

**Experts vs. Online Consumers:
A Comparative Credibility Study of
Health and Finance Web Sites**

**Julianne Stanford and Ellen R. Tauber
Sliced Bread Design, LLC**

**B.J. Fogg, Ph.D.
Stanford Persuasive Technology Lab**

**Leslie Marable
Consumer WebWatch**

October, 2002

TABLE OF CONTENTS

EXECUTIVE SUMMARY	3
INTRODUCTION	6
METHODS	8
STUDY DESIGN	8
CHOOSING CATEGORIES AND SELECTING WEB SITES	9
SUBJECTS	10
RUNNING THE STUDY	10
ANALYZING THE DATA	13
PART 1: PAIR-RANKING TASK ANALYSIS	14
PART 2: 1-TO-10 RANKING TASK ANALYSIS	15
DATA CONSISTENCY BETWEEN PART 1 AND PART 2	15
PART 3: VARIABLES	16
PART 4: OPTIONAL QUESTIONS	16
RESULTS AND DISCUSSION	16
PARTS 1 AND 2: SITE RANKINGS RESULTS AND DISCUSSION	16
HOW HEALTH EXPERTS RANKED SITES	17
OVERALL HEALTH-SITE EVALUATION TRENDS	20
HOW FINANCE EXPERTS RANKED SITES	22
OVERALL FINANCE-SITE EVALUATION TRENDS	24
DIFFERENCES BETWEEN FINANCE AND HEALTH EXPERTS	27
DIFFERENCES BETWEEN EXPERTS AND CONSUMERS	28
EXPERT VS. CONSUMER HEALTH RANKINGS	28
EXPERT VS. CONSUMER FINANCE RANKINGS	30
PART 3: VARIABLE RANKINGS RESULTS AND DISCUSSION	32
VARIABLE SCORES OF HEALTH VS. FINANCE EXPERTS	34
VARIABLE SCORES OF EXPERTS VS. CONSUMERS	34
PART 4: OPTIONAL QUESTIONS RESULTS AND DISCUSSION	39
CONSUMER EDUCATION	39
GUIDELINES AND DESIGN IMPLICATIONS	40
ANALYSIS OF EXISTING CONSUMER WEBWATCH GUIDELINES	40
ADDITIONAL FIELD-SPECIFIC GUIDELINES	42
FURTHER RESEARCH	43
APPENDIX A: SITES WITH WEB ADDRESSES	44
APPENDIX B: LIST OF SUBJECTS	45
APPENDIX C: OPTIONAL QUESTIONS	48
APPENDIX D: CODING CATEGORIES	49
APPENDIX E: TABLES OF SITE-PAIRING SCORES	56
APPENDIX F: TABLE OF EXPERT VS. CONSUMER VARIABLE SCORES	57
REFERENCES	59

LIST OF FIGURES

Figure 1: Description of comment categories	14
Figure 2: Health-site rankings	17
Figure 3: Percentage of health-expert comments, by category	20
Figure 4: Finance-site rankings	22
Figure 5: Percentage of finance-expert comments, by category	24
Figure 6: Expert vs. consumer health-site rankings	28
Figure 7: Expert vs. consumer health-site comments	29
Figure 8: Expert vs. consumer finance-site rankings	30
Figure 9: Expert vs. consumer finance-site comments	31
Figure 10: Expert variable scores	33
Figure 11: Finance-expert vs. consumer variable scores	34
Figure 12: Health-expert vs. consumer variable scores	36

EXECUTIVE SUMMARY

Many researchers have conducted studies on consumers to understand various aspects of Web site credibility. However, some consumers may not be well equipped to make informed decisions about the accuracy of information in technical fields such as health or finance. In fact, previous research in other mediums showed that consumers have had mixed results when judging credibility of information. Given the varying results of previous studies of consumer credibility assessments, we are left wondering whether consumers' credibility evaluations of Web sites are correct. These evaluations are increasingly important as people use the Web today to carry out a variety of vital tasks and research. Consumers are faced with important decisions about the information sources that they choose to believe for making important health or financial decisions. Do these everyday people know which Web sites are really credible, especially in vital areas such as finance and health? What do industry experts say about the credibility of sites in their fields? And, finally, how do the experts' assessments compare to how the average person decides which sites to trust?

To answer these credibility-related questions, Sliced Bread Design and Consumer WebWatch produced this expert study, titled *Experts vs. Online Consumers: A Comparative Credibility Study of Health and Finance Web Sites*, in collaboration with Stanford University's Persuasive Technology Lab (Stanford PTL). As a usability and interaction design agency, Sliced Bread Design, LLC is interested in understanding how to create ethical online experiences that communicate credibility to their users. Consumer WebWatch, which commissioned this study, has the goal to investigate, inform, and improve the credibility of information published on the World Wide Web. Due to these shared goals, the two organizations decided to embark on a study of how health and finance experts judge the credibility of websites within their domains of expertise.

This expert study was launched jointly with a parallel, consumer-focused study, *How Do People Evaluate a Web Site's Credibility? Results from a Large Study*, conducted by the Stanford PTL and Consumer WebWatch, which asked over 2,600 average people to rate the credibility of Web sites in 10 content areas. Our study then compared how experts and consumers evaluated the same health and finance sites to understand if and how consumers failed in their assessments. By comparing the expert and consumer evaluations, we hoped to identify any gaps in consumer education and begin to design guidelines for improving consumer understanding of online credibility. Furthermore, by studying experts in two diverse fields — health and finance — we aimed to learn about field-specific credibility in order to inform Web design guidelines and consumer education needs.

About These Joint Studies

Sliced Bread Design: In this study, 15 experts from the health and finance fields were asked to assess the credibility of sites in their respective areas. A total of 8 health and 7 finance experts visited the same 10 health sites or 10 finance sites as the consumers in the Stanford PTL study. They were asked to rank from 1-to-10 the credibility of the sites specific to their area of expertise, as well as provide detailed written assessments of each site under review. (See Appendix B for a list of the expert participants and brief bios.)

Stanford Persuasive Technology Lab: In the parallel consumer study, 2,684 people evaluated the credibility of two Web sites randomly assigned from one of ten content categories (e-commerce, entertainment, finance, health, news, nonprofit, opinion or review, search engines, sports, and travel.) A total of 100 sites were assessed. The 2,440 comments that people wrote about each site's credibility were gathered and analyzed to track which features of a Web site were noticed — or went unnoticed — when consumers evaluated credibility online. (See Appendix A in the Stanford PTL study report for a complete list of Web sites selected for the consumer study.)

Key Findings

The online consumers in the Stanford PTL study and our expert panel of evaluators diverged greatly in their credibility assessment criteria. Overall, our experts were far less concerned about visual appeal as a marker of credibility than the Stanford PTL consumers, and they were more concerned about the quality of a site's information. Among other conclusions, this study found that health experts assigned more credibility to health sites that provided information from reputable sources and cited the names and credentials of authors for each article published. Finance experts assigned more credibility to finance sites that provided investors with a great deal of unbiased educational information and research, rather than nudging consumers toward their own products or services.

The Stanford PTL study found that consumers tended to rely heavily on overall visual design when assessing Web sites, including layout, typography and color schemes. More than half (54.6%) of the comments by the consumers regarding finance sites referred to *design look*, which relates to the visual appeal of a site's design, compared to only 16.4 percent of finance expert comments on this topic. Likewise, 41.8 percent of consumer comments regarding health sites made note of a site's *design look*, compared to 7.6 percent of surveyed health experts' comments which mentioned this topic.

Our health experts most often relied on the *name reputation* of a site, its operator, or that of its affiliates, when assessing the credibility of health Web sites (43.9% of health expert comments related to this credibility criterion). The next most common issues mentioned when evaluating health site credibility were *information source*, which relates to the citation of a site's information sources (25.8%), and *company motive*, which relates to a user's perception of the motive of the organization behind the site, whether good or bad (22.7%).

Our finance experts most often relied on a site's scope or *information focus* when assessing the credibility of finance Web sites, which includes consideration of the quantity of information provided (40.3% of finance expert comments related to this credibility criterion). The next most common issues mentioned when evaluating finance site credibility were *company motive* (35.8%), and *information bias* (29.9%), which relates to a user's perception of bias in the site's content.

This study also reveals which specific elements lend credibility to a site's perception, according to each health and finance expert group. In addition, the report provides recommendations to Web publishers, particularly those in the health and finance fields, which aim to increase the credibility of sites among each type of Internet audience.

About Consumer WebWatch

Consumer WebWatch is a project of Consumers Union, the non-profit publisher of *Consumer Reports* magazine and ConsumerReports.org. The project is supported by The Pew Charitable Trusts, which invests in ideas that fuel timely action and results; the John S. and James L. Knight Foundation, which promotes excellence in journalism worldwide and invests in the vitality of 26 U.S. communities; and the Open Society Institute, which encourages debate in areas in which one view of an issue dominates all others. Consumer WebWatch's Web site launched April 16, 2002. <http://www.consumerwebwatch.org>

About Sliced Bread Design, LLC

Sliced Bread Design is a usability and interaction design agency established by Julie Stanford and Ellen Tauber to help people effectively use and enjoy interactive products. Sliced Bread Design provides user study and interface design services to help organizations create compelling online, desktop, voice, and wireless software. Its work appears in a variety of products, ranging from online enterprise resource planning software for Fortune 500 companies to consumer software for mobile phones. The company is located in California's Silicon Valley. <http://www.slicedbreaddesign.com>

About Stanford Persuasive Technology Lab

The Stanford Persuasive Technology Lab (captology.stanford.edu) creates insight into how computers can be designed to change what people think and do, an area called "captology." Directed by experimental psychologist B.J. Fogg, the Stanford University team includes social scientists, designers, and technologists who research and design interactive products (from Web sites to mobile phones) that motivate and influence their users. Although this area can be controversial, the lab's focus is on using the power of persuasive technology to promote positive changes in domains such as health, education, business, and safety. As part of the lab's mission, since 1998 the team has investigated what causes people to believe -- or not believe -- what they find online. <http://credibility.stanford.edu>

INTRODUCTION

Various researchers have conducted studies on consumers to understand the different elements of Web site credibility. However, some consumers may not be well equipped to make informed decisions about the accuracy of information in technical fields such as health or finance. In fact, consumers have had mixed results when judging credibility of information in other mediums. A recent study compared teacher and scientist perspectives on the credibility of 31 science information sources, including print media such as *Discover* magazine, radio and television programs such as ABC News's *Nightline*, and environmental organizations such as the Sierra Club [Klemm]. This study found that teachers and scientists significantly differed in their perceptions of the credibility of different information sources. For example, while scientists in the study rated TV newsmagazines such as ABC News's *20/20* as having the least amount of credibility, elementary school teachers gave the program the highest credibility ratings.

Some studies have shown that consumers rely significantly on visual cues when judging credibility. In a study on the influence of color and graphics on the credibility of Yellow Pages directory advertising [Lohse], subjects perceived colored ads as more credible than black and white. Similarly, a 1990 study of the perceived trustworthiness of television spokespeople found that viewers believed that people with baby-looking faces and females delivered more trustworthy content than people with mature-looking faces and males [Brownlow]. On the other hand, a 1999 study by the American Society of Newspaper Editors examined consumer and journalist assessments of newspaper credibility and found that consumers were even more critical credibility judges than the journalists [Urban]. Consumers in the newspaper study were particularly concerned about factual errors, spelling and grammar mistakes, journalistic bias, manipulation of the news by parent companies, and sensationalism when assessing the credibility of newspaper articles.

Given the varying results of studies of consumer credibility assessments in other mediums, we are left wondering whether consumers' credibility evaluations of Web sites are correct. These evaluations are increasingly important as people use the Web today to carry out a variety of vital tasks and research. Consumers are faced with important decisions about the information sources that they choose to believe for making important health or financial decisions. So, do these everyday people know which Web sites are really credible, especially in vital areas such as finance and health? What do industry experts say about the credibility of sites in their fields? And, finally, how do the experts' assessments compare to how the average person decides which sites to trust?

We started to answer these questions by conducting a comparative study of expert versus consumer evaluations of Web site credibility. Our goal for this study was to approach credibility from a new perspective. Instead of focusing on how consumers judge credibility, we endeavored to understand if consumers' judgments are correct and if not, why not? To reach this goal, we asked experts in the finance and health fields to evaluate sites in their domains of expertise, describe how they appraised each site, and then rank each site's credibility in relation to other sites. This expert study paralleled a consumer study conducted by the Stanford Persuasive Technology Lab (Stanford PTL) and Consumer WebWatch that asked over 2,600 average people to rate the credibility of Web sites in 10 content areas [Fogg, 2002]. Our study then compared how experts and consumers evaluated the same health and finance sites to understand if and how consumers failed in their assessments. By comparing the expert and consumer evaluations, we

hoped to identify any gaps in consumer education and begin to design guidelines for improving consumer understanding of online credibility. Furthermore, by studying experts in two diverse fields — health and finance — we aimed to learn about field-specific credibility in order to inform Web design guidelines and consumer education needs.

Every study of credibility must first lay the groundwork by defining the term *credibility*. In this paper we adhere to the definition of credibility outlined by Fogg and Tseng (1999), with the following discussion drawing largely from this work. In their view, credibility can be defined as believability. Credible information is believable information. There are two main ideas that help one to understand the construct of credibility. First, credibility is a perceived quality. It is not an objective property of a Web site like how many words it contains or how many links are on the page. Instead, when one discusses credibility, it is always from the perspective of the observer's perception. Second, people perceive credibility by evaluating multiple dimensions simultaneously. These dimensions can be categorized into two key components: *trustworthiness* and *expertise*. The *trustworthiness* component refers to the goodness or morality of the source and can be described with terms such as well-intentioned, truthful, or unbiased. The *expertise* component refers to perceived knowledge of the source and can be described with terms such as knowledgeable, reputable, or competent. People combine assessments of both trustworthiness and expertise to arrive at a final credibility perception. Although this definition positions credibility as a subjective perception, we assumed for this study that subject-matter experts would be better judges of a site's credibility and content quality than a consumer with no particular expertise or knowledge specialty.

As a usability and interaction design agency, our organization is interested in understanding how to create ethical online experiences that communicate credibility to their users. Consumer WebWatch, a nonprofit project of Consumers Union, publisher of *Consumer Reports*, commissioned this study. Consumer WebWatch's goal is to investigate, inform, and improve the credibility of information published on the World Wide Web. It is supported by grants from The Pew Charitable Trusts, the John S. and James L. Knight Foundation, and the Open Society Institute.

We believe that this investigation will advance Consumer WebWatch's and Sliced Bread Design's goals by setting credibility benchmarks for finance and health sites. We do not claim that the rankings and comments in this study are the definitive evaluations of the credibility of the sites involved. In fact, it is essential to understand that this study is not an investigation of the *most* credible health and finance sites on the Web. Rather, it is a study of the credibility of the particular sites that we have chosen for the comparative purposes of this research.

Unfortunately, this study confirms some of our suspicions about consumers' poor credibility assessments. As you read on, you will find that experts carefully evaluated content while consumers relied on visual appeal for much of their credibility appraisal. However, the good news is that this study identifies specific opportunities for consumer education. By realistically identifying the sources of the problem and the directions for resolution, we can begin to inform consumers and enhance the usefulness of the Web, which is arguably the biggest, most accessible information source on Earth. After all, as the cliché says, knowledge is power.

METHODS

STUDY DESIGN

The format of this study of experts' assessment of Web site credibility was based on methods for Web credibility research developed by the Stanford Persuasive Technology Lab (Stanford PTL). By running pilot studies, Stanford PTL developed an online research method using paired comparisons of Web sites. Subjects were asked to review two Web sites, select one as more credible, and comment on his or her choice. When run with a large number of subjects, the paired method resulted in a relative ranking of the credibility of a group of Web sites and a rich base of comments. This paired comparison method was used in a consumer study by the Stanford PTL that ran at the same time as this expert study, in the summer of 2002 [Fogg, 2002]. The Stanford PTL 2002 consumer study asked average people to rate the credibility of a single pair of Web sites in the same category (i.e., health, finance, news, sports, etc.). Each Web site in the pair was randomly selected from a pool of 10 content categories with 10 sites in each category. The study collected over 2,600 consumer rankings of paired sites from these 10 categories. The health and finance categories received 228 and 408 consumer rankings, respectively.

