). Trust online: young adults' evaluation of web content. *International Journal of Communication*, 4, 468-494.

Head, A.J., & Eisenberg, M. (2009). What today's college students say about conducting research in the digital age. *Project Information Library Progress Report*, Retrieved July 17th, 2012 from http://www.projectinfolit.org/pdfs/PIL_ProgressReport_2_2009.pdf

Herman, E. & Nicholas, D. (2010). The information enfranchisement of the digital consumer. *Aslib Proceedings: New Information Perspectives*, 62(3), 245-260.

Hong, T. (2006). Contributing factors to the use of health-related websites. *Journal of Health Communication*, 11, 1149-165.

Hotchkiss, G. (2004, June 29). Inside the mind of the searcher, part II: search behavior exposed. *Search Engine Guide*. Retrieved July 16th, 2013 from www.searchengineguide.com/gord-hotchkiss/inside-the-mind-of-the-searcher-part-ii-search-behavior-explored.php

Hovland, C., Janis, I.J., & Kelley, H.H. (1953) *Communication and Persuasion.* New Haven, Connecticut: Yale University Press.

Iding, M.K., Crosby, M.E., Auernheimer, B., & Klemm, E.B. (2009). Web site credibility: why do people believe what they believe? *Instructional Science*, 37, 43-63.

Introna, L.D., & Nissenbaum, H. (2000). Shaping the web: why the politics of search engines matters. *The Information Society*, 16(3), 169-185.

Kim, K. (2007). *Effects of message sidedness, corporate goodwill and trustworthiness in weblogs: an experimental test of alternative public relations message strategies*. Paper presented at the International Communication Association Conference 2007.

Jansen, B.J., Brown, A., & Resnick, M. (2007). Factors relating to the decision to click on a sponsored link. *Decision Support Systems*, 44, 46-59.

Johnson, T. J., Kaye, B.K., Bichard, S.L., & Wong, W.J. (2007). Every blog has its day: politically-interested internet users' perceptions of blog credibility. *Journal of Computer-Mediated Communication*, 13, 100-122.

Kawulich, B.B. (2005). Participant observation as a data collection method. *Forum: qualitative social research*. Retrieved July 24, 2013 from http://www.qualitative-research.net/index.php/fqs/article/view/466/996

Kim, J. (2007). *Effects of message sidedness, corporate goodwill and trustworthiness in Weblogs: An experimental test of alternative public relations message strategies*. Presented at 2007 ICA, San Francisco, CA.

Koehn, S.C. & Crowell, T. (1996). *The development of a perceived caring scale for educators and students*. Presented at the annual convention of the Speech Communication Association, San Diego, CA.

Lamb, G.M. (2004, May 6). The new gatekeepers. *The Christian Science Monitor,* Retrieved July 12th, 2013 from www.csmonitor.com/2004/0506/p14s02-stin.html

Martin, K.N., & Johnson, M.A. (2010). Digital credibility & digital dynamism in public relations blogs. *Visual Communication Quarterly*, 17, 162-174.

McCroskey, J.C. (1966). Scales for the measurement of ethos. *Speech Monographs*, 23, 65-72.

McCroskey, J.C. & Young, T.J. (1981). Ethos and credibility: the construct and its measurement after three decades. *Central States Speech Journal*, 32, 24-32.

McCroskey, J.C. & Teven, J.J. (1999). Goodwill: a reexamination of the construct and its measurement. *Communication Monographs*, 77, 90-103.

Meyer, H.K., Marchionni, D., & Thorson, E. (2010). The journalist behind the news: credibility of straight, collaborative, opinionated, and blogged "news". *American Behavioral Scientist*, 54(2), 100-119.

Nakamura, S., et al. (2007). Trustworthiness analysis of web search results. *Research and advanced technology for digital libraries 11th European Conference proceedings*, 38-49.

Nathan, R.J. & Yeow, P.H.O. (2009). An empirical study of factors affecting the perceived usability of websites for student internet users. *Universal Access in the Information Society*, 8, 165-184.

Palfrey, J. and Gasser, U. (2008). *Born digital: Understanding the first generation of digital* natives. New York: Basic Books.

Pan, B., et al. (2007). In Google we trust: users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication, 12*(3), article 3. Retrieved March 31, 2013 from http://jcmc.indiana.edu/vol12/issue3/pan.html

Pearson, J.M. & Pearson, A.M. (2008). An exploratory study into determining the relative importance of key criteria in web usability: a multi-criteria approach. *The Journal of Computer Information Systems*, 48(4), 115-127.

Pew Internet (2012). *Search Engine Use 2012*. Retrieved July 1st, 2013 from www.pewinternet.org/~/media//Files/Reports/2012/PIP_Search_Engine_Use_2012.pdf

Petty, R.E., & Cacioppo, J.T. (1986). *Communication and persuasion: central and peripheral routes to attitude change*. New York: Springer-Verlag.

Pollach, I. (2005). Corporate self-presentation on the WWW: strategies for enhancing usability, credibility, and utility. *Corporate Communications*, 10(4), 285-301.

Rains, S.A., & Karmikel, C.D. (2009). Health information-seeking and perceptions of website credibility: examining web-use orientation, message characteristics, and structural features of websites. *Computers in Human Behavior*, 25, 544-553.

Roth, P.L. & Switzer III, F.S. (1999). *Missing data: instrument-level huffalumps and item-level woozles*. Academy of Management, Research Methods Division, 4, Retrieved June 27th, 2013 from http://division.aomonline.org/rm/1999_RMD_Forum_Missing_Data.htm

Rubin, D.B. (1996). Multiple imputation after 18+ years. *Journal of the American Statistical Association*, 91(434), 473-489.

Schwarz, J., & Morris, M. (2011). Augmenting web pages and search results to support credibility assessment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1245-1254.

Schafer, J.L. (1999). Multiple imputations: a primer. *Statistical Methods in Medical Research*, 8, 3-15.

Sherman, C. & Price, G. (2003). The invisible web: uncovering sources search engines can't see. *Library Trends*, 52(2), 282-298.

Spradley, J.P. (1979). *The ethnographic interview*. New York: Holt, Rinehart and Winston.

Tarafdar, M., & Zhang, J. (2005a). Analyzing the influence of website design parameters on web site usability. *Information Resources Management Journal*, 18(4), 62-80.