Our expert study examined a small number of experts' assessments of sites in their fields. In contrast to the consumer study, which asked an individual to rank one pair of sites from the 10 sites in a category, our expert study asked each expert to rank five pairs of sites in one session. This allowed each expert to review all 10 of the sites in his or her category of health or finance. Asking the experts to assess pair rankings allowed some comparison to the consumer study, but the small number of pairings was not sufficient to produce a reasonable overall ranking of sites. Both the Stanford PTL consumer pilot tests and our own expert pilot tests confirmed that having subjects start with a 1-to-10 ranking of all sites was not effective because subjects had trouble assessing more than two sites simultaneously on initial review. However, if we introduced the sites in randomized pairs first, we could successfully present the 1-to-10 ranking task next. Thus, to allow some direct comparison to the consumer study and produce an overall expert ranking of sites in our study, we started experts with the paired comparison method, and followed with a secondary overall ranking task in which the experts ordered all the sites from most to least credible.

We also added a section to the expert study that was based on a separate Stanford PTL credibility study, conducted in 2001, which examined the elements of Web sites that contributed to perceptions of credibility [Fogg, 2001]. The Stanford PTL 2001 study asked 1,400 consumers to evaluate elements of Web sites that helped or hurt credibility, such as "The site lists author's credentials for each article." For each variable, non-expert subjects indicated how that variable affected the credibility of Web sites in general by selecting a response along a 7-point Likert-type scale from -3 (much less believable) to +3 (much more believable). Our expert study presented just 30 of the previous 55 variables, because our expert pilot studies revealed that inclusion of all 55 variables would negatively impact the overall length and feel of the study session. As a result, we eliminated less relevant items, such as "The site was recommended to you by a friend." We also eliminated similar variables, such as "The site is difficult to navigate," vs. "The site is arranged in a way that makes sense to you." The inclusion of this variable task allowed us to measure experts' assessment of each variable's affect on credibility and then compare the expert responses to those of the Stanford PTL's large group of consumers. The consumers in the

Stanford PTL 2001 study were asked to evaluate each variable in relation to all Web sites in general, while experts in our study were asked to evaluate each variable only in relation to Web sites in their fields of expertise (finance or health). Also, note that the Stanford PTL 2001 study consumer group used for comparison in this variable evaluation task was a different group of people than the Stanford PTL 2002 study consumer group used for comparison in the Web site ranking task described in the previous paragraph.

In addition, we wanted to gather open-ended thoughts from the experts about credibility. To do so, we included an optional section in which we posed credibility-related questions for subjects to answer in an unrestricted text field. We created the questions by compiling a list of key credibility areas, such as visual design and site ownership, forming candidate questions, and reducing the number to five individual questions. We then refined the language for each question during pilot studies we conducted in our own pre-testing phase. The final five questions are listed in Appendix C.

CHOOSING CATEGORIES AND SELECTING WEB SITES

Working with Consumer WebWatch and the Stanford PTL, we selected two content categories from the 10 in the consumer study: health and finance. The remaining consumer categories were: e-commerce, entertainment, news, nonprofits, opinion or review, search engines, sports, and travel. We felt that the health and finance categories were particularly relevant for a credibility study because they were heavily-trafficked by consumers, allowed for identification of experts in the given fields, and presented an interesting difference between a primarily informational category (health) and a primarily transactional category (finance). In addition, we chose to study these categories because incorrect information on a health or finance site could be particularly harmful to a consumer if used for making an important health or financial decision.

For each category, Consumer WebWatch and Sliced Bread Design worked in conjunction with the Stanford PTL to select 10 Web sites to include for the expert study. The sites chosen were the same health and finance sites assessed by consumers in the parallel Stanford PTL 2002 study. The sites were all consumer-oriented sites, and were selected to present a variety of sites across a range of variables, including:

- Bricks-and-mortar company vs. Internet only
- Range of visual and information design quality and style
- Brand name vs. No name
- Variety in amount and presentation of advertisements
- Presence of awards or seals of approval

In addition, we preferred general sites over content-specific sites; for example, we included health sites that presented a wide range of topics rather than those that focused on one specific health condition such as breast cancer. We excluded sites that required a subscription to access premium content. However, we included sites that required free registration, and provided “dummy” accounts for those sites for use by our panel of experts. All of the finance sites allowed transactions and all of the health sites provided health information, not just product sales. Last and most importantly, we chose sites that we expected to provide a range of credibility ratings. Thus, the study *did not* aim to select the 10 *most* credible sites in a category, but rather 10 sites that would produce a range of credibility rankings and comments. See Appendix A for the list of sites with Web addresses.

Note that the sale of one of the health sites, DrKoop.com, to a Florida company was announced on July 15, 2002, and the site changed significantly because it is no longer affiliated with former U.S. Surgeon General C. Everett Koop. This study was completed before the announced sale.

SUBJECTS

This expert study included eight health experts and seven finance experts, for a total of 15 subjects. We recruited approximately 15 experts in each category for a minimum goal of seven completing the study in each industry area. We did not desire a larger sample size, because our goal was to obtain consistent opinions from a reasonable number of experts within our budget and time constraints. We believed that more experts would have produced redundant results for this study, and in fact our intuition was correct as the experts in each category agreed with the category group as a whole. We discuss our measure of subject agreement below.

With a small number of subjects, the expert selection was important. First, we identified candidates who were experts in their domains of expertise -- finance or health -- and had experience on the Web. We defined *expert* as an accomplished authority in his or her field; someone who an average reasonable person would describe as “expert” if provided with a description of the person’s experience and achievements in his or her field. Next, we reduced the list by screening for geographic diversity throughout the United States. Although we included one German-born doctor-researcher who currently works in Canada, international participation was not a prerequisite in order to match the consumer study subjects. We also balanced the experts for types of expertise to include a mixture of academics and industry practitioners. For example, our finance experts included a professor, a journalist, a financial advisor, and an industry analyst. This process resulted in a list of approximately 15 experts in each field whom we recruited by e-mail and phone. During recruitment, we specifically mentioned Consumers Union and Consumer WebWatch to encourage expert participation.

Subjects who completed the study were paid an honorarium of \$100 and a matching \$100 was donated to a charity of his or her choice selected from a list of the 50 largest U.S. charities, ranked by total income in fiscal year 2000. This list of charities was compiled by *The NonProfit Times*. Two subjects sent the total \$200 to the charity they selected. Please refer to Appendix B for a list of subjects and a description of each person’s expertise.

RUNNING THE STUDY

The study took place between June 14 and July 12, 2002. Subjects accessed the study online via a Web site (finance and health experts were given two separate Web addresses). The first page required the subject to enter his or her name, which reinforced that it was not anonymous, and that he or she agreed to the terms of the study. This Welcome Page specifically mentioned Consumers Union and Consumer WebWatch to reinforce the integrity of the study and encourage the subject to spend more time and provide quality comments. This page also emphasized that the subject must complete the study in one browser session. The second page described the four parts of the study with the estimated time for completing each part, and provided the dummy login account for any sites requiring free registration.

Part 1 of the study presented the five randomly selected site pairs. Each of the pairs was presented on one page, for a total of five pages (study pages 3 through 7). For each pair, the subject chose

one site as “1 = more credible” and the other as “2 = less credible” by selecting from radio buttons. The subject was asked to share comments in a text field.

Sample screen from Part 1

The screenshot shows the 'CONSUMER WEBWATCH' logo at the top left. Below it, the title 'Finance Web Sites' is displayed. A paragraph of instructions asks the user to visit the listed sites and rank them from most to least credible. The interface is titled 'Pair 1 of 5' and contains a table with three columns: 'Web Site', 'Credibility Ranking', and 'Please share your comments'. The 'Credibility Ranking' column includes a legend: '1=more credible' and '2=less credible'. Two rows are shown: one for 'Schwab' (http://www.schwab.com) and one for 'ShareBuilder' (http://www.sharebuilder.com). Each row has radio buttons for ranking '1' and '2'. The 'Schwab' row has the '1' radio button selected. To the right of each row is a text area for comments. A 'Continue' button is located at the bottom right of the interface.

Web Site	Credibility Ranking 1=more credible 2=less credible	Please share your comments
Schwab http://www.schwab.com	1 <input checked="" type="radio"/> 2 <input type="radio"/>	
ShareBuilder http://www.sharebuilder.com	1 <input type="radio"/> 2 <input type="radio"/>	

Next, Part 2 presented the 1-to-10 ranking task on one page (study page 8). For each site, the subject selected the 1-to-10 rank from a drop-down menu. Error checking prevented the subject from leaving the page without assigning unique numbers to each of the 10 sites. Again, comment fields were provided for each site.

Portion of the screen for Part 2

CONSUMER WEBWATCH

Part 2 of 4: Rank and comment upon the sites

Now that you have compared five pairs of finance sites, please order the same ten sites from most to least credible, with 1 being the most credible site and 10 being the least credible. Do not assign more than one site to each ranking. This can be a difficult task, so please provide any new comments that explain your overall ordering decisions. You do not need to repeat comments that you provided earlier. This part should take about 10 minutes.

Web Site	Credibility Ranking 1=most credible 10=least credible	Please share your comments
Christian Brothers Investment Services, Inc. http://www.cbis-fsc.com/index.asp	Rank this site	
ChoicePicks http://www.choicepicks.com	Rank this site	
Domini Social Investments http://www.domini.com	Rank this site	
E-Trade http://us.etrade.com/e/t/home	Rank this site	

Next, Part 3 of the study introduced the variable ranking task and explained the 7-point ranking scale (study page 9). The 30 variables were presented on the next two pages (study pages 10 and 11).

Portion of a sample screen for Part 3

CONSUMER WEBWATCH Instructions

What makes a web site in the Finance area believable to you?
Page 1 of 2

1 The site has search capabilities.
 Much Less Believable -3 -2 -1 0 +1 +2 +3 Much More Believable

2 The site displays an award it has won.
 Much Less Believable -3 -2 -1 0 +1 +2 +3 Much More Believable

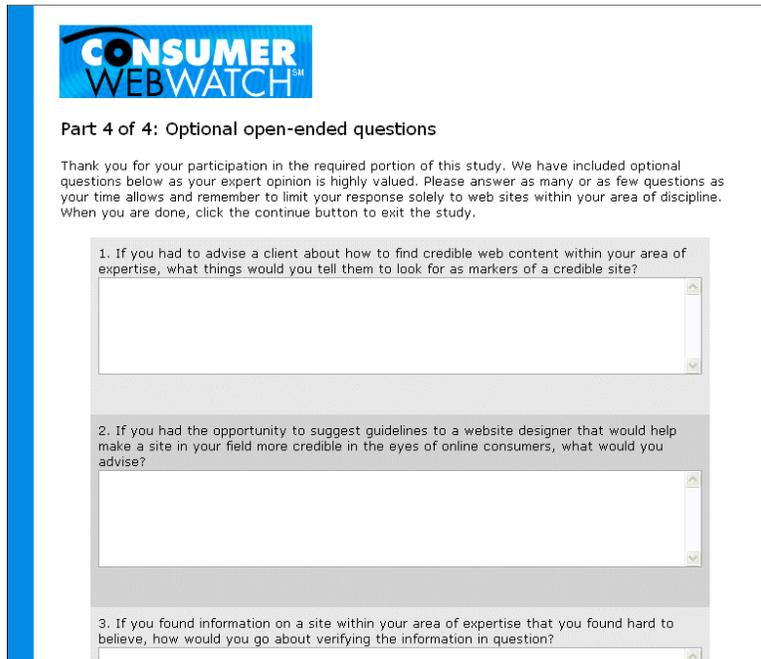
3 The site takes a long time to download.
 Much Less Believable -3 -2 -1 0 +1 +2 +3 Much More Believable

4 The site makes it hard to distinguish ads from content.
 Much Less Believable -3 -2 -1 0 +1 +2 +3 Much More Believable

5 The site contains content that differs from your opinions.
 Much Less Believable -3 -2 -1 0 +1 +2 +3 Much More Believable

Next, Part 4 (study page 12) presented five optional questions. Each subject was thanked for his or her participation in the required portion of this study, and asked to answer as many or as few optional questions as his or her time allowed.

Portion of the screen for Part 4 (Finance study shown)



CONSUMER WEBWATCH

Part 4 of 4: Optional open-ended questions

Thank you for your participation in the required portion of this study. We have included optional questions below as your expert opinion is highly valued. Please answer as many or as few questions as your time allows and remember to limit your response solely to web sites within your area of discipline. When you are done, click the continue button to exit the study.

1. If you had to advise a client about how to find credible web content within your area of expertise, what things would you tell them to look for as markers of a credible site?

2. If you had the opportunity to suggest guidelines to a website designer that would help make a site in your field more credible in the eyes of online consumers, what would you advise?

3. If you found information on a site within your area of expertise that you found hard to believe, how would you go about verifying the information in question?

Finally, the last page of the survey (study page 13) thanked the subjects again. The data for each subject was then stored in a database for analysis.

ANALYZING THE DATA

After the testing period ended, we reviewed the data to ensure that it was complete and reasonable. For completeness, we confirmed that the database showed valid information for each subject and for each task. For reasonableness, we analyzed the data to ensure that it made sense and noted any discrepancies that required discussion. This task varied by section, but included such things as comparing the order of each subject's pair rankings with his or her 1-to-10 rankings.

Once we were assured that the data fields looked valid, we analyzed the data to understand the results. First, we quantitatively analyzed the order of ranked sites and variables and calculated an average response for each group of experts. Next, qualitative data, which includes site-specific comments and answers to optional questions, required comprehensive review. In order to classify the comments and to compare them to the parallel Stanford PTL 2002 consumer study, we coded each qualitative comment based on a list of content codes developed by the Stanford PTL; we added one category, *information source*. The categories are explained in Figure 1 below.

Figure 1: Description of comment categories

Comment category	Description
Advertising	Comments relating to how users perceive advertising on the site
Company Motive	Comments relating to the perceived motive, good or bad, of the organization behind the site
Design Look	Comments relating to the look and visual appeal of the site
Information Bias	Comments relating to the perceived bias of information in the site
Information Design	Comments relating to how the information is structured or organized on the site, including navigation and site organization
Information Focus	Comments relating to the scope or focus of the site, including the quantity of information available
Information Accuracy	Comments relating to the accuracy of the information on the site that does not mention information sources
Information Source	Comments relating to the citation of sources
Name Recognition and Reputation/Affiliation	Comments relating to name recognition of the site or the reputation of the site's operator, affiliates, supporters, or partners
Writing Tone	Comments relating to the tone or attitude conveyed by the site's content

Each comment was assigned one or more of the 24 codes and could have been positive or negative with respect to the code. An example content code was “CMN,” standing for “company motive negative,” which means that the comment mentioned a concern that the motive of the company was negative. See Appendix D for the complete list of content codes with sample categorized comments. In addition, we grouped some individual codes into meaningful categories. The previous table defines the categories of expert comments that are referred to in the Results and Discussion section of this paper. Once all of the comments were coded, we calculated the frequency of each category. Thus, we are able to understand how often comments addressed a particular subject.

PART 1: PAIR-RANKING TASK ANALYSIS

The analysis of the results varied for each part of the study. Part 1, the pair rankings, provided quantitative and qualitative data based on pairs of sites. We read and coded qualitative data as described above. Quantitatively, the ranking system stored “+1” in the database when a site was ranked the more credible of the pair and “-1” when a site was ranked the less credible. We used the +1/-1 scores to calculate a mean score for each site. For example, a site that won all of its

pairings would have a final mean score of +1; winning half the pairings would result in a score of 0; and losing all of the pairings would result in a score of -1.

The parallel Stanford PTL 2002 consumer study created site rankings by using the mean score to order the sites in each category. However, as discussed in the Methods section above, the small number of pairings in this expert study and the random pairing method did not necessarily lead to an accurate ranking order of the sites for the experts. In order to understand the effect of the random pairings, we reviewed which sites were paired with which and found that indeed some of the pairings were “unfair.” For example, ShareBuilder was paired exclusively against sites that did well overall, and thus its pair ranking was lower than its 1-to-10 ranking from Part 2. In future studies, we can consider using a different pairing algorithm that seeks to match all pairs before allowing duplicates to get a wider range of comments.

PART 2: 1-TO-10 RANKING TASK ANALYSIS

Part 2, the 1-to-10 rankings, also provided quantitative and qualitative data on all of the sites. We reviewed and coded qualitative comments as described above. Quantitatively, the ranking system stored a number from 1 to 10 in the database for each site for each subject. Using the 1-to-10 numbers, we calculated the mean ranking for each site. See the Results and Discussion section, Figures 2 and 4, for the sites in each category ranked from 1 to 10.

DATA CONSISTENCY BETWEEN PART 1 AND PART 2

In order to verify the data consistency for each subject, we checked to see that each expert’s answers from the pairings agreed with his or her final 1-to-10 rankings. For example, a discrepancy existed if a subject rated site A more credible than site B in the pair rankings, but then ranked site B higher than site A in the 1-to-10 rankings. We refer to these dissonant occurrences as “reversals.” Of the 75 total pairs (15 subjects with five pairs each), there were seven pairs reversed. In all except two of these reversed cases, the subject ranked the sites adjacent to each other in the 1-to-10 ranking, thereby making the discrepancy fairly insignificant. Thus, there were only two pairings of the 75 with notable discrepancies, and in those two cases the subjects’ comments did not explain their rationale for changing the relative ranking of a site from Part 1 to Part 2.

We also scored how much each expert was in agreement with the other experts in his or her group. We desired a high level of agreement, meaning consistency among expert opinions; conformity shows both that the expert point of view was stable and that the results were meaningful with the number of experts we included in the study. We devised a rough calculation for agreement based on a comparison between an individual expert’s pair ranking of each site with the group’s 1-to-10 ranking of the site. If an expert rated site A overall more credible than site B in the pair task, and site A did better in the group 1-to-10 rankings, the expert got a point. With 10 sites in five pairs, the maximum score was five. Thus, if the expert’s pair rankings completely agreed with the group’s rankings, the expert would have scored a 5. The scores were as follows: finance: 5, 5, 5, 5, 4, 4, 3; health: 5, 5, 5, 5, 4, 4, 3, 3. In summary, we used this coarse measurement to understand whether our experts were generally in agreement, and we found that they were.

PART 3: VARIABLES

Part 3 provided quantitative data on how different elements of a Web site in the expert's field affected its believability. As discussed in the Study Design section, each item received a score from -3 (makes the site much less believable) to +3 (makes the site much more believable). Using those scores, we computed a mean for each variable, both overall and for health and finance categories separately. Next, we used those means to make two comparisons: the differences between the two groups of experts, and the differences between the experts and the consumers from the Stanford PTL 2001 study, which was described in the Methods section above.

To compare differences between the groups of experts, we noted items that were interestingly different, which we defined as a difference of at least 1 point in the mean rankings of the two groups. To compare differences between the experts and a large group of consumers, we compared the expert mean rankings to those of the consumers' mean rankings from the Stanford PTL 2001 study. These consumer mean rankings were based on 1,400 consumer assessments of Web credibility variables as described in the Study Design section on page 8. Ninety-five percent confidence intervals were computed from these consumer responses and compared to the means for each group of expert participants in our study. For the experts, in all but 5 of the 60 mean variable scores (each group of experts contributed 30 mean scores calculated for each of the 30 variables), there was a statistically significant difference from the general population at the 95% confidence interval ($p \leq .05$). However, while most of the items were statistically different, we identify and discuss only items that have greater practical difference using our 1-point-difference rule. These results are discussed below in the Variable Rankings Results and Discussion section.

PART 4: OPTIONAL QUESTIONS

The optional questions provided qualitative comments that we read and coded according to the coding scheme described in the Analyzing the Data section above. Thirteen of the 15 subjects -- 7 health experts and 6 finance experts -- answered the optional questions. The interesting themes found in these responses are discussed in the Results and Discussion section below.

RESULTS AND DISCUSSION

PARTS 1 AND 2: SITE RANKINGS RESULTS AND DISCUSSION

This section presents and discusses the results of the site rankings gathered in Parts 1 and 2 of the study. Figures 2 and 4 show the mean results from both ranking tasks, presented in order of the final 1-to-10 ranking from Part 1. The pair ranking from Part 1 of the study is displayed on the right side of the table for comparison with the final 1-to-10 ranking. Note that these final results are *not* the 10 *most* credible health and finance sites on the Web in each category, but only a ranking of the sites chosen for this study and judged by this small group of experts.

We did not include a table ordered by the results from the pair rankings from Part 1 of the study because, as discussed in the Study Design section, we believe that the expert pair rankings are insufficient to produce robust results. Since pairings used in Part 1 were random, and there was a

small sample size, not all sites were paired equally with other sites, and there were some duplicate pairings. ShareBuilder, for example, was ranked tenth in the pair rankings and improved to seventh in the 1-to-10 rankings. A review of the sites that ShareBuilder was paired with in Part 1 reveals that it was always matched with sites that were ranked more credible in the Part 2 (1-to-10) rankings. Therefore, in the final 1-to-10 rankings, ShareBuilder was eventually ranked ahead of two other sites with which it was never paired. An investigation of the other sites with a difference in the rankings between Parts 1 and 2 also revealed “unfair” pairings similar to the situation that occurred with ShareBuilder.

HOW HEALTH EXPERTS RANKED SITES

Figure 2: Health-site rankings

Final 1-to-10 Ranking	Web Site Name	Average 1-to-10 Ranking	Pair Ranking
1	NIH	1.0	1
2	MayoClinic.com*	2.9	2
3	WebMD	3.6	3
4	InteliHealth	3.8	3
5	MDChoice	5.0	3
6	Dr. Koop	6.5	6
7	HealthWorld	7.4	8
7	Dr. Weil	7.4	8
9	Oxygen Health and Fitness	8.4	7
10	Health Bulletin	9.1	10

*To distinguish from MayoClinic.org, which is the Web site for the Mayo Foundation and its associated clinics and hospitals, we refer to MayoClinic.com with the .com succeeding its name throughout this document.

In the health ranking, the site for the National Institutes of Health (NIH) was ranked number one by every participant in the expert study. Next, MayoClinic.com, WebMD, and InteliHealth all had fairly close mean rankings. These sites were perceived as generally credible but not as credible as the NIH site. MDChoice ranked right in the middle with a mean score ranking of 5.0, not quite making it into the top group of sites, yet not suffering from the credibility issues of the bottom five. The next group of sites -- Dr. Koop, HealthWorld Online, Dr. Weil, and Oxygen Health and Fitness -- comprise the bottom end of the rankings by our panel of experts. None of these sites completely lacked credibility, but each suffered at least one major flaw that is described in detail in this section. Finally, Health Bulletin was ranked at the bottom: Five of our eight health experts ranked this site last. To understand why each site fared as it did in the rankings, the following is a brief review of our experts' comments for each site.

1. NIH

Our panel of experts ranked the NIH first due to its “*sterling*” reputation; one expert characterized it as “*the gold standard.*” The NIH also scored high because of its “*lack of self-interest,*” *said one.* Another comment summarized the overall expert opinion of NIH: “*I don’t*

have to look at their Web site (although I did!) to know that I trust this site. It is more about knowing the source and the processes rather than credibility markers such as design.” Other experts pointed out that the site references peer-reviewed journals and that it is used as a source by other sites. There were no negative comments about this site.

2. MayoClinic.com, 3. WebMD, 4. IntelliHealth

Each of these sites had an affiliation with reputable sources that resulted in its favorable ranking. Experts ranked the MayoClinic.com site highly because of its reputation, commenting that it is “credible due to its affiliation with the highly regarded Mayo Clinic.” WebMD’s credibility was enhanced by its coverage of quality information from reputable sources cited for each article. This resulted in one expert’s writing, “Authorship and sometimes credentials of each posting listed. Some original material, a lot of listings from other credible (usually peer-reviewed) sites.” And, IntelliHealth gained credibility via its affiliation with the Harvard Medical School and a seal of approval from the nonprofit American Accreditation HealthCare Commission (URAC), a health organization and Web site accreditation agency. For-profit WebMD and IntelliHealth ranked slightly below the commercial site for the not-for-profit Mayo Clinic Foundation (MayoClinic.com) because some experts had slight concerns about its company motive. For example, one expert wrote of IntelliHealth, “Though they proudly display the Harvard logo, the URAC Accreditation Seal and many other trust marks, I would still be wary as they are commercially driven.”

5. MDChoice

MDChoice is an interesting middle-ranked site where the comments clearly reflected its midpoint ranking. Our experts did not seem to *distrust* the content on this site, because it was either reviewed by outside sources or included a direct link to credible outside sources. Instead, our experts ranked MDChoice lower than others because it seemed to contain only material culled from other sites. One expert explained his ranking as follows: “Material culled mostly from other reasonable sources and apparently reviewed by editorial board. Appears to be less original material and review than #2 & 3 ranked choices.” This lack of original content paired with suspicions “about a company backed by venture capital money” and a large number of drug company ads caused it to be ranked below other for-profit health sites.

6. Dr. Koop

Our experts were mixed in their responses to Dr. Koop, although most said they were suspicious of its credibility because of the “many ads” or the lack of reference citations. Some cited concerns about the site’s earlier negative press, such as one who opined, “After the Dr. Koop scandal a couple of years ago, this site has no credibility whatsoever.” In the summer of 2000, Dr.Koop.com, Inc. and four corporate officers were sued in securities fraud class-action lawsuits filed by investors in Texas (USA), which alleged that the company made false promises when it began selling its stock to the public during its initial public offering. Former U.S. Surgeon General C. Everett Koop was *not* named in the litigation. The lawsuits were eventually settled during the summer of 2001.

7. HealthWorld and Dr. Weil (Tie)

HealthWorld and Dr. Weil tied in their rankings by our panel of experts, with both ranked in the bottom half because of a perception that each is too commercial in focus. The majority of our experts stated that the credibility of these sites suffered because the sites were not only providing information about alternative health products, but also *selling* these same products. *“I do not like product information appearing mid-screen (monitor) when I am browsing selecting information.... I was expecting content,”* wrote one expert in reference to HealthWorld. Several experts also mentioned a lack of references or outside review as a problem; one such comment is, *“I feel this is a suspicious Web site, with non-evidence based health information, not from evidence-based sources.”*

9. Oxygen Health

Oxygen Health ranked ninth primarily because our health experts felt it focused less on health and more on fitness and commercial articles. Experts assessing the site in our study had trouble finding real health information. One said:

“This seems to be a fitness/stress reduction/feel good site as opposed to a ‘health’ site. I thought I would attempt to find some health information. As this is a site designed for women, I figured ‘breast cancer’ information would be easy to locate. After spending some time looking at link options under the ‘health and fitness’ tab, I did a search for ‘breast cancer,’ and was not able to locate any substantive information regarding the disease.”

Some experts pointed out that the articles appeared to be *“... authored mostly by non-MDs and [there’s] no sign of subsequent review.”* However, our experts did not find information that was factually incorrect, which saved Oxygen Health from being ranked last.

10. Health Bulletin

Finally, our panel of experts ranked Health Bulletin last because it contained what one expert characterized as *“biased information from the alternative/homeopathic point of view”* presented in a flashy manner and *“without actual credible authorship present.”* There were no positive comments about this site, as our experts had a fairly unanimous negative opinion of the information. Said one: *“Numerous spelling errors. Sentences such as ‘One thing would be to find out whether your headaches are more serious medically. Get an exam. Second would be to try out different preparations from the store or from mail-order distributors. Good luck.’ make the site not very credible.”*

OVERALL HEALTH-SITE EVALUATION TRENDS

Figure 3: Percentage of health-expert comments, by category

Comment category	Percent of health-expert comments
Name/Reputation/Affiliation	43.9%
Information Source	25.8%
Company Motive	22.7%
Information Focus	19.7%
Advertising	13.6%
Design Look	7.6%
Information Bias	4.6%
Information Design	3.0%
Writing Tone	3.0%
Information Accuracy	1.5%

Figure 3 provides a quantitative summary of the types of comments the health experts made during the site-pairing and ranking tasks. The comment percentages do not add up to 100%, as some comments were coded with multiple categories and not all categories tracked appear in the table. This quantitative analysis paired with our review of the comments revealed several trends in the way these health experts judged the credibility of health sites.

First and foremost, our health experts gave the most credibility to sites that provided information from reputable sources, as illustrated by the high number of comments in the name/reputation/affiliation (43.9%) and information source (25.8%) categories. Our experts' credibility perceptions of health sites are influenced more by the expertise dimension of our credibility definition than the trustworthiness dimension (see page 7 for our credibility definition). For example, one expert explained his credibility ranking for IntelliHealth by writing, "...material written/reviewed by Harvard Medical School docs. Some info from standard source (like USP for patients)." In contrast, Dr. Koop.com's perceived lesser credibility seems to be harmed by issues of reputation and sourcing: "Authorship of individual articles difficult if not impossible to ascertain. Leadership heavy, using Dr. Koop's name, only 2 other MDs on panel listed as 'authors and experts,' none of the three appear to be practicing clinicians." Reputation and sourcing greatly influenced the final health-site rankings, which generally trended -- from highest to lowest -- from sourcing by reputable authors (NIH, WebMD), to general review (InteliHealth), to no author credentials or citations (Health Bulletin).

Sources develop a good reputation in the health field by having a history of providing quality information. Our health experts used reputation as a key indicator of credibility by assuming that such sources are motivated to continue to provide good information to protect their reputation. For example, MayoClinic.com ranked in the top tier almost purely by its past reputation, with comments such as, "Again, [this site is] credible due to its affiliation with the highly regarded Mayo Clinic. This site does not, however, appear to contain as much in-depth health information as IntelliHealth." In contrast, the reputation Dr. Weil brought to his site was countered by the

site's commercial focus: *"Dr. Weil is credible, but the site is more commercial -- i.e., trying to sell you upgrades, vitamins, etc."* This leads to the next judgment criterion: company motive.

Company motive was the third-highest commented area by our health experts (22.7%). Our experts felt that health sites should operate with the interests of the readers as their first priority, not their balance sheet. One expert summarized this viewpoint, writing, *"I find health Web sites that sell or market products less credible than those that relay information only."* Our experts' rank reflected this credibility trend by rating the U.S. government's NIH, which works solely for the benefit of people's health, at the top of the credibility scale. The overall ranking then proceeds -- from highest to lowest -- from commercial interests that were presented in a non-sales-oriented manner (WebMD and IntelliHealth), to sites where our experts had concerns about commercial motive driving the content (Dr. Weil), and finally to sites where products were sold in line with the content, making it difficult for users to distinguish ads from editorial content (Health Bulletin).

While credibility could be tainted by commercial motive and product sales, our experts' comments did not indicate that a profit motive alone precluded credibility. In fact, according to our experts, if the site has deep, sourced information, it could still be for-profit and remain mostly credible. This was the case with WebMD, which ranked third: *"Good overall site, but I'm always cautious about publicly traded companies underwriting health Web sites and any possible ulterior motives to the content."* In the case of Dr. Weil, a few experts commented that they looked for sourced information to counteract the product sales, but could not find it, which thereby hurt its credibility among our health panel of experts. One expert wrote, *"Mostly, this site is about selling vitamins. There's no research that I can find."*

Finally, from the perspective of surface credibility criteria such as writing and visual design, our health experts were concerned with language-presentation issues such as editing, poor grammar, and typos. Although they did not base their credibility assessments purely on the quality of the visual design (7.6%), they were wary of sites that had a flashy design. For example, one expert commented that Health Bulletin *"seems 'flashy' which makes you question the information it contains -- is it really sound scientifically or just the latest sound bite?"* Writing tone comprised only 3.0% of the comments, but when this area was mentioned it was typically associated with a very negative affect on credibility. One expert commented about Health Bulletin: *"Besides suffering from poor editing (which makes me further question their education and authority), this site seems to present material in a sensationalized manner. Their reviews under the 'fountain of youth' are 'PROVEN' and from 'peer-reviewed journals' ... but no citations are provided."*

HOW FINANCE EXPERTS RANKED SITES

Figure 4: Finance-site rankings

Final 1-to-10 Ranking	Web Site Name	Average 1-to-10 Ranking	Pair Ranking
1	Fidelity	1.9	1
1	Schwab	1.9	1
3	Merrill Lynch	3.7	3
4	E-Trade	3.9	6
5	Domini Social	5.6	3
6	Scotttrade	6.3	3
7	Christian Brothers	7.3	6
7	ShareBuilder	7.3	10
9	ChoicePicks	8.1	9
10	Stocks at Bottom	9.1	8

Unlike the health sites, Figure 4 shows that our panel of finance experts did not clearly rank one site first, and they produced two other very close rankings. According to the finance experts in our study, Fidelity Investments and Charles Schwab tied for the top position. Third-place Merrill Lynch and fourth-place E-Trade had very close mean scores. Christian Brothers Investment Services and ShareBuilder tied for seventh place in our experts' rankings. The rest of the Web sites were evenly spread out with Stocks at Bottom ranked last. In fact, six of our seven finance experts ranked it last. To understand why each site fared as it did in the rankings, what follows is a brief review of expert comments for each site.