Tarafdar, M., & Zhang, J. (2005b). Analysis of critical website characteristics: a cross-category study of successful websites. *The Journal of Computer Information Systems*, 46(2), 14-24.

Tewksbury, R. (2009). Qualitative versus quantitative methods: understanding why qualitative methods are superior for criminology and criminal justice. *Journal of Theoretical and Philosophical Criminology*, 1(1), 38-58.

Teven, J.J. & McCroskey, J.C. (1997). The relationship of perceived teacher caring with student learning and teacher evaluation. *Communication Education*, 46, 1-9.

U.S. Census Bureau. (2013, June 27). *State & County Quickfacts: Monroe County, N.Y.* Retrieved July 24, 2013, from http://quickfacts.census.gov

Vrontis, D., Ktoridou, D., & Melanthiou, Y. (2006). Website design and development as an effective and efficient promotional tool: a case study in the hotel industry in Cyprus. *Journal of Website Promotion*, 2(3/4), 125-139.

Whitehead, J.L. (1968). Factors of source credibility. The Quarterly Journal of Speech, 54, 59-63.

Appendix A: Tables and Figures

Table A1

Cronbach's Alpha for Reliability

Item	Cronbach's Alpha
Expertise	.882
Goodwill	.811
Trustworthiness	.901
Reputation	.943
Aesthetics	.800

Table A2

Descriptive Statistics and Construct Correlations (N= 141)

Construct	Mean	*SD*	Expertise	Goodwill	Trustworthiness	Reputation	Aesthetics
Expertise	5.40	.98					
Goodwill	4.53	1.02	.288***				
Trustworthiness	5.51	.99	.604***	.506***			
Reputation	3.98	.59	.550***	.356***	.572***		
Aesthetics	3.73	.51	.374***	.324***	.404***	.575***	

Note: SD = Standard Deviation.
*p < .05, **p < .01, ***p < .001.

Table A3

Predicting Expertise from Reputation and Aesthetics (N= 141)

	Model 1		Model 2		Model 3	
	b	Beta	*b*	Beta	*b*	Beta
Constant	1.82 ***		1.53		-2.52	
	(.47)		(.55)		(3.12)	
Reputation	.90***	.55	.82***	.50	1.80*	1.10
	(.12)		(.14)		(.76)	
Aesthetics			.16	.09	1.34	.70
			(.16)		(.91)	
Rep*Aes					-.28	-1.08
					(.21)	
R^2	.30		.31		.32	
Adjusted R^2	.30		.30		.30	
F(df)	60.17 (1,139)		30.54 (2,138)		21.14 (3,137)	
p	< .001		< .001		< .001	

Note: Standard errors are in parentheses; *b* = unstandardized coefficient; Beta = standardized coefficient; Rep*Aes = Reputation*Aesthetics.
*p < .05, **p < .01, ***p < .001.

Table A4

Predicting Goodwill from Reputation and Aesthetics (N= 141)

	Model 1		Model 2		Model 3	
	b	Beta	*b*	Beta	*b*	Beta
Constant	2.09 ***		1.46		5.06	
	(.55)		(.64)		(3.65)	
Reputation	.62***	.36	.44	.25	-.43	-.25
	(.14)		(.17)		(.88)	
Aesthetics			.36	.18	-.69	-.34
			(.19)		(1.06)	
Rep*Aes					.25	.92
					(.25)	
R^2	.13		.15		.16	
Adjusted R^2	.12		.14		.14	
F(df)	20.21 (1,139)		12.00 (2,138)		8.36 (3,137)	
p	< .001		< .001		< .001	

Note: Standard errors are in parentheses; *b* = unstandardized coefficient; Beta = standardized coefficient; Rep*Aes = Reputation*Aesthetics.
*$p < .05$, **$p < .01$, ***$p < .001$.

Table A5

Predicting Trustworthiness from Reputation and Aesthetics (N= 141)

	Model 1		Model 2		Model 3	
	b	Beta	*b*	Beta	*b*	Beta
Constant	1.73 ***		1.35		-2.72	
	(.46)		(.54)		(3.10)	
Reputation	.95***	.57	.84***	.51	1.83	1.10
	(.12)		(.14)		(.75)	
Aesthetics			.22	.11	1.40	.73
			(.16)		(.90)	
Rep*Aes					-.28	-1.08
					(.21)	
R	.33		.34		.34	
Adjusted R^2	.32		.33		.33	
F(df)	67.73 (1,139)		34.89 (2,138)		24.00 (3,137)	
p	< .001		< .001		< .001	

Note: Standard errors are in parentheses; *b* = unstandardized coefficient; Beta = standardized coefficient; Rep*Aes = Reputation*Aesthetics.
*$p < .05$, **$p < .01$, ***$p < .001$.

Table A6

Qualitative Categories

Question	Categories									
Familiarity	No idea how they work (2)	Yes (9)								
Other search tools	Peers (4)	Direct website (2)	Social networks (3)	Wikipedia (3)	Library (4)	Database (3)	Academic institutions (1)	Forums (1)		
Academic research	Yes - google (7)	Google Scholar (2)	No (1)	Yes (2)	Bing (1)	Yahoo is bad (1)				
Results read on SERP	Top first (7)	2-3 pages (6)	1st page only (4)	Will scroll (3)	New tabs (3)					
New page or modified query	Next page (2)	Modify query (9)								
Time spent on single search	Quick skim (5)	Examine relevancy (1)								
Determine website trust	Dates (1)	Consistency (1)	Previous knowledge (4)	Third party sources(4)	URL (1)	Reputation (5)	Brand identity (2)	Top SERP (1)	Credibility (1)	Garbled text (safety) (1)
Where looking	Dates (5)	Headline (14)	URL (9)	Abstract-only in deep searches (3)	Keywords (4)	Abstract (3)				
Search engine influence	Dates (1)	Aesthetics (1)	New info - Self-doubt (1)	Puts words in place (2)	It knows what you want (2)	Skeptical of top spots bc of companies paying (2)	Convenience (2)	First page/top more trustworthy (3)	Total trust (1)	Misdirection (1)
Determining search engine trust	Previous experiences (5)	Backlogs (1)	Consistent information (1)	Why not? (1)	Wary of ads (3)	Accurate (1)	Numb to ads (3)	Search for facts (1)		
Why click?	Familiar URLs (1)	Match to head (3)	Dates (2)	Pictures (4)	Scan for numbers (1)					
Search intent - different levels for:	Information only (1)	In-depth information (5)	More pages clicked if questions arise (1)	Specific information (1)	Academics (2)	Cursory knowledge (2)	Acceptable error (1)	Purchasing (2)	Exploring (1)	