1. Fidelity and Schwab (Tie)

Fidelity and Schwab tied for the top ranking from our finance experts and received similar positive comments about the amount of information both sites offered. Our experts commented that Fidelity was up-front about its expertise in mutual funds and provided a *“focus on education and risk management,”* said one. Others noted Fidelity's stellar reputation, such as one expert who wrote, *“Fidelity doesn't need to use the site to build credibility, as its reputation is established.”* Similarly, Schwab received praise for its focus on education, such as: *“Less focus on transaction costs and more focus on better investment decisions based on independent or neutral analyses. Content is the focus.”* Another expert wrote that the Schwab site was, *“one of the very best ... clear, concise, quick, all-purpose.”*

3. Merrill Lynch

Like Fidelity and Schwab, Merrill Lynch also presented a lot of information, but it was actually criticized by these experts for presenting too much. Our experts commented that the content was very broad and unfocused, with too many different audiences, but at least it did not just push too many marketing materials. One expert commented, *“Too much information for too diverse an investor demographic, which makes it harder to hone in on one's specific needs. But nonetheless,*

the info is there and it's believable." Another expert suggested that Merrill Lynch *"needs to set up a special site just for retail investing."*

4. E-Trade

Our experts ranked E-Trade higher than other discount brokerages in our study, due to its reputation and attempt at an educational offering for novice investors. However, it was ranked lower than Fidelity and Schwab because our experts thought that it focused more on marketing to investors rather than educating them. Comparing Schwab and E-Trade, one expert wrote that E-Trade *"seems more hard-sell than the Schwab site, which is more low-key and informational in tone."* Another expert who reviewed E-Trade's education section wrote that E-Trade *"should focus on investors' well-being with education and risk management. It is hidden deep within their Web site."*

5. Domini Social

While some admired Domini Social Investment's focus on social values and its clean design, our experts ranked it in the middle overall because they thought it is not a site that is appropriate for all levels of investors. However, these experts ranked Domini Social over other narrowly focused sites because it very clearly spelled out its objectives and viewpoint. One expert summarized the overall opinion with this comment: *"I admire Amy Domini & the goals of her funds, but this seems to be an odd pairing with Schwab. The site does a good job of explaining her funds and their purpose, but is not for all investors."*

6. Scottrade

Our experts called Scottrade *"a basic super-discount broker"* and ranked it lower than discount broker E-Trade because they thought it had a pushier marketing message and no investor education. One expert commented, *"I have to admit the focus at Scottrade is simply to promote frequent trading. There is no emphasis on risks associated with frequent trading."* However, Scottrade ranked higher than discount broker ShareBuilder because our experts thought Scottrade provided clearer content and explanations of their marketing strategies: *"It looks straightforward, with routine marketing material on the costs of executing trades with Scott compared with regular discount brokerages,"* said one.

7. Christian Brothers and ShareBuilder (Tie)

Like Domini Social, experts liked that Christian Brothers explicitly states the values that drive its investment policy. However, our experts ranked it lower than Domini Social because the site did not explain how its social goals translate into a specific investment strategy, and there was *"scant information"* about investing. One expert summarized this view by commenting, *"Will appeal to socially driven investors, but weak on research, investor details."*

ShareBuilder tied with Christian Brothers in the ranking, but was ranked seventh overall for different reasons. Experts in our study criticized ShareBuilder for overly promoting day trading with heavy marketing language, prompting comments such as, *"Too much focus on transaction costs. More interested in getting investors to trade more often."* Although these experts thought that ShareBuilder provided more investor education than Scottrade, they felt that the information

was biased. For example, one finance expert commented, *“There is not much explicit believable information.”*

9. ChoicePicks:

Our panel of finance experts questioned ChoicePicks’s credibility and ranked it ninth because of its flashy, pushy marketing message that *“speaks with the voice of a carnival barker urging investors to trade frequently,”* noted one. However, unlike the bottom-ranked site, these experts appreciated that, as one put it, *“the site clearly promotes its services to active traders and not for all individual investors.”*

10. Stocks at Bottom

Stocks at Bottom was ranked last by six of our seven experts because they thought it presented a biased point of view with a gaudy design. One expert summarily wrote, *“Too much glorification of successes without any reference to potential failures.”* Two subjects could not connect or had link failures on the site, which lead to their decreased-credibility opinion of the site.

OVERALL FINANCE-SITE EVALUATION TRENDS

Figure 5: Percentage of finance-expert comments, by category

Comment category	Percent of finance-expert comments
Information Focus	40.3%
Company Motive	35.8%
Information Bias	29.9%
Design Look	16.4%
Information Design	13.4%
Name/Reputation/Affiliation	10.5%
Writing Tone	10.5%
Advertising	1.5%
Information Accuracy	0.0%
Information Source	0.0%

Figure 5 provides a quantitative summary of the types of comments finance experts made during the site-pairing and ranking tasks (see page 14 for a description of the categories). This quantitative analysis paired with our review of the comments revealed several trends in the way in which finance experts judge the credibility of sites.

SITE CONTENT MOST INFLUENCED FINANCE EXPERTS

Overall, finance experts in our study based much of their credibility assessments on the breadth, depth, and transparency of information that a site presented. Their credibility perceptions relied more on a trustworthiness assessment of the Web site than on an assessment of expertise. Our experts used a site’s name/reputation/ affiliation (10.5%) as additional criteria to judge a site’s

credibility after first evaluating the site's content. The site rankings reflect how these experts evaluated each site's content for educational focus, clarity, quantity, bias, tone, and design. According to our finance experts, Fidelity and Schwab, ranked evenly at the top, provide unbiased, clear, well-designed information, and do not have much marketing-focused content. Sites in the middle of the credibility ranking have more content focused on marketing with less education, but are not overly sales-oriented. Narrowly focused sites in this midrange tended to take special care to clearly explain their biases. Finally, sites at the bottom of the ranking contain significantly more marketing content with pushy language and either nonexistent or biased educational information. The following is a discussion of these content-focused trends derived from our experts' finance site evaluations.

Above all, our finance experts expected credible sites to provide investors with a great deal of educational information and research. The largest percentage of these experts' comments (40.3%) related to the information focus or quantity of information available on a site, and many of these comments focused on educational content. For example, when explaining why Merrill Lynch was ranked over E-Trade, one expert wrote that at E-Trade, *"Investors are not anchored to learning and analyses before making investment decisions."* In contrast, Merrill Lynch *"focuses more on education and research than transaction costs."* Another expert noted that the Domini Social site *"is rather bare-bones in terms of information. Even the 'Learning Center,' which was a featured button on the home page, had nothing on it when clicked."* Many of our experts had trouble trusting a site without a substantial educational offering.

Finance experts also evaluated site content, especially educational information, for bias. Information bias was the third most-commented category by our finance experts, with 29.9% of total comments. Schwab received high marks for providing unbiased information, with comments such as: *"Schwab is clear that their recommendations are unbiased and doesn't rely on analyst 'biased' recommendations. They have done a good job communicating their intent -- that is, the best interests of the individual investors."*

Some expert comments were particularly strict about avoiding information bias even for otherwise highly credible sites. One expert, who did not rank Fidelity first, explained why:

"The biggest drawback of Fidelity's site is that is geared to promote Fidelity products. That's understandable, but it's something investors should not lose sight of because Fidelity has its share of pricey funds. For instance, the fund selector displays only Fidelity funds. Its information about 529 plans alludes to the fact that non-Fidelity products exist, but that's not highlighted. It's also important that investors understand that Fidelity is built on the concept that active fund management pays off better than investing in index funds."

Perceived information bias was one of the leading factors that lead to the low ranking of ShareBuilder by our panel of finance experts. These experts commented that ShareBuilder offered only one point of view in its content. Said one: *"I don't want to say ShareBuilder is not credible. But it's newer, and it focuses only on a sliver of investing, one which has triggered a certain amount of debate. As such, investors need to know the pros and cons of whether this style of investing is better or worse than buying mutual funds or stocks directly."*

Another criterion our finance experts used for evaluating information was clarity of purpose. Our experts required that a credible site clearly communicate its target audience, what the goals of its services are, and how these goals translate into an investment strategy. These types of comments were found in a combination of the top three comment categories (information focus, company motive, and information bias), but relate predominantly to company motive (35.8%). Clarity is especially important for sites that focus on one area of investing, such as Domini Social and Christian Brothers. Our experts noted Domini Social for how it clearly identified its social investing viewpoint. Said one: *“Perhaps because of its social goals, this investment service is exceedingly up-front with information and muted in its marketing come-on, which inspires credibility.”* In contrast, according to our experts, Christian Brothers did not have as thorough explanations of how its viewpoint relates to its investment strategy, leading to concerns. Said one: *“While the social goals of the fund may be admirable, I would’ve appreciated more about how those goals are reconciled with the investment objectives.”*

To promote user understanding of information, these experts also said sites need to have good information design (13.4%). Information overload due to perceived poor design particularly hurt Merrill Lynch’s credibility with some experts: *“Of course, Merrill Lynch serves such a broad market of investors, it has to jam EVERYTHING into the site. This makes it a bit unwieldy, too.”* In contrast, the higher-ranked Schwab, which also serves a wide investor audience, received positive comments such as, *“Clear design; useful links arranged in a compact way.”* However, good information design was not enough to overcome other content-presentation faults. ShareBuilder illustrated this, according to one finance expert, who said:

“ShareBuilder exemplifies the paradox that a simpler GUI [Graphical User Interface], while welcome, also means less information available. It seems clear that ShareBuilder is simply a discount brokerage, not a full-service investment/financial services operation. For what it is, the site is clear and easy to use. Its depth of info resources seems scant, however.”

Finally, these experts commented that finance sites need to avoid pushy marketing language and glitzy presentation. Failing in site design (16.4% of finance expert comments) or writing tone (10.5% of finance expert comments) greatly injured the credibility of the bottom two finance sites (ChoicePicks and Stocks at Bottom). Our experts were concerned with the look and feel of ChoicePicks, as one described it as a *“flashy site pushing risky strategies and gimmicky theories ... and selling software, other products.”* Likewise, Stocks at Bottom received comments such as, *“Serious people shouldn’t need gaudy colors, moving crawls (via JavaScript) and other such silliness.”* Even the highly ranked Merrill Lynch site was faulted by one expert for having too much marketing language:

“Starting with a pervasive message of how ‘bullish’ the firm is, the site uses language (sometimes off-putting) that drives home how smart Merrill Lynch representatives are. The firm also regularly tries to steer investors to complete a ‘brief checklist,’ which gives the firm information that could be used to make a sales pitch, I presume. In other words, I can’t help be wary that I’m being tugged into a sales call at every turn.”

WHAT ABOUT REPUTATION?

Some finance experts (10.5%) used a site’s name or reputation in addition to assessing the value of its content to evaluate a site’s credibility. Comments from our experts showed that higher-

ranked sites like Fidelity, E-Trade, and Domini Social were able to rely on their reputations to overcome other faults and boost overall credibility. For example, one expert wrote:

“What E-Trade has going for it is a broader scope and reputation. Make no mistake that it’s a discount brokerage that wants to encourage more frequent trading. The ads and links to products and services for the brokerage and its banking arms are inescapable. But credibility isn’t an issue that struck me.”

Another expert compared Fidelity and ShareBuilder by commenting, “...*the Fidelity brand recognition and credibility precedes it...*” However, reputation was not as primary to our finance experts as it was to our health experts. This difference is explored further in the next section.

DIFFERENCES BETWEEN FINANCE AND HEALTH EXPERTS

Differences between these finance and health experts’ assessments of credibility are related to the inherent differences in these industry categories. The primary motive of most informational health sites is to provide useful health information. While there can certainly be commercial or transactional objectives through advertising revenue or product sales, the stated goal of many informational health sites is generally benevolent. In contrast, transactional financial sites are primarily profit-motivated. While the financial sites do share the goal of financial profit with their client, most of the sites profit through fees that are dependent upon the number of transactions and amount of assets managed. Another innate contrast between health and finance sites is in the nature of the information provided. Informational health sites are generally trying to provide “proven” information resulting from scientific studies or regarding state-of-the-art medical practice. Finance sites are generally providing insights about past performance to guide decisions about purchasing financial products with an unknown future return.

These differences highlight the disparity in the guiding principles that each group of experts in this study used when evaluating sites in their fields. Returning to our definition of credibility, our *finance experts’* perceptions of credibility were more heavily influenced by the *trustworthiness component* of expertise while our *health experts* were more influenced by the *expertise component*. Finance experts’ primary concern was that financial sites present trustworthy, balanced information. This aligns with the idea that if a finance site’s motive is to make money by influencing people to buy its products, then a highly credible, well-intentioned site must provide information from multiple points of view that a consumer can balance to make a wise investment decision. If a financial site presents only one perspective, it can increase its credibility by being up-front about its motive and target audience. In contrast, if a health site’s motive is to provide the best available health information, then a credible site must provide information from “good” sources that are specifically cited. Good, credible, knowledgeable sources are generally reputable health institutions that have earned their reputations through reliable research bolstered by the goal of advancing medical knowledge.

Second, these finance experts find information design and visual presentation very important, because being able to find information easily and having a professional look helps a Web site inspire confidence and trust -- qualities that are especially important within the context of unknown long-term returns. In contrast, our health experts find that visual design is secondary to verifying the factuality of the information, which makes a non-flashy presentation and good

grammar especially important. Within these contrasts, both groups remark that an easy-to-use information design is crucial.

DIFFERENCES BETWEEN EXPERTS AND CONSUMERS

As discussed in the Methods section (see page 8), the Stanford Persuasive Technology Lab (Stanford PTL) conducted a parallel study of over 2,600 consumer rankings of Web site pairs with comments. During its study, average consumers compared 204 finance site pairs and 114 health site pairs. Figures 6 through 9 compare expert and consumer Web site evaluations.

EXPERT VS. CONSUMER HEALTH RANKINGS

Figure 6: Expert vs. consumer health-site rankings

Web Site Name	Expert Rank	Consumer Rank	Difference
NIH	1	3	-2
MayoClinic.com	2	1	+1
WebMD	3	6	-3
InteliHealth	4	2	+2
MDChoice	5	4	+1
Dr. Koop	6	5	+1
HealthWorld	7	7	0
Dr. Weil	7	8	-1
Oxygen Health	9	10	-1
Health Bulletin	10	9	+1

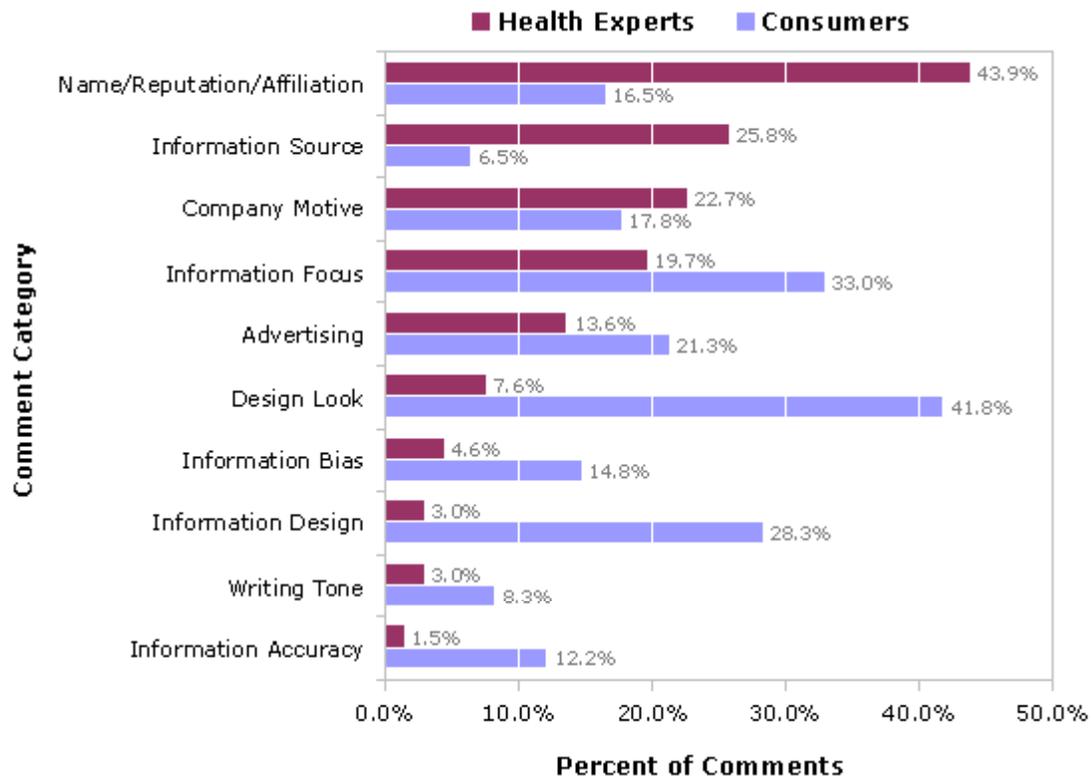
Interesting differences, defined as a difference of more than 1 point in the mean score, are marked in bold.