Table A7

Summary of Findings

Quantitative		Qualitative	
Confirmatory	Contradictory	Confirmatory	Contradictory
Expertise and Trustworthiness are most important credibility elements (Kim, 2007; Freeman & Spyridakis, 2004; Iding, Crosby, Auernheimer, & Klemm, 2009; Rains & Karmikel, 2009)	Good will can be measured (McCrosky & Teven, 1999) – debatable based on "not applicable" answers	Users do not understand how search engines work (Jansen, Brown, & Resnick, 2007)	Navigation menus and links increase credibility (Rains & Karmikel, 2009) – this did not come up in the interviews, but that could be because the navigation structure was competent
	Faith in search engines decreases with age (Pew Internet, 2012) – this study found no difference in respondents based on ages	Trustworthiness is important because of malicious websites (Schwarz & Morris, 2011)	Ads decrease credibility judgments (Jansen, Brown, & Resnick, 2007) – ads were ignored or viewed as annoying but did not affect judgments
	Men and women have differing search behaviors (Hotchkiss, 2004) – this study found no difference between genders	High service can compensate for product placement (Bhargava & Feng, 2002)	
		Users skim quickly and move on (Dubicki, 2011; Fox, 2012)	
		Specific site in mind while searching (Dubicki, 2011)	
		Accurate and current information builds trust (Vrontis, Ktoridou, & Melanthiou, 2006)	
		Need for third-party sources Iding, Crosby, Auernheimer, & Klemm, 2009)	

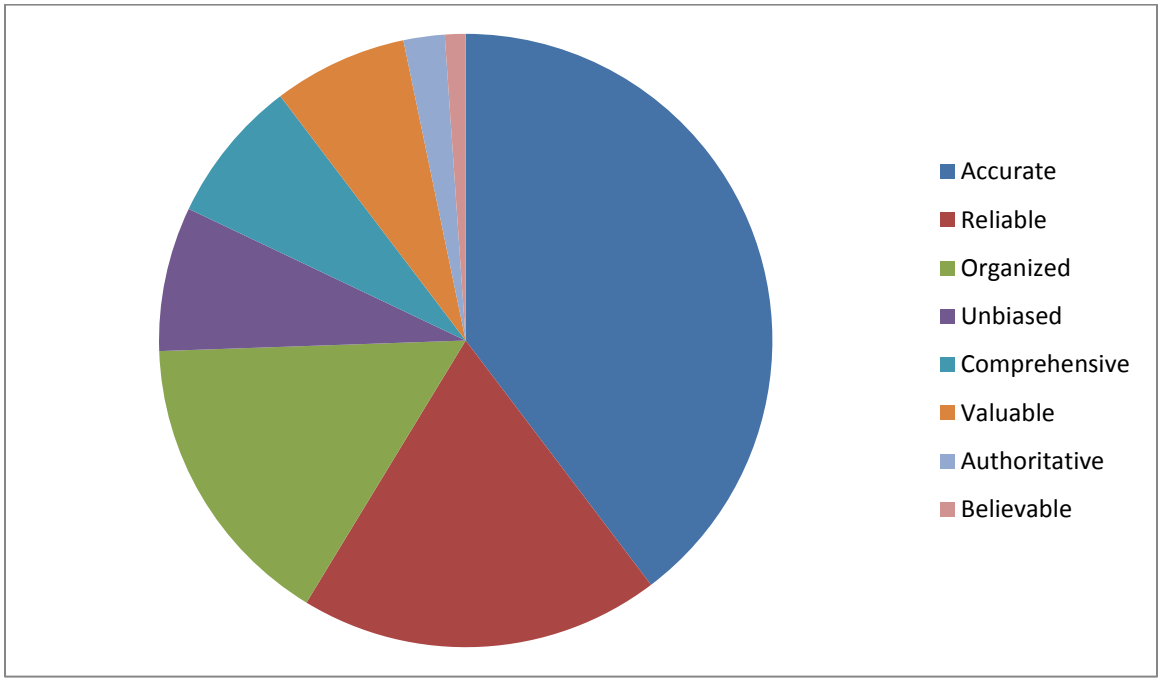

*Figure A*1.Online survey participants' ranking of expertise factors. Percentage of participants who ranked each item as most important.

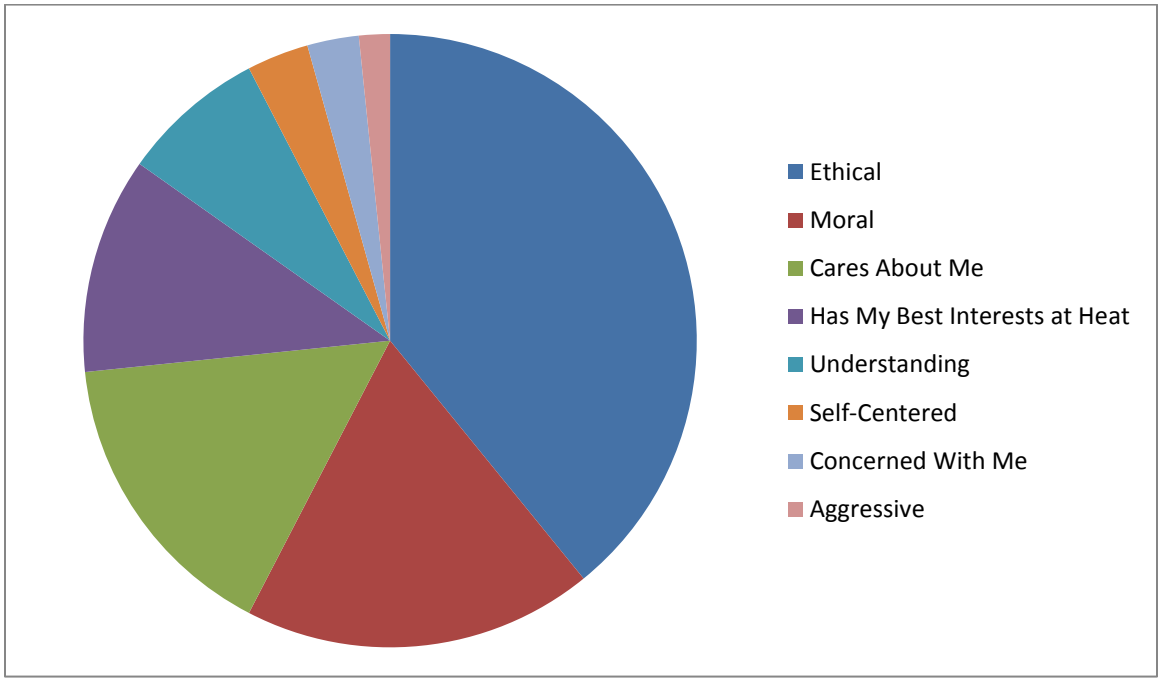

Figure A2. Online survey participants' ranking of goodwill factors. Percentage of participants who ranked each item as most important.

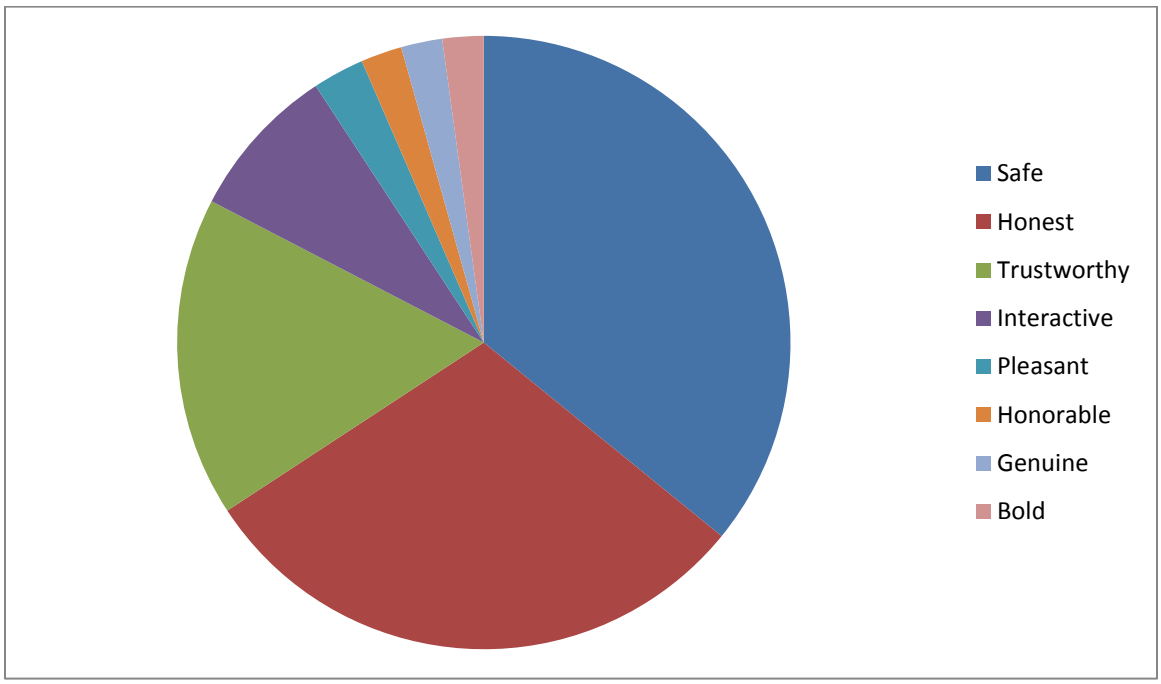

Figure A3. Online survey participants' ranking of trustworthiness factors. Percentage of participants who ranked each item as most important.

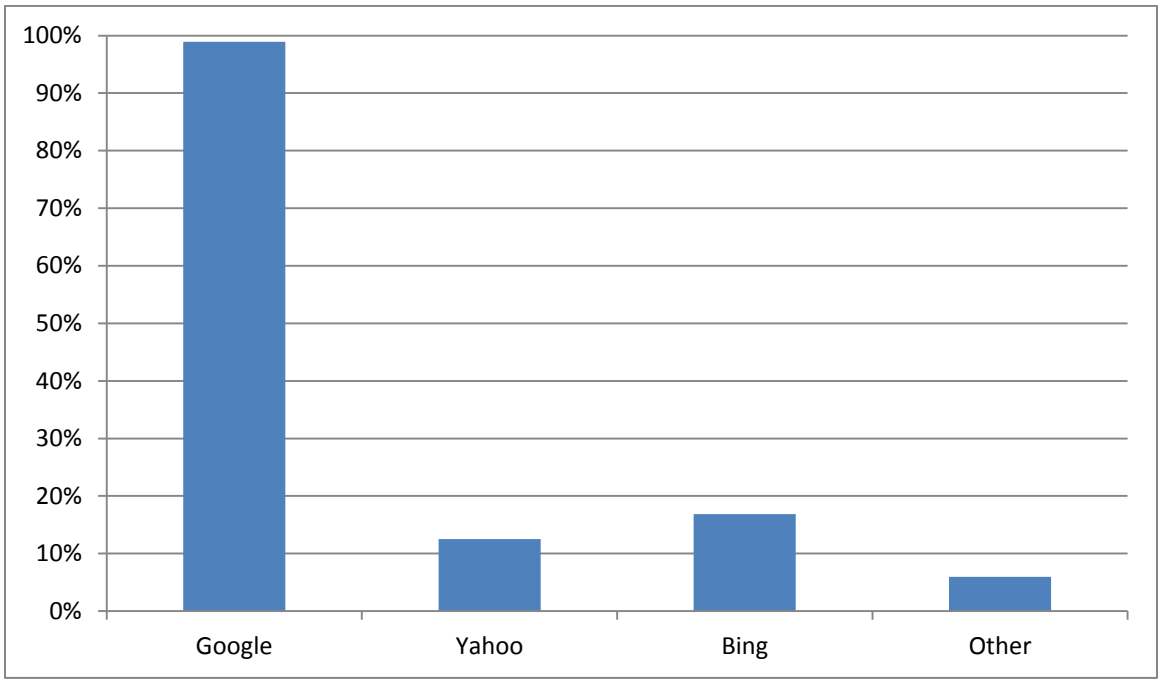

Figure A4. Search engines used by online survey participants. Percentage of participants who had used the search engine in the past week.