The most interesting difference in the health rankings is the NIH, which received a unanimous first-place ranking from our health experts and only a third-place ranking from the Stanford PTL's consumers. The consumers' first-ranked site, MayoClinic.com, was ranked second by these experts. Another significant difference is WebMD, which our experts ranked third, while the Stanford PTL's consumers ranked it sixth, not even in the top half. Also, while our experts ranked InteliHealth fourth, the Stanford PTL's consumers ranked this same site second, above even the NIH.

The two groups directly agreed on only one site, HealthWorld, which was ranked seventh overall by each group. The remainder of the health sites -- MDChoice, Dr. Koop, Dr. Weil, Oxygen Health, and Health Bulletin -- did not differ remarkably in the rankings between expert and consumer evaluators. However, it is interesting to note that Health Bulletin, which our panel of health experts noted might contain questionable information, ranked one spot above Oxygen Health among the Stanford PTL's consumers.

What accounted for some of these differences in the rankings? We answer that question by examining the comments made by the two groups. Figure 7 compares comments about the health sites assessed by these consumers and experts.

Figure 7: Expert vs. consumer health-site comments



Overall, the most striking difference in the percentage of comments is that the consumers commented much more frequently on the design look and information design of health sites than our experts (34.2% and 25.3% more, respectively). Consumers also commented 13.3% more frequently than our experts about information focus. In contrast, health experts valued name/reputation/affiliation and information source significantly more than consumers (27.4% and 19.3% more, respectively). We discuss these differences in detail below.

Health-site design was much more influential for consumers than for experts when judging credibility. Consumers commented on health-site design look and information design a combined 59.5% more than these experts, a tremendous difference. One representative consumer explained his ranking of IntelliHealth over MayoClinic.com as follows: *“I preferred IntelliHealth and I cannot really say why. I think the appearance of the page setups or colors. It just seemed more believable to me.”* Another consumer assessed that WebMD was less credible than HealthWorld because the *“home page is so brief. Just a black-and-white image of a girl looking into a mirror.”*

While the consumers focused on evaluating the look of health sites, the experts focused on evaluating the reputation of the site’s owner, associations, and the specific sources of each article. Our health experts commented 27.4% more frequently on the reputation of the sites than the Stanford PTL’s consumers and 19.3% more frequently about information source. When assessing accuracy, the consumers do not primarily question the source of the information. Instead, they make a personal judgment on how accurate they believe the information is, such as, *“Weil has a*

point of view, to be sure, but he upholds scientific standards as well. His answer to the letter regarding ephedra and creatine is comprehensive and informative.” Conversely, even when the consumers reviewed the NIH site, many did not comment on its reputation. Said one consumer: “This site looks more professional and although it looks more boring, I expect a truly scientific/health sciences site to look boring. Anything too flashy can take away from the seriousness. I know a Web site needs to look good to attract visitors.” This significant and alarming difference in consumer and expert credibility evaluation must be addressed in order to help consumers accurately assess the validity of online health information.

Finally, the consumers commented 13.3% more frequently than these experts on the information focus of health sites. Although some experts noted the quantity of information on a site, most were more concerned with evaluating the sources of that information rather than the amount of it available. The consumers, however, felt more comfortable when there was a lot of information present on multiple health topics. For example, after ranking MDChoice higher than the MayoClinic.com, one consumer noted that on MDChoice, “[I] Think I would get more information on my problems.” A different consumer ranked HealthWorld higher than IntelliHealth, writing, “I enjoyed the doctors’ input on the different ailments and other alternative possibilities. I liked this site [HealthWorld] better, because it gave me more options.”

EXPERT VS. CONSUMER FINANCE RANKINGS

Figure 8: Expert vs. consumer finance-site rankings

Web Site Name	Expert Rank	Consumer Rank	Difference
Schwab	1	1	0
Fidelity	1	2	-1
Merrill Lynch	3	4	-1
E-Trade	4	6	-2
Domini Social	5	5	0
Scottrade	6	7	-1
Christian Brothers	7	9	-2
ShareBuilder	7	3	+4
ChoicePicks	9	8	+1
Stocks at Bottom	10	10	0

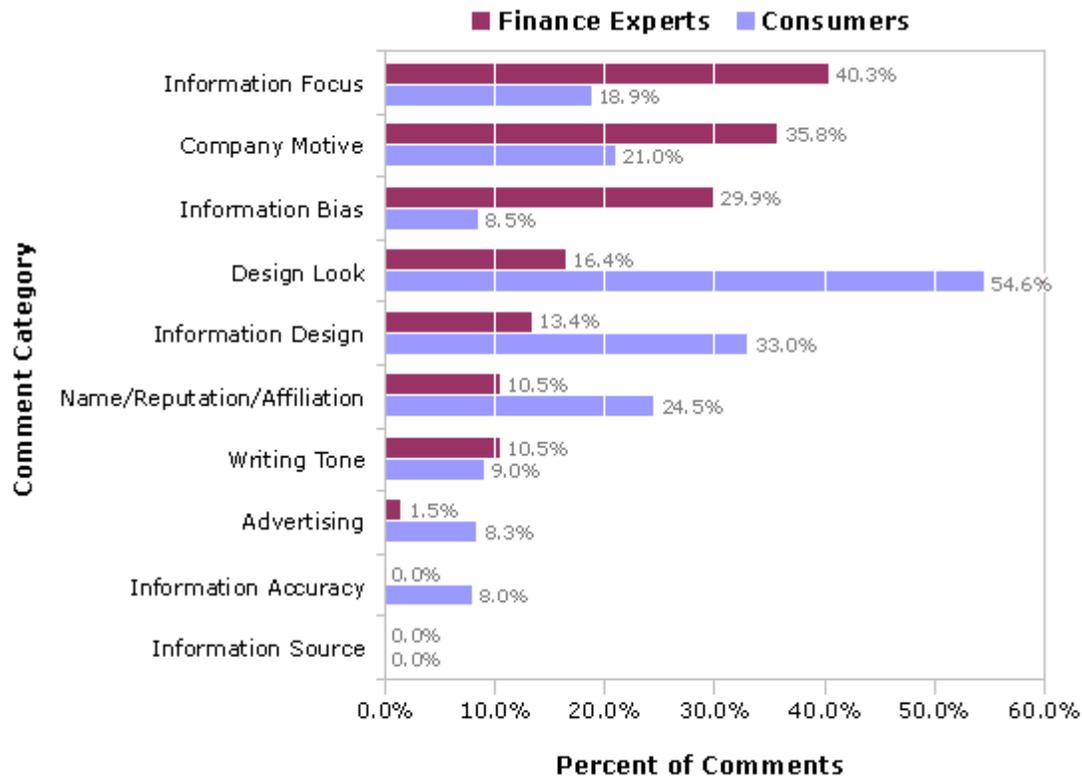
Interesting differences, defined as a difference of more than 1 point in the mean score, are marked in bold.

In the finance category, Schwab and Fidelity tied for first place with our experts. The Stanford PTL’s consumers, however, ranked Schwab higher than Fidelity in their rankings, with an average ranking of +.57 for Schwab vs. +.28 for Fidelity (see Part 1: Pair-Ranking Task Analysis section on page 14 for a description of these numeric values). The consumers actually ranked Fidelity closer to Merrill Lynch, which had an average ranking of +.20. Another big surprise was the consumers’ third-ranked site, ShareBuilder, which was ranked seventh by our experts in a tie with Christian Brothers. If ShareBuilder and its affect on the rankings were removed, the remaining sites would be ranked very similarly by the two groups. The only remaining differences would be the pairs of E-Trade/Domini Social and ChoicePicks/Christian Brothers, which each differed by one rank position for consumers vs. experts. All other sites (Merrill

Lynch, Scottrade, and Stocks at Bottom) had the same ranking across both groups.

Figure 9 compares comments from consumers and experts on finance sites to aid in understanding the ranking differences. Consumer and expert evaluators of finance sites differed greatly in their credibility evaluation criteria. First, the consumers commented a lot more than these finance experts on design look, information design, and name/reputation/affiliation. In contrast, our finance experts focused more than the consumers on in-depth content evaluation, including assessment of the information focus, company motive, and information bias.

Figure 9: Expert vs. consumer finance-site comments



The largest difference between consumer and expert finance site evaluation was in the predominant consumer use of surface markers of credibility, mainly design look and information design issues. Notably, the consumers reviewing finance sites made more comments on site design (54.6%) than the average for consumers across all 10 content categories (46.1%) in the parallel Stanford PTL 2002 consumer study. Some of our finance experts did assess site design as a credibility marker, as design look and information design were the fourth and fifth most-commented-upon areas by our finance experts, respectively. However, these criteria were used for expert evaluation of finance sites in addition to a thorough review of the content, and they were mostly used to identify non-credible sites with flashy graphics. In contrast, the consumers often relied on site design as a primary marker of the content credibility. This was exemplified by ShareBuilder, which our experts ranked seventh and the Stanford PTL’s consumers ranked third. One consumer who wrote about ShareBuilder said, *“Having a smiling guy in the center makes it look like an ad for a bank, or something -- feeds off of that kind of credibility. Easy to read text,*

very organized-looking... Seems they have less to hide.” Yet another consumer commented, “This Web site is great! It really spells things out for a rookie investor. Everything was easy to find and understand.”

The next big area of difference between our finance experts and the Stanford PTL’s consumers is information focus (21.4% difference between the two groups), information bias (21.4% difference), and company motive (14.8% difference), reflecting the greater importance that our experts placed on content analysis. While these experts often cited concerns about the availability of education and presentation of balanced viewpoints, the consumers did not seem as concerned about these factors for evaluating finance sites. For example, Fidelity and Schwab tied for first place among our expert panel because of their comparable educational offerings and excellent content. However, Schwab outranked Fidelity among the consumers because of its superior visual appeal. For example, one consumer described Schwab’s site as follows: *“Very professional. To-the-point design, options expand as you seek information. Solid, proven design. Content appears accurate, due in part to smart design.”* Another consumer thought that Schwab’s site was credible because, *“The smiling guy somehow adds something to this site, I can trust them more than a slogan....”*

Finally, the consumers commented 14.0 percent more frequently on the reputation of finance sites than our finance experts. While this difference is not as significant as other areas, such as information focus and information bias, it supports the observation that finance consumers are more likely to rely on surface markers of credibility, such as known reputation, than to read and evaluate the information on a site.

PART 3: VARIABLE RANKINGS RESULTS AND DISCUSSION

In the last part of the report we talked about rankings and comments from experts and consumers. In this part of the report we turn our attention to a different part of the study: How experts responded to statements describing different site elements, such as a statement that a *site “lists the organization’s physical address.”*

The variable ranking section of the study asked experts to provide quantitative information about how different elements of a Web site affected its credibility. Figure 10 shows the results of variable score ranking ordered from items that made a site much more believable (maximum score of +3) to those that made a site much less believable (minimum score of –3). The table also shows the results for each of our industry-specific expert groups, as well as the difference between the two expert groups.

Overall, the variable that scored the most influential for making a site believable was the same for both of our expert groups: *“The site is by an organization that is well respected.”* Our finance experts gave their next-highest scores to sites that list the organization’s physical address, state a privacy policy, list authors’ credentials for each article, and provide a quick response to customer service questions. In comparison, our health experts gave their next-highest scores to sites that list authors’ credentials for each article, and those that represent a nonprofit organization. At the bottom of the list, two of the three lowest-scored *elements for both groups* were sites with pop-up ads, and sites that made it hard to distinguish ads from content. The other item that made sites least credible for these finance experts was unexpected site unavailability, and for these health experts it was sites with a commercial purpose.

Figure 10: Expert variable scores

(+3= much more believable, -3 = much less believable)

Variable text	All Experts	Finance Experts	Health Experts	Difference (Finance vs. Health)
The site is by an organization that is well respected.	2.87	3.00	2.75	0.25
The site lists authors' credentials for each article.	2.47	2.43	2.50	-0.07
The site lists the organization's physical address.	2.20	2.57	1.88	0.70
The site states its privacy policy.	2.07	2.43	1.75	0.68
The site gives a contact e-mail address.	1.87	1.57	2.13	-0.55
The site gives a contact phone number.	1.73	2.00	1.50	0.50
The site represents a nonprofit organization.	1.73	0.86	2.50	-1.64
The site provides a quick response to your customer service questions.	1.40	2.29	0.63	1.66
The site looks professionally designed.	1.40	1.71	1.13	0.59
The site is arranged in a way that makes sense to you.	1.33	1.71	1.00	0.71
The site is linked to by a site you think is believable.	1.33	1.57	1.13	0.45
The site has been updated since your last visit.	1.27	1.86	0.75	1.11
The site provides links to its competitors' sites.	1.27	1.57	1.00	0.57
The site has search capabilities.	1.13	1.57	0.75	0.82
The site was recommended to you by a friend.	1.07	1.29	0.88	0.41
The site was recommended by a news media outlet, such as a newspaper, magazine, or e-mail newsletter.	1.07	1.14	1.00	0.14
The site displays an award it has won.	1.07	1.00	1.13	-0.13
The site provides live chat with a company representative.	0.60	1.14	0.13	1.02
The site is advertised on the radio, billboards, or other media.	0.07	0.29	-0.13	0.41
The site contains content that differs from your opinions.	-0.20	0.71	-1.00	1.71
The site requires you to register or log in.	-0.27	<i>-0.14</i>	-0.38	0.23
The site requires a paid subscription to gain access.	-0.40	-0.43	-0.38	-0.05
The site takes a long time to download.	-0.93	-1.14	-0.75	-0.39
The site has a typographical error.	-1.27	<i>-1.29</i>	<i>-1.25</i>	-0.04
The site has a link that doesn't work.	-1.33	-1.57	-1.13	-0.45
The site has a commercial purpose.	-1.47	-0.14	-2.63	2.48
The site has one or more ads on each page.	-1.47	-1.00	-1.88	0.88
The site is sometimes unexpectedly unavailable.	-1.67	-2.29	-1.13	-1.16
The site automatically pops up new windows with ads.	-2.27	-2.43	-2.13	-0.30
The site makes it hard to distinguish ads from content.	-2.87	-2.86	-2.88	0.02

Interesting differences, defined as a difference of more than 1 point in the mean score, are marked in bold. Results where finance experts' responses were not statistically different from health experts' responses are marked in italics ($p \leq 0.5$).

VARIABLE SCORES OF HEALTH VS. FINANCE EXPERTS

The variable scores revealed some notable differences between our health and finance expert groups. As expected, these health experts placed more believability in a site that represented a nonprofit organization (2.5 for health experts vs. 0.86 for finance experts). While our finance experts were neutral about commercial purpose (-0.14), our health experts were very negative about it (-2.63). These variable scores were in agreement with the 1-to-10 site ranking findings.

Another difference between the two expert groups was in the “*sites that contain content that differs from your opinions*” variable, which increased credibility for these finance experts (0.71) and decreased it for these health experts (-1.00). This was an interesting difference that was also in agreement with the 1-to-10 rankings. Our finance experts wanted multiple sides of an issue presented, while our health experts wanted peer-reviewed, sourced material that presumably would not differ from the consensus of other health experts’ opinions.

A third difference was that these finance experts assigned greater credibility to sites that provided a quick response to customer service questions, were updated since a user’s last visit, provided live chat, and were not unexpectedly unavailable. These are all highly important items for timely financial transactions, making them more important for finance sites than for health.

VARIABLE SCORES OF EXPERTS VS. CONSUMERS

Expert completion of the variable scoring task enabled us to compare these experts’ responses to those of a large group of consumers, as discussed above in the Methods section. The following Figures 11 and 12 show the results, ordered by consumer rankings of the items. The results for each expert group are compared to the consumers in separate tables. A figure that combines all of this data into one table is available in Appendix F.

Figure 11: Finance-expert vs. consumer variable scores

(+3 = much more believable, -3 = much less believable)

Variable text	Consumers	Finance Experts	Difference (Experts vs. Consumers)
The site is by an organization that is well respected.	1.97	3.00	1.03
The site provides a quick response to your customer service questions.	1.83	2.29	0.46
The site lists the organization’s physical address.	1.67	2.57	0.90
The site has been updated since your last visit.	1.65	1.86	0.21
The site gives a contact phone number.	1.56	2.00	0.44
The site looks professionally designed.	1.54	1.71	0.17
The site gives a contact e-mail address.	1.47	1.57	0.10

Variable text	Consumers	Finance Experts	Difference (Experts vs. Consumers)
The site is arranged in a way that makes sense to you.	1.46	1.71	0.25
The site lists authors' credentials for each article.	1.31	2.43	1.12
The site is linked to by a site you think is believable.	1.26	1.57	0.31
The site states its privacy policy.	1.21	2.43	1.22
The site has search capabilities.	1.18	1.57	0.39
The site was recommended by a news media outlet, such as a newspaper, magazine, or e-mail newsletter.	1.14	1.14	0.00
The site was recommended to you by a friend.	1.03	1.29	0.26
The site provides links to its competitors' sites.	0.99	1.57	0.58
The site is advertised on the radio, billboards, or other media.	0.77	0.29	-0.48
The site represents a nonprofit organization.	0.66	0.86	0.20
The site provides live chat with a company representative.	0.62	1.14	0.52
The site displays an award it has won.	0.31	1.00	0.69
The site contains content that differs from your opinions.	0.11	0.71	0.60
The site requires you to register or log in.	-0.14	<i>-0.14</i>	0.00
The site has a commercial purpose.	-0.28	-0.14	0.14
The site has one or more ads on each page.	-0.60	-1.00	-0.40
The site requires a paid subscription to gain access.	-0.88	-0.43	0.45
The site takes a long time to download.	-1.00	-1.14	-0.14
The site has a typographical error.	-1.26	<i>-1.29</i>	-0.03
The site is sometimes unexpectedly unavailable.	-1.29	-2.29	-1.00
The site has a link that doesn't work.	-1.42	-1.57	-0.15
The site automatically pops up new windows with ads.	-1.64	-2.43	-0.79
The site makes it hard to distinguish ads from content.	-1.90	-2.86	-0.96

Interesting differences, defined as a difference of more than 1 point in the mean score, are marked in bold. Results where expert responses were not statistically different from consumer responses are marked in italics ($p \leq 0.5$).

Figure 12: Health-expert vs. consumer variable scores

(+3 = much more believable, -3 = much less believable)

Variable text	Consumers	Health Experts	Difference (Experts vs. Consumers)
The site is by an organization that is well respected.	1.97	2.75	0.78
The site provides a quick response to your customer service questions.	1.83	0.63	-1.21
The site lists the organization's physical address.	1.67	1.88	0.21
The site has been updated since your last visit.	1.65	0.75	-0.90
The site gives a contact phone number.	1.56	1.50	-0.06
The site looks professionally designed.	1.54	1.13	-0.42
The site gives a contact e-mail address.	1.47	2.13	0.66
The site is arranged in a way that makes sense to you.	1.46	1.00	-0.46
The site lists authors' credentials for each article.	1.31	2.50	1.19
The site is linked to by a site you think is believable.	1.26	1.13	-0.14
The site states its privacy policy.	1.21	1.75	0.54
The site has search capabilities.	1.18	0.75	-0.43
The site was recommended by a news media outlet, such as a newspaper, magazine, or e-mail newsletter.	1.14	1.00	-0.14
The site was recommended to you by a friend.	1.03	0.88	-0.16
The site provides links to its competitors' sites.	0.99	1.00	0.01
The site is advertised on the radio, billboards, or other media.	0.77	-0.13	-0.90
The site represents a nonprofit organization.	0.66	2.50	1.84
The site provides live chat with a company representative.	0.62	0.13	-0.50
The site displays an award it has won.	0.31	1.13	0.82
The site contains content that differs from your opinions.	0.11	-1.00	-1.11

Variable text	Consumers	Health Experts	Difference (Experts vs. Consumers)
The site requires you to register or log in.	-0.14	-0.38	-0.24
The site has a commercial purpose.	-0.28	-2.63	-2.35
The site has one or more ads on each page.	-0.60	-1.88	-1.28
The site requires a paid subscription to gain access.	-0.88	-0.38	0.51
The site takes a long time to download.	-1.00	-0.75	0.25
The site has a typographical error.	-1.26	<i>-1.25</i>	0.01
The site is sometimes unexpectedly unavailable.	-1.29	-1.13	0.17
The site has a link that doesn't work.	-1.42	-1.13	0.30
The site automatically pops up new windows with ads.	-1.64	-2.13	-0.49
The site makes it hard to distinguish ads from content.	-1.90	-2.88	-0.98

Interesting differences, defined as a difference of more than 1 point in the mean score, are marked in bold. Results where expert responses were not statistically different from consumer responses are marked in italics ($p \leq 0.5$).

Experts and consumers ranked the same variable, “*The site is by an organization that is well respected,*” first, and also agreed on the two bottom answers, “*The site automatically pops up new windows with ads,*” and “*The site makes it hard to distinguish ads from content.*” The one item that all three -- consumers and *both* expert groups -- scored the same was “The site has a typographical error,” which somewhat diminished credibility for all groups (-1.2).

Interesting differences, which we define as a difference of more than one point in the mean score, were as follows:

- Sites of well-respected organizations deemed more credible:
While consumers and our finance experts both ranked a well-respected organization first, finance experts assigned a greater increase in credibility (3.00) than consumers (1.97) or health experts (2.75). This was an interesting difference as the percentage of consumer comments about this area during their ranking of the same finance sites (24.5%) was significantly greater than comments from finance experts (10.5%).
- Customer service more concerning to consumers and finance experts:
Consumers ranked quick response to customer service questions second most important (1.83), while finance experts ranked this fourth most important (2.29). Health experts ranked it seventeenth (0.63). This difference is understandable given that informational health sites of the type we chose to study generally have less of a need for customer service.

- Information sources and author expertise lends credibility:
Listing authors' credentials for each article on a site added credibility for consumers (1.31), but not as significantly as it did for experts (2.43 for finance experts and 2.50 for health experts). This suggests that experts might value the source of information and expertise much more than the average Web user.
- Presence of a privacy policy boosts credibility:
Finance experts scored a statement-of-privacy policy much higher (2.43) than did the consumers (1.21) or our health experts (1.75). Interestingly, neither the experts nor the consumers mentioned privacy in any of their comments during the ranking tasks.
- Health experts assign more credibility to non-profit sites:
Health experts assigned nonprofit sites much more importance (2.50) than did either our consumers (0.66) or our finance experts (0.86). This finding is consistent with the conclusion that our health experts tend to value a site with a noncommercial motive.
- Health experts more concerned about content that differs from their opinions:
Health experts were more concerned with site content that differed from their own opinions (-1.00) than consumers (0.11). Finance experts were in the middle of the two groups (0.71). This agreed with the overall health panel's assessment that, as experts, they tend to have correct knowledge of available health information.
- Health experts more concerned about commercial purpose:
Health experts were more concerned with a site's commercial purpose (-2.63) than consumers (-0.28). The consumer score was similar to that of the finance experts' score (-0.14).
- Too many online ads a credibility detractor for health sites:
Health experts were more concerned with the presence of one or more ads on a site (-1.88) than the consumers (-0.60) or the finance experts (-1.00). This agrees with the comments we received from our health experts who summarily stated that they are less likely to believe a site's information if it has the potential of being biased by advertiser influence.
- Unexpected site unavailability tarnishes credibility:
While all three groups ranked a site that was sometimes unexpectedly unavailable as a credibility detractor, finance experts were especially critical of this item (-2.29), compared to consumers (-1.29) and to health experts (-1.13). This finding suggests that unexpected site unavailability regarding financial matters is perhaps a more serious problem than the unexpected unavailability other types of sites.

PART 4: OPTIONAL QUESTIONS RESULTS AND DISCUSSION

After completing all the site rankings and variable scoring, we asked experts to answer five optional, open-ended questions that allowed us to identify any issues that did not arise in their earlier comments. In general, we found that our experts' answers agreed with their previous comments and did not uncover any new issues. The following is a review of responses gathered from these questions. Please see Appendix C for the full text of each question.

Questions 1 and 2 asked the experts to recommend guidelines, in their area of expertise, for an end user to find credible Web content, or for a site designer to create a credible site. Most finance experts advised that sites should avoid marketing language, make the company motive clear, and present information in an unbiased way. Interestingly, they did not suggest that a site should provide a lot of consumer education information, even though this is something they often mentioned during the ranking tasks. Consistent with their comments in the rankings, however, health experts suggested that health sites should cite sources and be affiliated with reputable medical-related organizations.

Question 3 addressed an area that was not covered elsewhere in the study: how an expert would verify dubious information found on a Web site. Our finance and health experts both said that they would search on the Web for verifying information, send an e-mail query to the original Web source, or check with a reputable third party, such as the U.S. Securities and Exchange Commission, the U.S. National Library of Medicine's MEDLINE, or another industry expert.

Question 4 asked the experts to specifically comment on surface presentation credibility markers of sites such as graphics, navigation, or grammar. These finance and health experts both commented that the most important thing is that information should be easy to find. Many health experts referred to language presentation, stating that sites should make sure they have good editing, grammar, and thorough proofreading for typographical errors. They also suggested that sites should refrain from using flashy graphics, but never specifically mentioned professional graphic design. In contrast, our finance experts thought that finance sites should have a clear, professional visual presentation yet never mentioned any grammar-related concerns.

Finally, Question 5 asked how the owner, sponsor, or advertiser of a site affected a site's credibility. Responses to this question were consistent with comments in the ranking sections. Both groups of experts thought that an ideal site should have no advertisements. But if sponsors or advertising is needed, it must be clearly separated from the content, not lead to any editorial bias, and not detract from the purpose of the site. Our health experts felt that national or academic sponsors such as the U.S. government or a university would have the best access to good medical-related research with the least information bias. Our finance experts thought that the reputation of the company was more important than what ads or sponsors were displayed on a site.

CONSUMER EDUCATION

This study has revealed a few areas in which further consumer education can address the credibility perception differences between experts and consumers. In critical areas such as health and finance, consumers need to better discern when they should be skeptical of information. Most importantly, consumers need to be educated that for any Web site, visual design should not be the

primary marker of credibility. Although it is true that some less-credible sites in our study have flashy, distracting graphics, it is not inversely true that clean, professional-looking sites are necessarily credible. This was exemplified by ShareBuilder, a site with a noted attractive design that resulted in a high credibility rating among the Stanford PTL's consumers yet content that made the finance experts on our panel skeptical.

For health sites in particular, consumers should be aware that many people are inclined to evaluate Web site credibility using visual design. By recognizing this inclination, a consumer can seek to focus on more substantive credibility criteria, such as the source of information. Reputable, nonprofit sources with no agenda or bias, other than the distribution of information, are the best sources of health information. Just as consumers would not go to a doctor without a medical degree, they should not accept information that does not come from a qualified, reputable source. Furthermore, just because a site contains a great quantity of health-related information does not mean that that information is in fact reliable or credible.

For finance, the opposite recommendation is true: Users should remain skeptical of finance information even if they recognize the source. The majority of financial sites in our study are trying to sell something. Even socially responsible finance sites, such as Domini Social, are promoting a specific agenda. Consumers must evaluate the content themselves to verify that the site presents a well-rounded discussion of each financial issue. If a site does not, they should increase their knowledge base by visiting multiple sites in order to gather and understand multiple viewpoints before making a financial decision.

To summarize, we propose that consumer education should cover three main points:

- Visual design should not be the key marker of credibility.
- Credibility of health sites should be based on an assessment of the expertise; seek information from a qualified, reputable source.
- Credibility of finance sites should be based on an assessment of the trustworthiness of the content; seek well-rounded information.

GUIDELINES AND DESIGN IMPLICATIONS

We are hopeful that education will encourage consumers to better evaluate credibility by employing the judgment criteria used by the experts. Consequently, we propose expert-inspired credibility guidelines for health and finance sites to illuminate the credibility of their content in meaningful ways.

ANALYSIS OF EXISTING CONSUMER WEBWATCH GUIDELINES

Consumer WebWatch has proposed a set of five general guidelines for improving credibility on the Web (see <http://www.consumerwebwatch.org/bestpractices/index.html>). The following is an analysis of the existing guidelines with suggestions for refinement, based on the results of this study.

1. Identity

This guideline states that Web sites should disclose their physical location, contact information, ownership, purpose and mission. Of all the Consumer WebWatch guidelines, this category was the most often mentioned by both finance and health experts, which is an indication of its importance. All Web sites, but particularly the transaction-heavy finance sites, need to clearly identify their purpose and mission and inform the user about any site-specific bias in their content. Also, given our finance experts' high scoring of the credibility variable regarding "physical address," finance sites should also consider placing contact information, including physical address, in the footer of every page. For health sites, where our experts revealed that reputation is more influential than in finance, sites need to clearly identify any associations with credible organizations. In addition, health sites should strive to include the names and credentials of its health or medical writers as well as list their professional affiliations, which our health experts have shown leads to a site's credibility. Consumer WebWatch should consider adding the identification of any associations or affiliations, in addition to ownership, to this identity guideline.

2. Advertising and Sponsorships

This guideline states that sites should clearly distinguish ads from site content and disclose paid placements, sponsors, and relevant business relationships, including sponsored links to other sites. This policy is particularly important for health sites, where our experts were suspicious of sites with clear commercial motives, especially when ads were hard to distinguish from the content. Consumer WebWatch guidelines should be augmented to address the inherent marketing and advertising promotional content found on finance sites, which is often not presented as a "traditional" ad. Instead, users need to be aware of "content-based" advertising that presents "articles" that solely promote the company's point of view, and thus skews the consumer toward its own product lines or services.

3. Customer Service

This guideline states that sites involved in e-commerce transactions should clearly disclose their relevant financial relationships with other sites. They should also disclose all fees charged before a transaction begins, and state and enforce product or service return and cancellation policies. This guideline is more important for the finance sites we chose for this study because they are primarily transactional in nature. Although finance experts did not directly mention customer service in their comments, they did, however, rank this issue highly in the variable rankings. This same guideline can and should be applied to health sites with transactional components, such as Dr. Koop and Dr. Weil that provide online stores selling vitamins and other supplements.

4. Corrections

This guideline states that sites should seek to correct false information, display a page or section where incorrect information is corrected or clarified, and state their policy on consumers' rights if a purchase is made based on previously published incorrect information. Content corrections were not specifically mentioned by any experts in the study. In fact, responses to optional Question 3 indicated that our experts were more likely to check inaccuracies with the article author or a third-party source than to contact the site itself. That said, citing the article source and providing contact information for the author (with his or her credentials) should be included for each published article or written reference, in addition to the posting of corrections or

clarifications on the site when necessary. No expert in this study commented on the portion of this guideline that addresses online purchases, and thus we cannot address that part of the guideline.

5. Privacy

This guideline states that a site should have an easy-to-find, clearly stated privacy policy that discloses how personal data, cookies, and other information-collection mechanisms will be used; sites should notify users of changes to the privacy policy. Although both groups of industry experts ranked a privacy policy high on the importance scale in the variable portion of the study, neither group mentioned the importance of privacy in their comments. In fact, Christian Brothers, HealthWorld, and Health Bulletin did not have privacy policies at the time of this study, yet this was not mentioned as a reason for ranking these sites lower than others in their respective areas. Thus, privacy is an important concept, but not the first place that Web users -- whether novice or expert professionals -- look to assess credibility, perhaps because privacy policies have become a pervasive Web standard.