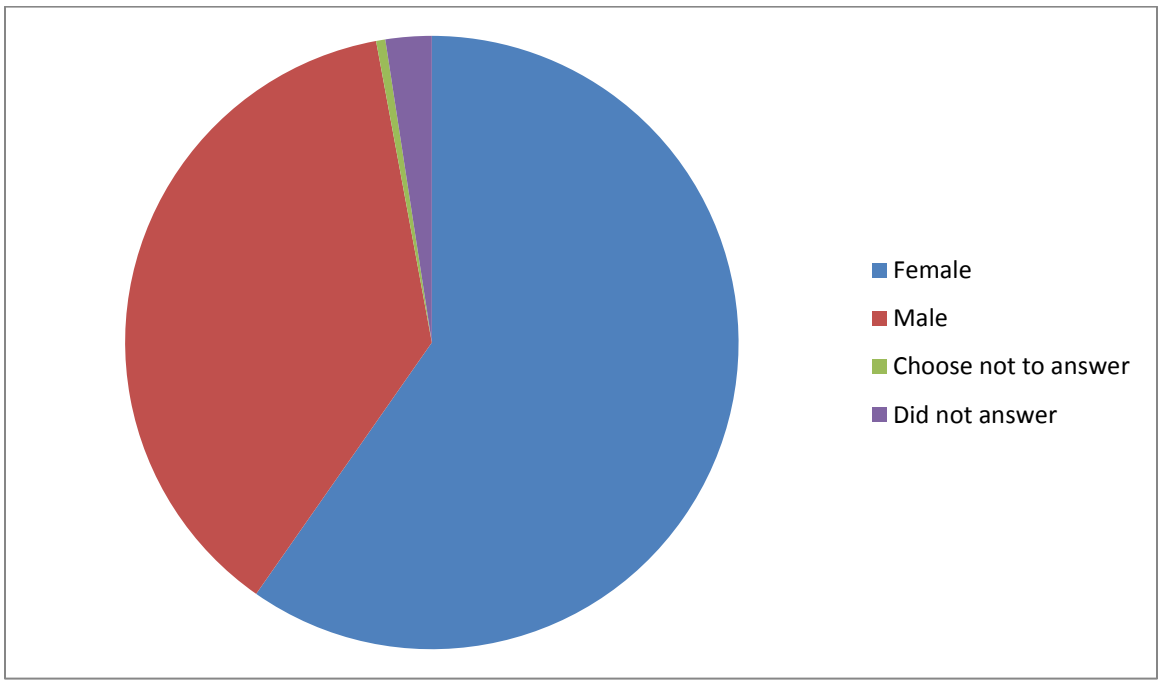

Figure A5. Gender of online survey participants.

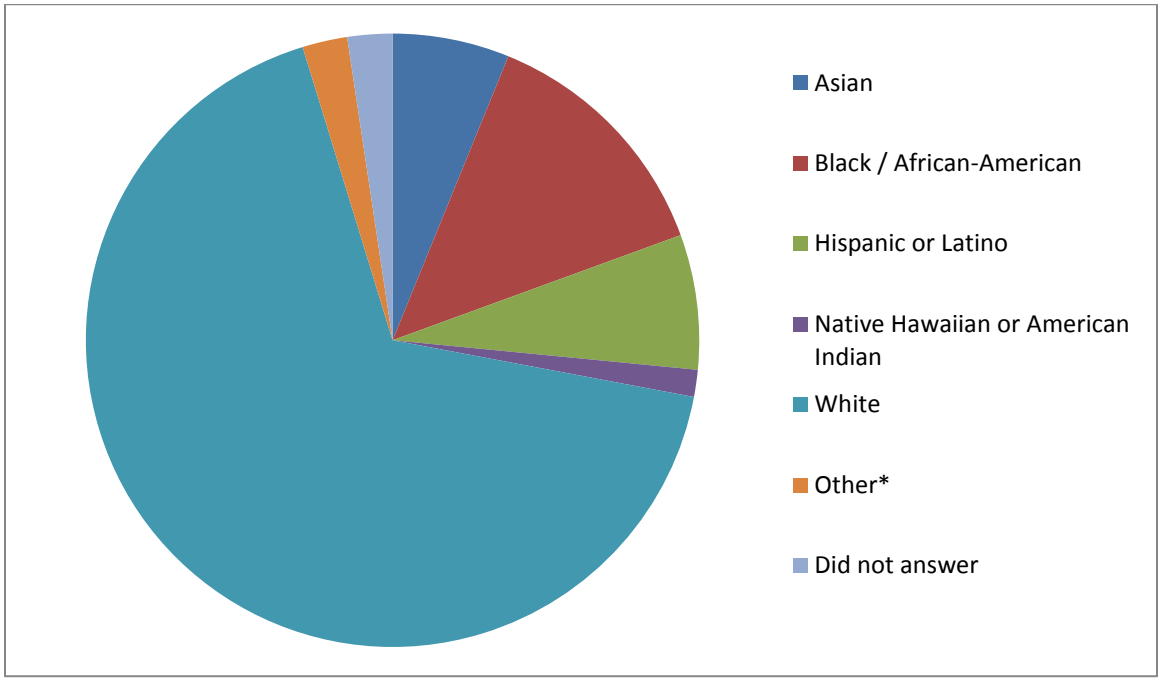

Figure A6. Race/Ethnicity of online survey participants. *Other includes write-in responses of American, Arab, Don't Know, Irish-American, and South Asian.