ADDITIONAL FIELD-SPECIFIC GUIDELINES

In addition to the existing Consumer WebWatch guidelines, we also suggest the following additional guidelines for health and finance sites:

Guidelines for health sites

- Cite author-specific sources for all information.
 - Provide the author's name and affiliation for all published information.
 - Use reputable sources (from well-known institutions, peer-reviewed journals, etc.) when possible.
 - Employ a review board of known experts to review content from lesser-known sources.
 - Although site ownership is important, it is perhaps even more important to identify the exact source of each article posted.
- Present a noncommercial face.
 - If the site has a commercial focus, avoid marketing-based articles, and explain how your for-profit interests (product sales, sponsors, etc.) may affect the site's content.
 - Use detailed information from unbiased, reputable, third-party sources to provide supporting information about products for sale.
 - Do not allow excessive, inline, or pop-up ad placement; ads include both in-house and third-party product promotions.
- Edit the content of your site carefully to insure good grammar and spelling.
- Avoid flashy design that detracts from the serious health content.

Guidelines for finance sites

- Explain the site's commercial goals, target audience, and potential biases.

- Provide articles that describe multiple viewpoints on financial issues
 - Clearly explain financial product recommendations.
 - Provide education to help investors make informed decisions.
 - If your site has a specific agenda, explain the agenda itself and how it might affect the site's product recommendations or services provided.
- Try to avoid, or keep to a minimum, having your site unexpectedly unavailable, so that people can complete timely financial transactions.
- Provide quick, effective customer service, including direct-contact information and reasonable hours of operation.

FURTHER RESEARCH

This study of health and finance experts successfully revealed how each expert group appraises the credibility of Web sites in their domain. The ability of these health and finance experts to produce consistent, robust data suggests that similar studies in other areas can also produce Web site credibility and education guidelines for a variety of fields. Future expert studies can leverage existing work by selecting from the other eight areas already examined in the parallel consumer study (e-commerce, entertainment, news, nonprofits, opinion or review, search engines, sports, and travel). Or one may choose among the many specific areas that present key credibility challenges, such as results of browser searches, financial advice, or automotive sales.

This study also suggests another direction for research: consumer education. We would like to understand the most effective ways of educating the general population about methods for credibility assessment. In order for consumers to effectively evaluate Web site credibility, we must understand how to change people's methods of credibility assessment. We hope these findings support the efforts of librarians and other information professionals who currently take on the enormous task of educating the public about how to find quality information on the Web. As we move away from a pen-and-paper society to more of a paperless daily environment, we suspect these efforts will become even more of a career goal for informational professionals worldwide.

APPENDIX A: SITES WITH WEB ADDRESSES

Health Sites

Dr. Koop	http://www.drkoop.com *
Dr. Weil	http://www.drweil.com
Health Bulletin	http://www.healthbulletin.org
HealthWorld Online	http://www.healthy.net
InteliHealth	http://www.intelihealth.com
MayoClinic.com	http://www.mayohealth.org **
MDChoice	http://www.mdchoice.com
National Institutes of Health	http://www.nih.gov
Oxygen Health and Fitness	http://www.oxygen.com/health
WebMD	http://www.webmd.com

* Note that the sale of one health site, DrKoop.com, to a Florida company was announced on July 15, 2002, and the site changed significantly because it is no longer affiliated with former U.S. Surgeon General C. Everett Koop. This study was completed before the announced sale.

** Note that the Web address that we used for MayoClinic.com, www.mayohealth.org, automatically redirects to www.mayoclinic.com. We refer to this site as MayoClinic.com throughout this paper to distinguish from MayoClinic.org, which is the Web site for the Mayo Foundation and its associated clinics and hospitals.

Finance Sites

ChoicePicks	http://www.choicepicks.com
Christian Brothers Investment Services	http://www.cbis-fsc.com
Domini Social Investments	http://www.domini.com
E*Trade	http://www.etrade.com
Fidelity Investments	http://www.fidelity.com
Merrill Lynch	http://www.ml.com
Charles Schwab	http://www.schwab.com
Scottrade	http://www.scottrade.com
ShareBuilder	http://www.sharebuilder.com
Stocks At Bottom	http://www.stocksatbottom.com

APPENDIX B: LIST OF SUBJECTS

Health Experts

J. Sybil Biermann, MD

Associate Professor of Orthopaedic Surgery, University of Michigan

Dr. J. Sybil Biermann is the Director of Orthopaedic Oncology at the University of Michigan. Her research interests include quality of information on the Internet and the impact of the Internet on the patient-physician relationship.

Susan J. Coburn

Consultant and expert in the online health care industry

Susan J. Coburn has over 20 years experience in the health care industry and is currently the author of a weekly newspaper column reviewing health care sites on the Internet.

Gunther Eysenbach, MD

Associate Professor at the Department of Health Policy, Management, and Evaluation, University of Toronto;

Senior Scientist at the Centre for Global eHealth Innovation, Toronto General Hospital, and Division of Medical Decision Making and Health Care Research, Toronto General Research Institute

Dr. Gunther Eysenbach, who has been dubbed "cyberphysician" by the *British Medical Journal*, has authored more than 90 publications, including several pioneer studies and comments on e-health and consumer health informatics published in respected international journals such as the *JAMA*, the *BMJ*, and *The Lancet*. He is principal investigator of two European Union projects to evaluate and improve the quality of health information on the Web, MedCERTAIN and MedCIRCLE.

Howard Fuller

Librarian, Stanford Health Library, Stanford Hospital & Clinics

Howard Fuller is the Librarian of the Stanford Health Library. He has been involved in consumer and patient health libraries in both the private and nonprofit sectors for over 10 years. Prior to his work in libraries, Mr. Fuller was a health and wellness educator.

Cynthia Helphingstine, PhD, MBA

President, The Biotron Group, Inc.

Dr. Cynthia Helphingstine is the President and founder of The Biotron Group, Inc., a consulting firm serving health care and biotechnology companies. Since its formation in 1986, The Biotron Group has successfully completed numerous assignments that required assessing the market for new medical devices, diagnostics, pharmaceuticals and life science instrumentation systems. Prior

to forming The Biotron Group, Dr. Helplingstine spent nine years with Baxter Healthcare Corporation.

Ellen Offner

Health Care Executive, Boston, Massachusetts

Ellen Offner is director of health plans, finance, marketing, and planning in a university medical center. Previously, she was Vice President for Medicare programs and, earlier, Vice President for Product Development at a large HMO.

Ronni Sandroff

Health editor, Consumer Reports; Editor, Consumer Reports on Health

Ronni Sandroff is health editor of *Consumer Reports* magazine and chief editor of *Consumer Reports on Health*, a monthly 12-page newsletter. Before coming to *Consumer Reports* three years ago, she wrote for such consumer publications as *The New York Times Magazine*, *Working Woman*, *Sports Illustrated*, and *McCall's* and was editor of several publications for health professionals, including *Oncology Times*. Ms. Sandroff has also published a novel, two short novellas, and numerous short pieces of fiction.

Daniel Z. Sands, MD, MPH

Clinical Director of Electronic Patient Records and Communication, Beth Israel Deaconess Medical Center;

Assistant Professor of Medicine, Harvard Medical School

Daniel Sands is an internationally recognized lecturer, consultant, and thought leader in the area of clinical computing and patient and clinician empowerment through the use of computer technology. In 1998, Dr. Sands received the President's Award from the American Medical Informatics Association in recognition of his leadership in coauthoring the first national guidelines for the use of e-mail in patient care. Since 1998, he has maintained the Electronic Patient Centered Communication Resource Center at <http://www.e-PCC.org>.

Finance Experts

Carolyn Bell, CFP

Vice President and Principal, Stanford Investment Group, Inc.

Carolyn Bell is an owner and a principal of Stanford Investment Group, Inc., a Registered Investment Advisor and broker-dealer located in Mountain View, California. SIG provides financial planning and investment management services to wealthy professionals in Silicon Valley.

David Bollier

Author, activist and consultant on civic implications of digital technologies

David Bollier is Senior Fellow at the Norman Lear Center at the USC Annenberg School for Communication, and cofounder of Public Knowledge, a new advocacy organization that champions the public's stake in copyright, the Internet and digital technologies. He is also author of *Silent Theft: The Private Plunder of Our Common Wealth* (Routledge).

Bill Gannon

Editorial Director, Financial Engines, Inc.

Bill Gannon is the Editorial Director of Financial Engines Inc., of Palo Alto, California. Prior to joining Financial Engines, Mr. Gannon was a national correspondent for Newhouse Newspapers and *The Star-Ledger* of Newark, New Jersey, where he covered international conflict and terrorism, and national politics, and was an investigative reporter. As a journalist, he received a number of national awards and honors including The John S. Knight Fellowship at Stanford University.

Prabhudev Konana, PhD

Associate Professor of MIS;

*Assistant Director, Center for Research in Electronic Commerce,
McCombs School of Business, The University of Texas at Austin*

Dr. Prabhudev Konana is a recipient of numerous research awards and Grants, including the prestigious NSF CAREER Award and NSF Information Technology Research grant. His wide range of research interests include study of beliefs and attitudes of online investors, determinants of satisfaction in online investing, and next generation electronic brokerages. Dr. Konana's research is cited in popular media including CNNfn, *Business 2.0*, *InternetWeek*, and Bloomberg Financial Network.

Lynn Langway

Adjunct Assistant Professor of Journalism, New York University;

Editorial Consultant

Lynn Langway is a former *Newsweek* business writer and senior editor. She is now an adjunct assistant professor of journalism at New York University and an editorial consultant whose clients have included *Fortune Small Business*, *Business Week Enterprise*, *Working Woman*, and *Women.com*.

Mark Schwanhausser

Personal finance reporter and columnist, San Jose Mercury News

Mark Schwanhausser writes about money, which includes how to save it, invest it and keep it from the taxman. In addition to mapping out the by-the-numbers solutions, he explores the psychological barriers that influence how you manage your money. He writes "The Money Manual" column, which appears every other Sunday in the *San Jose Mercury News* (San Jose, California).

Bill Whyman

President, Precursor Group

Bill Whyman is President of the Precursor Group, which serves many of the United State's largest institutional investors. He is also the firm's technology strategist. Precursor is an independent research company that does no investment banking, trading or asset management and whose analysts are not allowed to trade individual stocks. Precursor is a founding member of the Investorside Research Association.

APPENDIX C: OPTIONAL QUESTIONS

1. Health: If you had to advise a patient about how to find credible Web content within your area of expertise, what things would you tell them to look for as markers of a credible site?

Finance: If you had to advise a client about how to find credible Web content within your area of expertise, what things would you tell them to look for as markers of a credible site?

2. If you had the opportunity to suggest guidelines to a Web site designer that would help make a site in your field more credible in the eyes of online consumers, what would you advise?
3. If you found information on a site within your area of expertise that you found hard to believe, how would you go about verifying the information in question?
4. How does the surface presentation of a site affect the credibility? For example, consider the site's visual design, navigation, and grammar.
5. In your area of expertise, in what way does a Web site's owner, or the presence or absence of a sponsor or advertiser, affect your view of the site's credibility?

APPENDIX D: CODING CATEGORIES

The following coding categories table is reproduced in its exact form from the Stanford Persuasive Technology Lab's 2002 Consumer Study report with one exception: We added the "Information Source" category for coding the expert responses to our study and also recoded the consumer responses accordingly to include this category. All example comments in the table below are from the consumers who participated in the PTL study.

Issue & Explanation	Codes Used	Examples from the consumer data (Note: For some comments we've corrected typos and punctuation.)
Identity Comments relating to Consumer WebWatch Guideline #1, which addresses identity issues.	IP Identity positive	<ul style="list-style-type: none"> <i>This Web site has clear information about its company and policies.</i> <i>I could easily find contact information, a physical address, and phone number, which would help me to verify their legitimacy.</i> <i>Very transparent and informative, not only about investment possibilities but also about philosophy.</i>
	IN Identity negative	<ul style="list-style-type: none"> <i>I do not know enough about them to feel safe giving them my money.</i> <i>There was no phone number to call.</i> <i>They weren't really specific about what they do, other than some broad generalities.</i>
Sponsorship Comments relating to Consumer WebWatch Guideline #2, which addresses issues of sponsorship.	SP Sponsorship positive	<ul style="list-style-type: none"> <i>Tells who the sponsoring organizations are.</i> <i>This search engine makes it clear what is paid advertising and what is not.</i>
	SN Sponsorship negative	<ul style="list-style-type: none"> <i>The "sponsors" and ads made me a bit skeptical.</i> <i>Too difficult to tell sponsored ads from other hits retrieved!</i>
Customer Service Comments relating to Consumer WebWatch Guideline #3, which addresses issues of customer service. Also, comments relating to how an organization operated were coded in this category.	CSP Customer service positive	<ul style="list-style-type: none"> <i>They take pains to let you know how their service works.</i> <i>I like the business being up-front about costs, etc.</i> <i>Useful to the point of defining customer service issues. And consumer complaints.</i>
	CSN Customer service negative	<ul style="list-style-type: none"> <i>This site seemed to have less accountability to its customers on the items that can be purchased.</i> <i>I don't like sites where you can't see exactly what you are paying for.</i> <i>I would like to find something that will tell me that the quality of services is high.</i>

<p>Corrections Comments relating to Consumer WebWatch Guideline #4, which addresses issues of correcting false information.</p>	<p>COR</p>	<p>[There were no comments on this issue.]</p>
<p>Privacy Comments relating to Consumer WebWatch Guideline #5, which addresses issues of privacy.</p>	<p>PP Privacy positive</p>	<ul style="list-style-type: none"> • <i>Had a privacy statement.</i> • <i>Liked the Privacy Policy.</i> • <i>I liked the large "Security" button with info re: shared user data.</i>
	<p>PN Privacy negative</p>	<ul style="list-style-type: none"> • <i>No privacy policy evident.</i> • <i>Various cookies materialized.</i> • <i>I hate having to log in and give information when I'm searching for tickets.</i>
<p>Design Look Comments relating to the look of the site.</p>	<p>DLP Design look positive</p>	<ul style="list-style-type: none"> • <i>I like the look of the Web site. Looks professional.</i> • <i>The layout is cleaner.</i> • <i>At first I thought that this was going to be the more credible site, with its layout and design, but after reading a few reviews I discovered that it is not.</i>
	<p>DLN Design look negative</p>	<ul style="list-style-type: none"> • <i>Looks very unprofessional.</i> • <i>Seems flashy, like a used-car salesman.</i> • <i>Graphics are of an earlier generation of e-commerce UI [User Interface].</i>
<p>Information Design Comments relating to how the information is structured or organized on the site.</p>	<p>IDP Info design positive</p>	<ul style="list-style-type: none"> • <i>It is well organized and has a straightforward design--you understand it within seconds.</i> • <i>This is more organized.</i> • <i>Everything the Web site visitor needed was right on the front page.</i>
	<p>IDN Info design negative</p>	<ul style="list-style-type: none"> • <i>Core information (such as contact details) can take a little while to find.</i> • <i>Too difficult to navigate.</i> • <i>Information badly presented. They try to put everything on the front page.</i>
<p>Information Focus Comments relating to the scope or focus of the site.</p>	<p>IFP Info focus positive</p>	<ul style="list-style-type: none"> • <i>Somehow looks as if more information is being offered at this site.</i> • <i>Well focused.</i> • <i>This site was totally devoted to news and the dissemination of information.</i>