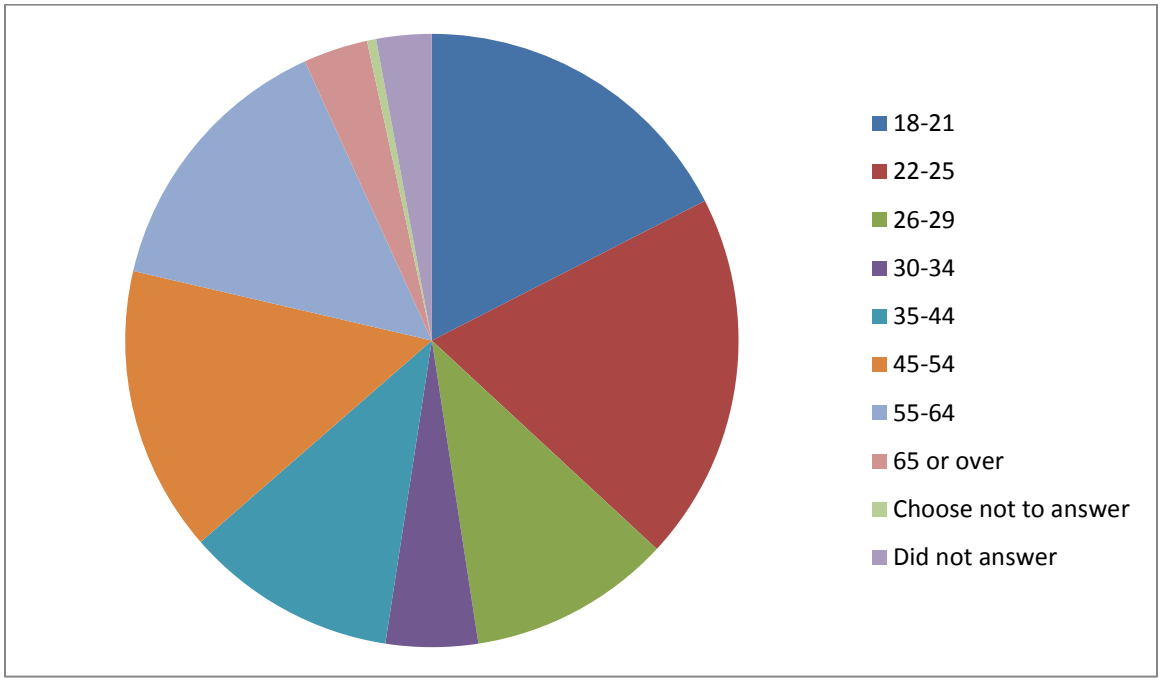

Figure A7. Ages of online survey participants.

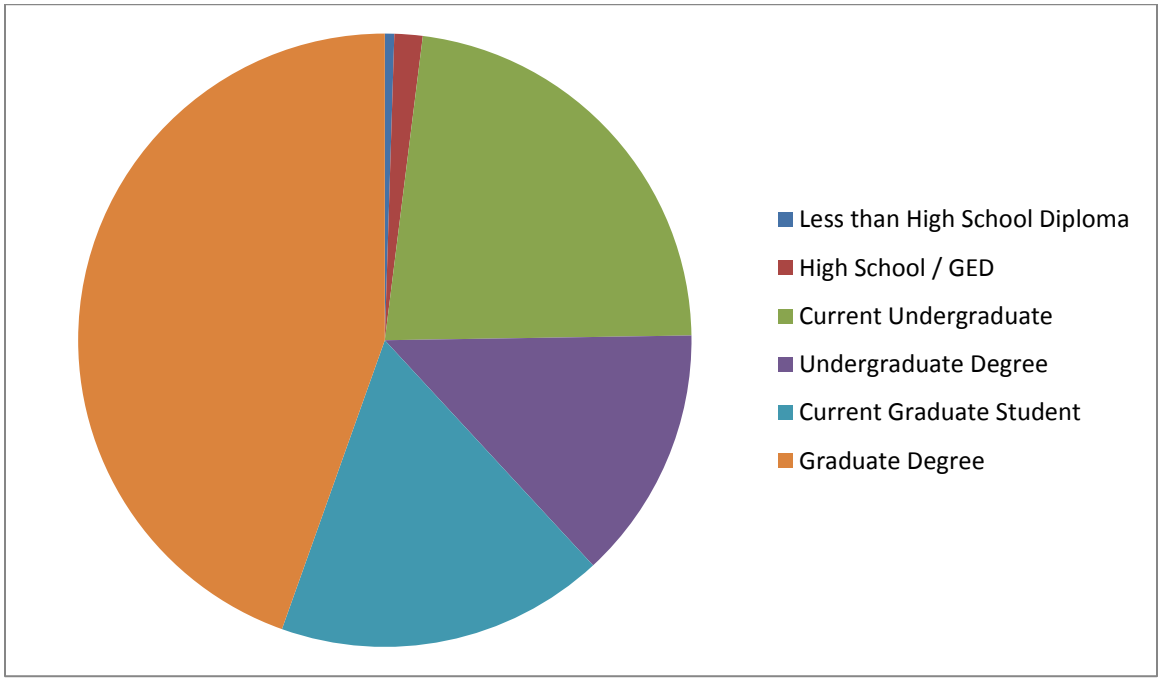

Figure A8. Education level of online survey participants.

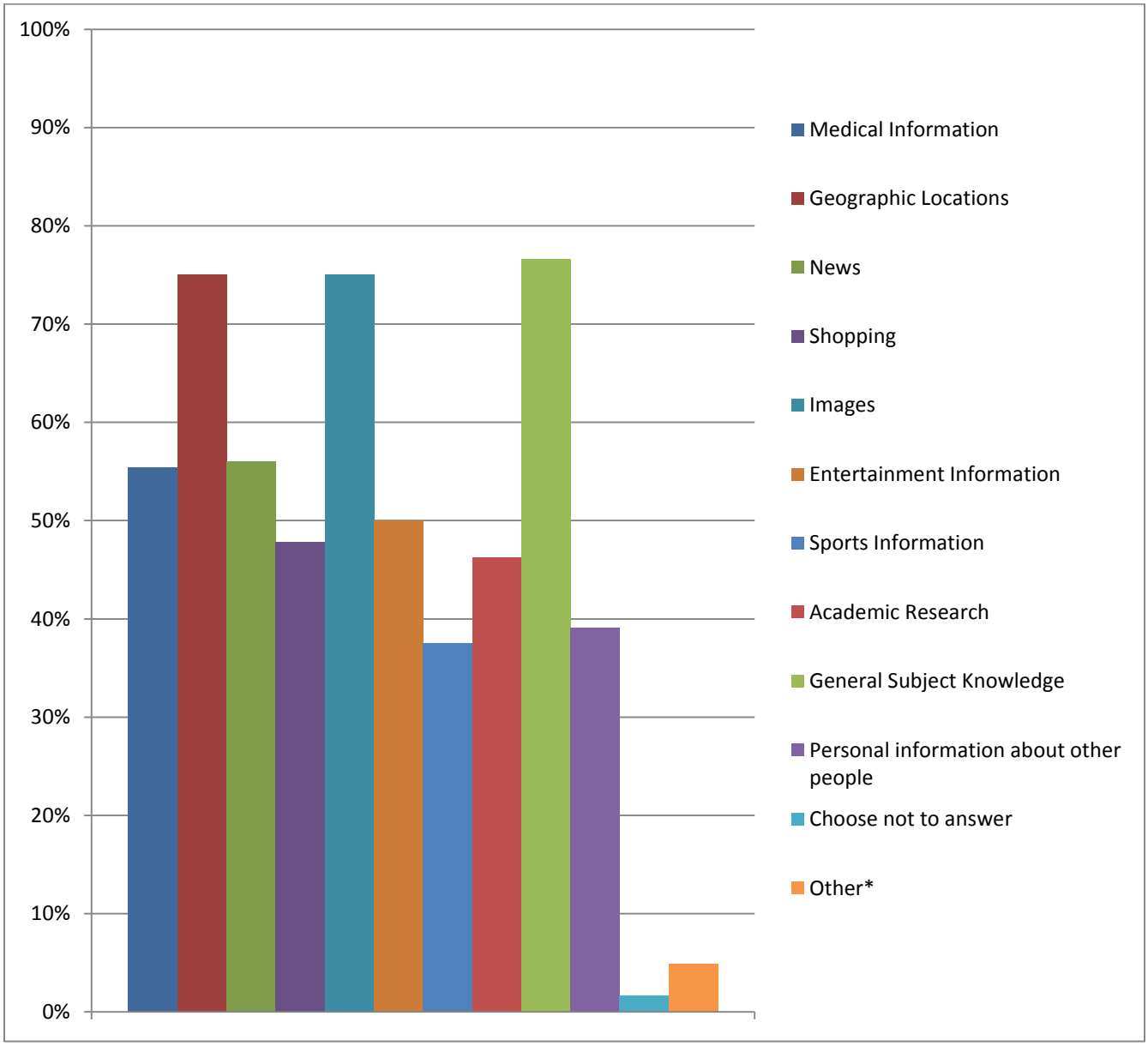

Figure A9. Online survey participants' use of search engines. *Other includes write-in responses of: Food and nutrition; work related (artists); email, calendar, documents, videos; RSS feeds; GoogleScholar; Legal Professionals, Services, Home Repair, Products; Ravel venues [sic].

Appendix B: Survey Instrument

Search Engine Reputation, Aesthetics, and Goodwill

Welcome, folks! This survey is designed to assess your interpretation of search engines: the validity of their results, your feelings towards the companies creating the engine, and lastly, the overall design of the website itself. Your responses will allow us to better understand how internet users seek and interpret information online. The survey is broken up into five main parts, followed by some demographic questions. It should not take longer than 10 minutes to complete.

This survery is part of a research study. Your participation in this study is voluntary and you have the right at any point in time to choose not to participate or to stop participating for any reason with no penalty or loss of benefits. The researcher anticipates that your participation will not cause any major harm or pose a significant health risk. You may benefit from participating in this research by gaining a greater understanding of your own online information seeking behavior.

If you have any questions or concerns about your rights, the study itself, or its outcome you may contact Bryan Christiansen at (585) 313-8537 or bcc6845@rit.edu. Your results will be kept anonymous and confidential.

1. Which search engines have you used in the past week in order to search for information online? You may check as many as applicable.

- ☐ Google
- ☐ Yahoo
- ☐ Bing
- ☐ Other

2. Part I: The Search Engine Results

Instructions: Consider the search engine(s) you have used the most in the past week. Select on a scale from 1 to 7 how well you think the search engine results fit the following descriptions. For instance, for the first item, "Organized," a 1 would mean you think the information is very disorganized and a 7 would mean you think the information is very organized.

	(Not at all) 1	2	3	(Neutral) 4	5	6	(Very) 7	Not applicable
Organized	○	○	○	○	○	○	○	○
Believable	○	○	○	○	○	○	○	○
Authoritative	○	○	○	○	○	○	○	○
Valuable	○	○	○	○	○	○	○	○
Unbiased	○	○	○	○	○	○	○	○
Accurate	○	○	○	○	○	○	○	○
Comprehensive	○	○	○	○	○	○	○	○
Reliable	○	○	○	○	○	○	○	○

3. Rank the items in order of most important to least important to you:

drag the items to the desired positions

- ✣ Organized
- ✣ Believable
- ✣ Authoritative
- ✣ Valuable
- ✣ Unbiased
- ✣ Accurate
- ✣ Comprehensive
- ✣ Reliable

4. Part II: The Search Engine Company

Instructions: Now, still considering the search engine(s) you've used the most over the past week, select on a scale from 1 to 7 how well the search engine company fits with the following descriptions. For instance, for under the first item, "Cares about me," a 1 would indicate that you think the company does not care about you at all and a 7 would mean you think the company cares about you very much.

	(Not at all) 1	2	3	(Neutral) 4	5	6	(Very) 7	Not applicable
Cares about me	○	○	○	○	○	○	○	○
Has my interests at heart	○	○	○	○	○	○	○	○
Self-centered	○	○	○	○	○	○	○	○
Concerned with me	○	○	○	○	○	○	○	○
Moral	○	○	○	○	○	○	○	○
Understanding	○	○	○	○	○	○	○	○
Ethical	○	○	○	○	○	○	○	○
Aggressive	○	○	○	○	○	○	○	○

5. Rank the items in order of most important to least important to you:

drag the items to the desired positions

- Cares about me
- Has my best interests at heart
- Self-Centered
- Concerned with me
- Moral
- Understanding
- Ethical
- Aggressive

6. Part III: Search Engine Brand

Instructions: For the final part of this section, still considering your most-used search engine(s), select on a scale from 1 to 7 how well the search engine brand fits the following characteristics. For instance, under the first item, "Honest," a 1 would indicate that you think the brand is very dishonest and a 7 would mean you think the brand is very honest.