	<p>IFN Info focus negative</p>	<ul style="list-style-type: none"> • <i>This site was interesting, but too broad.</i> • <i>There was not a whole lot of information on these pages.</i> • <i>Too much at a glance.</i>
<p>Information Accuracy Comments relating to the accuracy of the information on the site that do not mention information source.</p>	<p>IAP Info accuracy positive</p>	<ul style="list-style-type: none"> • <i>Information sounds credible and corroborates what little I know about this area of work.</i> • <i>This site listed more facts.</i> • <i>The information is generally solid.</i>
	<p>IAN Info accuracy negative</p>	<ul style="list-style-type: none"> • <i>Not necessarily providing me with the most accurate or best information.</i> • <i>The data is harder to verify.</i> • <i>I don't believe in market timing that is advocated on this site.</i>
<p>Information Source Comments relating to the citation of sources.</p>	<p>ISP Info source positive</p>	<ul style="list-style-type: none"> • <i>Articles peer-reviewed by credible doctors.</i> • <i>List of authors and their credentials.</i> • <i>Cites sources presented.</i>
	<p>ISN Info source negative</p>	<ul style="list-style-type: none"> • <i>He could do a better job of citing his references.</i> • <i>Limited bibliographies with article and credentials of authors are not listed.</i> • <i>Unfounded, inconsistent quotations with no links to source.</i>
<p>Information Bias Comments relating to the perceived bias of information on the site.</p>	<p>IB Information biased</p>	<ul style="list-style-type: none"> • <i>I feel their view is colored by their desire to boost their advertising revenue.</i> • <i>Obvious slant in types of stories covered as well as headlines.</i> • <i>Sentimentalist jingoistic crap.</i>
	<p>IU Information unbiased</p>	<ul style="list-style-type: none"> • <i>It is credible because the opinions contained therein are based on unbiased research.</i> • <i>Students speaking in an uncensored forum.</i> • <i>Heavily political, represents the left point of view, but reports fairly and opinion is reasonable.</i>
<p>Information Usefulness Comments relating to the usefulness of the information on the site.</p>	<p>IUP Info usefulness positive</p>	<ul style="list-style-type: none"> • <i>This Web site provided useful and interesting knowledge about events in sports.</i> • <i>I found this site very helpful and informative. I will visit it again to get information I need for myself and my young son.</i> • <i>This site is great if you're going to travel. You can get the "lay of the land."</i>
<p>Experts vs. Online Consumers: A Comparative October 2002</p>	<p>Credibility Study of Health and Finance Web Sites</p>	<p>Page 51 of 59</p>

	<p>IUN Info usefulness negative</p>	<ul style="list-style-type: none"> • <i>There appears to be no useful information on this Web site.</i> • <i>I didn't see what sense the page was unless it was purely for fun.</i> • <i>This site was not very useful other than the bulletin board.</i>
<p>Information Clarity Comments relating to how the site's content is (or is not) clear and understandable.</p>	<p>ICP Info clarity positive</p>	<ul style="list-style-type: none"> • <i>Clear, concise information on home page.</i> • <i>This Web site is great! It really spells things out for a rookie investor. Everything was easy to find and understand.</i> • <i>Presents information clearly and cleanly.</i>
	<p>ICN Info clarity negative</p>	<ul style="list-style-type: none"> • <i>I found this site unclear and confusing.</i> • <i>Nothing on the site but computer language.</i> • <i>The Web site talks about making things easy for investors but then speaks over most people's heads.</i>
<p>Readability Comments relating to the site's readability-- how easy or hard it was to read what was on the pages.</p>	<p>ETR Easy to read</p>	<ul style="list-style-type: none"> • <i>The layout and the graphics are easy to read.</i> • <i>The format was easier for me to read and follow along with the stories.</i> • <i>Easy to read.</i>
	<p>HTR Hard to read</p>	<ul style="list-style-type: none"> • <i>First blush rates it better on credibility, but hard-to-find and hard-to-read critics ratings make it appear a poor use of my time.</i> • <i>Type too small. Ad smack in the middle of page, with editorial material, was repugnant. Headline over picture of Tiger Woods should have extended over type, too, or type should have been moved up.</i> • <i>The page is not easily readable. The font "Courier" contributed to this.</i>
<p>Currency of Information Comments relating to how current (up to date) the information is on the site.</p>	<p>CRP Currency positive</p>	<ul style="list-style-type: none"> • <i>Their information seems to be more up-to-date.</i> • <i>This one would appear to be updated more often and kept current.</i> • <i>Contains date of update.</i>
	<p>CRN Currency negative</p>	<ul style="list-style-type: none"> • <i>The photos of the children were from the mid '90s; they should be more up-to-date.</i> • <i>No update date.</i> • <i>Somewhat dated material, it seems, which gives more time for spinning news and event reports.</i>

Writing Tone Comments relating to the tone or attitude conveyed by the site's content	TWP Tone of writing positive	<ul style="list-style-type: none"> • <i>The headlines were dry and to the point. The Web site had so much to it that it felt "trustworthy."</i> • <i>They have a friendly, no-nonsense, straightforward tone to their page and their letter.</i> • <i>Restrained language and lack of buzz and flash make this a more credible site.</i>
	TWN Tone of writing negative	<ul style="list-style-type: none"> • <i>There was too much sarcasm when it came to headlines.</i> • <i>"Holy Crap" and other slang or poor language harms credibility. Credible people tend to understate.</i> • <i>Seems more optimistically oriented, more like advertising than actual review.</i>
Site Functionality Comments relating to how the site functions, both technical performance and the services the site offers.	FP Functionality positive	<ul style="list-style-type: none"> • <i>Downloads quickly.</i> • <i>Always uses Inktomi, therefore more consistent, therefore more credible.</i> • <i>It was the most believable because of the little movie blip they had on the site.</i>
	FN Functionality negative	<ul style="list-style-type: none"> • <i>I find the site more credible in theory but it's awfully buggy. When I entered the city of Rome, Italy, it gave me hotel info about Nice, France.</i> • <i>The site was down for site maintenance.</i> • <i>The news stories section did not show any stories even though I tried it twice.</i>
Performance on Test by User Comments that tell about a test the user performed to evaluate the site's credibility.	PTP Performed test positive	<ul style="list-style-type: none"> • <i>I searched for MRI and was able to view images and relevant Web pages.</i> • <i>Did 2 searches. Both yielded most relevant sites for the subject at top of list.</i> • <i>This site is interesting, I found a fare on this site that was actually cheaper than what I found on the Web site of the airline itself! I bookmarked this one and will use it again.</i>
	PTN Performed test negative	<ul style="list-style-type: none"> • <i>I attempted to access information on shingles at this site and no information was available.</i> • <i>I looked for Anaplastic Astrocytoma. The site had only a couple sites to look at, and they didn't give a whole lot of info.</i> • <i>It was the least believable because it could not find a coat of armor for my last name.</i>

Past Experience with Site Comments relating to previous experiences people had with the site under evaluation.	PEP Past experience positive	<ul style="list-style-type: none"> <i>I've used it so much that I know and trust this organization.</i> <i>I can ALWAYS find anything I need at this site! Have never had difficulties here.</i> <i>I have used this Web site in the past and have found it to have sound information.</i>
	PEN Past experience negative	<ul style="list-style-type: none"> <i>I am familiar with them and usually find their critiques to be accurate but not portrayed in the best way.</i> <i>I've used this site before and it did not meet my expectations.</i>
Name Recognition and Reputation Comments relating to name recognition of the site or the reputation of the operator.	NRP/RP Name recognition positive/ Reputation positive	<ul style="list-style-type: none"> <i>Maybe because it is a well-known outfit it looks to me to be more credible.</i> <i>This is a more credible site. Not because of simple name recognition but because having a local branch gives this company an image of being more real.</i> <i>Name recognition gives this site the edge. Otherwise my appraisal was almost equal.</i>
	NRN/RN Name recognition negative/ Reputation negative	<ul style="list-style-type: none"> <i>This site is less credible because the name is unfamiliar.</i> <i>Both look credible, but this isn't a brand name, the other is.</i> <i>I simply don't know anything about this company. I would have to do more research on the company before I trusted it with my business.</i>
Advertising Comments relating to how users perceive advertising on the site.	AP/AV Advertising positive	<ul style="list-style-type: none"> <i>Site did not have pop-ups.</i> <i>Banner ads on the side are more discreet and help make the site feel independent.</i> <i>There are ads along the page but they don't all jump out at you and if you are really interested in them you can go check them out, but you don't lose track of why you came to the site.</i>
	AN Advertising negative	<ul style="list-style-type: none"> <i>Site had pop-ups and spam in my face. This makes it less appealing and less credible.</i> <i>This site is full of ads, the home page is dominated by the Need Cash promo. Advertising is bound to affect objectivity.</i> <i>Sold out their site to advertising.</i>

Motive of Organization Comments relating to the perceived motive--good or bad--of the organization behind the site.	CMP Company motive positive	<ul style="list-style-type: none"> • <i>501C3 status makes this organization appear more credible.</i> • <i>This site is just about opinions and isn't trying to sell anything. Much more of a chance that people are giving honest opinions here since money isn't involved.</i> • <i>I would trust this site because it's run by a religious denomination whose aim is socially responsible investing.</i>
	CMN Company motive negative	<ul style="list-style-type: none"> • <i>Doesn't seem credible when they give a product a good review and give you a link to order it, too.</i> • <i>This is an entertainment Web site? More like an ad-dollar-driven method of driving people into lame venues. Not likely.</i> • <i>The fact that this one is done by an insurance company makes it seem less credible.</i>
Affiliations Comments relating to the site's affiliates, supporters, or partners.	ASP Association positive	<ul style="list-style-type: none"> • <i>Affiliation with a prestigious university adds to a sense of objectivity.</i> • <i>Backed by CNN</i> • <i>Links to other major sites. Amazon, Honda. Would these sites lend their credibility to a scam site? Not likely.</i>
	ASN Association negative	<ul style="list-style-type: none"> • <i>Lately, anything related to religion and money sends up a red flag.</i> • <i>I tend to distrust anyone or thing that brags about any awards right off the get-go. The J.D. Powers awards seem to be passed out like Halloween candy.</i> • <i>I don't believe much that GW Bush says, so his presence on the Web site detracts from its credibility.</i>
General Suspicion General comments about being suspicious of the site or its operators.	JPS Just plain suspicious	<ul style="list-style-type: none"> • <i>Sounds almost too good to be true.</i> • <i>I wouldn't put either in the "credible" category. . . . I would be suspect of both.</i> • <i>Because it involves buying and selling I tend to believe that their system is bound to have faults.</i>
General Dislike General comments about not liking the site or its operators	JPD Just plain dislike	<ul style="list-style-type: none"> • <i>About as cheap and nasty as any form of publication in any media can be, crass and corporate, avoid like the plague.</i> • <i>I don't like Web sites that exploit women's bodies.</i> • <i>The American "entertainment industry" generally fails to entertain me.</i>

APPENDIX E: TABLES OF SITE-PAIRING SCORES

Health-site rankings with pair-ranking score

Final 1-to-10 Ranking	Web Site Name	Pair Ranking	Average Pair Score
1	NIH	1	+1.0
2	MayoClinic.com	2	+0.5
3	WebMD	3	+0.3
4	InteliHealth	3	+0.3
5	MDChoice	3	+0.3
6	Dr. Koop	6	0.0
7	HealthWorld	8	-0.5
7	Dr. Weil	8	-0.5
9	Oxygen Health	7	-0.3
10	Health Bulletin	10	-1.0

Finance-site rankings with pair-ranking score

Final 1-to-10 Ranking	Web Site Name	Pair Ranking	Average Pair Score
1	Fidelity	1	+1.0
1	Schwab	1	+1.0
3	Merrill Lynch	3	+0.1
4	E-Trade	6	-0.1
5	Domini Social	3	+0.1
6	Scottrade	3	+0.1
7	Christian Brothers	6	-0.1
7	ShareBuilder	10	-1.0
9	ChoicePicks	9	-0.7
10	Stocks at Bottom	8	-0.4

APPENDIX F: TABLE OF EXPERT VS. CONSUMER VARIABLE SCORES

(+3 = much more believable, -3 = much less believable)

Variable text	Consumers	Finance Experts	Difference (Finance vs. Consumers)	Health Experts	Difference (Health vs. Consumers)
The site is by an organization that is well respected.	1.97	3.00	1.03	2.75	0.78
The site provides a quick response to your customer service questions.	1.83	2.29	0.46	0.63	-1.21
The site lists the organization's physical address.	1.67	2.57	0.90	1.88	0.21
The site has been updated since your last visit.	1.65	1.86	0.21	0.75	-0.90
The site gives a contact phone number.	1.56	2.00	0.44	1.50	-0.06
The site looks professionally designed.	1.54	1.71	0.17	1.13	-0.42
The site gives a contact e-mail address.	1.47	1.57	0.10	2.13	0.66
The site is arranged in a way that makes sense to you.	1.46	1.71	0.25	1.00	-0.46
The site lists authors' credentials for each article.	1.31	2.43	1.12	2.50	1.19
The site is linked to by a site you think is believable.	1.26	1.57	0.31	1.13	-0.14
The site states its privacy policy.	1.21	2.43	1.22	1.75	0.54
The site has search capabilities.	1.18	1.57	0.39	0.75	-0.43
The site was recommended by a news media outlet, such as a newspaper, magazine, or e-mail newsletter.	1.14	1.14	0.00	1.00	-0.14
The site was recommended to you by a friend.	1.03	1.29	0.26	0.88	-0.16
The site provides links to its competitors' sites.	0.99	1.57	0.58	1.00	0.01
The site is advertised on the radio, billboards, or other media.	0.77	0.29	-0.48	-0.13	-0.90
The site represents a nonprofit organization.	0.66	0.86	0.20	2.50	1.84
The site provides live chat with a company representative.	0.62	1.14	0.52	0.13	-0.50
The site displays an award it has won.	0.31	1.00	0.69	1.13	0.82
The site contains content that differs from your opinions.	0.11	0.71	0.60	-1.00	-1.11

Variable text	Consumers	Finance	Difference	Health	Difference
		Experts	(Finance vs. Consumers)	Experts	(Health vs. Consumers)
The site requires you to register or log in.	-0.14	<i>-0.14</i>	0.00	-0.38	-0.24
The site has a commercial purpose.	-0.28	-0.14	0.14	-2.63	-2.35
The site has one or more ads on each page.	-0.60	-1.00	-0.40	-1.88	-1.28
The site requires a paid subscription to gain access.	-0.88	-0.43	0.45	-0.38	0.51
The site takes a long time to download.	-1.00	-1.14	-0.14	-0.75	0.25
The site has a typographical error.	-1.26	<i>-1.29</i>	-0.03	<i>-1.25</i>	0.01
The site is sometimes unexpectedly unavailable.	-1.29	-2.29	-1.00	-1.13	0.17
The site has a link that doesn't work.	-1.42	-1.57	-0.15	-1.13	0.30
The site automatically pops up new windows with ads.	-1.64	-2.43	-0.79	-2.13	-0.49
The site makes it hard to distinguish ads from content.	-1.90	-2.86	-0.96	-2.88	-0.98

Interesting differences, defined as a difference of more than 1 point in the mean score, are marked in bold. Results where expert responses were not statistically different from consumer responses are marked in italics ($p \leq 0.5$).

REFERENCES

- Brownlow, S., & Zebrowitz, L. Facial appearance, gender, and credibility in television commercials. *Journal of Nonverbal Behavior*, 14.1 (Spring 1990): 51-61.
- The NPT 100. *The NonProfit Times* (November 1, 2001). Retrieved May 7, 2002, from <http://www.nptimes.com/Nov01/charts.pdf>
- Consumer WebWatch Guidelines. *Consumer WebWatch*. Retrieved September 18, 2002 from <http://www.consumerwebwatch.org/bestpractices/index.html>
- Fogg, B.J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., Paul, J., Rangnekar, A., Shon, J., Swani, P., & Treinen, M. What Makes Web Sites Credible? A Report on a Large Quantitative Study. *Proceedings of ACM CHI 2001 Conference on Human Factors in Computing Systems* (New York: ACM Press, 2001): 61-68.
- Fogg, B.J., Soohoo, C., Danielson, D., Marable, L., Stanford, J., & Tauber, E. How Do People Evaluate a Web Site's Credibility? Results from a Large Study. *Consumer WebWatch*. (2002). Retrieved October 28, 2002 from <http://www.consumerwebwatch.org>
- Fogg, B.J., & Tseng, H. The Elements of Computer Credibility. *Proceedings of ACM CHI 99 Conference on Human Factors in Computing Systems, v. 1* (New York: ACM Press, 1999): pp. 80-87.
<http://www.acm.org/pubs/articles/proceedings/chi/302979/p80-fogg/p80-fogg.pdf>
- Klemm, B., Iding, M., & Speitel, T. Do scientists and teachers agree on the credibility of media information sources? *International Journal of Instructional Media*, 20.1 (Winter 2001): 83+.
- Lohse, G., & Rosen, D. Signaling Quality and Credibility in Directional Advertising: The Influence of Color and Graphics on Choice. *Journal of Advertising*, 30.2 (July 1, 2001): 73+.
- Urban, C. Examining Our Credibility: Perspectives of the Public and the Press. *American Society of Newspaper Editors* (1999). Retrieved October 16, 2002, from <http://www.asne.org/kiosk/reports/99reports/1999examiningourcredibility/index.htm>