	(Not at all) 1	2	3	(Neutral) 4	5	6	(Very) 7	Not applicable
Honest	○	○	○	○	○	○	○	○
Trustworthy	○	○	○	○	○	○	○	○
Honorable	○	○	○	○	○	○	○	○
Pleasant	○	○	○	○	○	○	○	○
Safe	○	○	○	○	○	○	○	○
Genuine	○	○	○	○	○	○	○	○
Bold	○	○	○	○	○	○	○	○
Interactive	○	○	○	○	○	○	○	○

7. Rank the items in order of most important to least important to you:

drag the items to the desired positions

- Honest
- Trustworthy
- Honorable
- Pleasant
- Safe
- Genuine
- Bold
- Interactive

8. Part IV: Reputation

Instructions: Now, answer to what degree you agree with the following statements concerning the search engine(s) you have used the most in the past week. If you have no opinion or cannot determine an opinion, you may select Undecided. If you do not feel the question is applicable, you may select Not Applicable.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
This search engine has a good reputation	○	○	○	○	○	○
This search engine has a good reputation compared to other rival search engines	○	○	○	○	○	○
This search engine has a reputation for offering good products and services	○	○	○	○	○	○
This search engine has a reputation for being fair in its relationship with its users	○	○	○	○	○	○
I think that this search engine usually fulfils the commitments it assumes	○	○	○	○	○	○
I think that the information offered by this site is sincere and honest	○	○	○	○	○	○
I think I can have confidence in the results that this search engine returns	○	○	○	○	○	○
The search engine is known for the transparency in the services it offers the consumer	○	○	○	○	○	○
I think that the advice and recommendations given on this search engine benefit both the search engine and myself	○	○	○	○	○	○
I think that this search engine is concerned with the present and future interests of its users	○	○	○	○	○	○
I think that this search engine takes into account the repercussions that their actions could have on the consumer	○	○	○	○	○	○
I think that the design and commercial offers of this search engine take into account the desires and needs of its users	○	○	○	○	○	○
I think that this search engine is receptive to the needs of its users	○	○	○	○	○	○
I think that this search engine has the necessary technology to work properly	○	○	○	○	○	○
I think that this search engine has the necessary resources to successfully carry out its activities	○	○	○	○	○	○
I think that this search engine knows its users well enough to offer them products and services adapted to their needs	○	○	○	○	○	○
I think that I will continue my relationship with this search engine in the next few months	○	○	○	○	○	○
I would recommend this search engine to other people	○	○	○	○	○	○
I would point out the positive aspects of this search engine if anybody criticised it	○	○	○	○	○	○
I would not stop my relationship with this search engine if it made any mistakes	○	○	○	○	○	○
Although new online alternatives appear, I will continue my relationship with this search engine	○	○	○	○	○	○

9. Part V: Aesthetics

Instructions: For the last part of the survey, please again answer to what degree you agree with the following statements concerning the search engine you just used. If you have no opinion or cannot determine an opinion, you may select Undecided. If you do not feel the question is applicable, you may select Not Applicable.

	Strongly Disagree	Disagree	Undecided	Agree	Strongly Agree	Not Applicable
I think the colors used in the search engine were appealing to me	○	○	○	○	○	○
I think that the fonts used on the search engine pages was attractive to me	○	○	○	○	○	○
I think that the graphics used on the search engine appealed to me	○	○	○	○	○	○
I think that the arrangement of both text and graphics together was appealing to me	○	○	○	○	○	○
The search engine displayed links to other search options, such as image or news searches	○	○	○	○	○	○
The search engine displayed links to other search options prominently	○	○	○	○	○	○
It was easy to find other search options on the page	○	○	○	○	○	○
The display of these links fits with the general scheme of the search engine	○	○	○	○	○	○
The search engine minimizes vertical scrolling	○	○	○	○	○	○
The search engine minimizes horizontal scrolling	○	○	○	○	○	○
It was intuitive to find the links I was looking for	○	○	○	○	○	○
The presence of ads were intrusive	○	○	○	○	○	○
I pay attention to sponsored content	○	○	○	○	○	○

10. What is your sex?

- ○ Female
- ○ Male
- ○ Choose not to answer

11. What is your race / ethnicity? You may choose more than one answer.

- ☐ Asian
- ☐ Black / African-American
- ☐ Hispanic or Latino
- ☐ Native Hawaiian or American Indian
- ☐ White
- ☐ Other: _____

12. How old are you?

- ○ 17 or under
- ○ 18-21
- ○ 22-25
- ○ 26-29
- ○ 30-34
- ○ 35-44
- ○ 45-54
- ○ 55-64
- ○ 65 or over
- ○ Choose not to answer

13. What is your highest level of education?

- Less than High School Diploma
- High School / GED
- Currently enrolled as an Undergraduate College Student
- Undergraduate Degree
- Currently enrolled as a Graduate Student
- Graduate Degree
- Choose not to answer

14. If you are currently enrolled in classes, what is your current program of study?

15. One more! To gain information about any of the following, would you turn to a search engine FIRST? You may select more than one answer.

- Medical information
- Geographic locations
- News
- Shopping
- Images
- Entertainment information
- Sports information
- Academic research
- General subject knowledge
- Personal information about other people
- Choose not to answer
- Other:

Submit

Copyright © 2013 Rochester Institute of Technology. All Rights Reserved. | Disclaimer | Copyright Infringement

Color Scheme: orange (default)

Appendix C: Qualitative Guiding Questions

1) Are you familiar with how search engines work and how to use them?

2) What tools have you used prior to today to conduct research?

3) Do you generally use search engines for academic research?

4) Which results will you read on a Search Engine Results Page? Only the first result at the top, the first couple, or everything?

5) If you do not find the answer you are looking for on the first SERP, are you more likely to go on to the next page or modify your query?

6) How much time will you take on a single query before you search for something else?

7) How do you determine which results contain trustworthy information?

8) Are you more attracted to the title or the abstract when you decide where to click?

9) Does the search engine itself have an influence on your search?

10) How do you trust search engines?

11) Why do you click where you do?

12) How does intent impact the rigor of your result evaluations?