What Makes Web Sites Credible?
A Report on a Large Quantitative Study

BJ Fogg, Jonathan Marshall, Othman Laraki, Alex Osipovich, Chris Varma, Nicholas Fang,

Jyoti Paul, Akshay Rangnekar, John Shon, Preeti Swani, Marissa Treinen

Persuasive Technology Lab
Stanford University
Cordura Hall, Stanford, CA 94309 USA
bjfogg@stanford.edu, www.webcredibility.org

ABSTRACT

The credibility of web sites is becoming an increasingly important area to understand. To expand knowledge in this domain, we conducted an online study that investigated how different elements of Web sites affect people's perception of credibility. Over 1400 people participated in this study, both from the U.S. and Europe, evaluating 51 different Web site elements. The data showed which elements boost and which elements hurt perceptions of Web credibility. Through analysis we found these elements fell into one of seven factors. In order of impact, the five types of elements that increased credibility perceptions were "real-world feel," "ease of use," "expertise," "trustworthiness," and "tailoring." The two types of elements that hurt credibility were "commercial implications" and "amateurism." This large-scale study lays the groundwork for further research into the elements that affect Web credibility. The results also suggest implications for designing credible Web sites.

Keywords

Credibility, World Wide Web, trustworthiness, expertise, captology, usability, online research, Web design.

INTRODUCTION

In the last few years we have seen a proliferation of Web sites with low-quality information, as well as sites that are outright misleading. Anecdotal evidence suggests that Web users are becoming more skeptical of the information they find online and may be wary of Web-based experiences. As a result, Web designers now face increasing pressure to enhance the credibility of their sites [6, 8]. But to this point, designing for Web credibility has been more art than science, leaving many Web designers to rely on intuition alone, without any support from quantitative research.

Given the importance of Web credibility, it is remarkable that so few quantitative studies exist. Although some private companies have created proprietary knowledge about designing for maximum Web credibility, only a handful of studies in this area are public. In other words, little research has been published on why people believe information on some Web sites but not others.

Some recent work has examined factors that affect trust in certain types of Web sites, such as e-commerce sites [2, 3]. Other studies have looked at credibility but in ways too limited to draw robust conclusions [4]. Drawing from small studies, researchers have suggested means of evaluating the quality of Web information [13], while Web site consultants have proposed ways to make Web sites more credible [e.g., 9]. However, no large study has yet been published that investigates and outlines the factors which influence the perceived credibility of Web sites. We felt this work was needed by both researchers and designers.

Therefore, over the course of a year our team studied the impact of a broad range of elements on perceived Web site credibility. We included Web site elements from various categories, such as aesthetic features, content variables, and technical factors. As researchers, our hope was to create a basic and broad understanding of what makes Web sites credible, as well as providing a foundation—both in method and content—for future quantitative investigations. As designers, our hope was to enhance the practice of designing for Web credibility, offering Web site creators concrete data for design decisions made on a daily basis.

WHAT IS "CREDIBILITY"?

To set the stage for the methods and results of our study, we first need to define "credibility," the focus of our research. Simply put, *credibility* can be defined as *believability*. Credible people are believable people; credible information is believable information. In fact, some languages use the same word for these two English terms.

Two additional points help clarify the credibility construct. First, credibility is a perceived quality [6]; it doesn't reside in an object, a person, or a piece of information. Therefore, in

discussing the credibility of a computer product, one is always discussing the *perception* of credibility [10, 11].

Next, scholars agree that credibility perceptions result from evaluating multiple dimensions simultaneously [1, 7, 10-12]. Although the literature varies on how many dimensions contribute to credibility evaluations, the vast majority of researchers identify two key components of credibility:

- trustworthiness
- expertise

What this means is that in evaluating credibility, a person makes an assessment of both trustworthiness and expertise to arrive at an overall credibility assessment.

Trustworthiness, a key element in the credibility calculus, is defined by the terms *well-intentioned*, *truthful*, *unbiased*, and so on. The trustworthiness dimension of credibility captures the perceived goodness or morality of the source.

Expertise, the other dimension of credibility, is defined by terms such as *knowledgeable*, *experienced*, *competent*, and so on. The expertise dimension of credibility captures the perceived knowledge and skill of the source.

Taken together, these ideas suggest that *highly credible Web sites will be perceived to have high levels of both trustworthiness and expertise* [5, 6].

With this background in mind, our team set out to research the specific elements of Web sites that would lead people to perceive a site as credible or not.

RESEARCH METHOD

In this research we gathered user data through a carefully constructed questionnaire. This self-report method is common in many types of research, but we acknowledge that this approach has inherent weaknesses. However, given the breadth of information we aimed to generate, using an online questionnaire was the most appropriate way to accomplish our goals. As described below, we used an iterative strategy to design the study.

Designing the Questionnaire – Items and Format

We designed the questionnaire elements in four stages.

In stage one, we created a bank of 300 items that related to an aspect of Web credibility (e.g., "the site has a typographical error," "the site requires you to log in"). We identified these items by interviewing Web users, reviewing writings of experts, and brainstorming as a team. In this stage, we also created 15 questions to assess demographic information.

In the second stage, our reseach team selected 90 of the 300 credibility items as possible elements for the study. Our selections were based on (1) keeping a wide variety of items, while at the same time (2) keeping items that raised important practical issues. We then ran a pilot study with 20 people, testing these 90 credibility items and the demographic questions. Much as we expected, people found the questionnaire too long and at times confusing.

In the third stage, we shortened and clarified the questionnaire and again pilot tested the study, this time with 30 people.

In our fourth stage, we selected what we determined to be the 51 most important credibility items. We also narrowed the demographic questions to eight total. Five questions gathered information on participants' age, nationality, education level, gender, and income. Three questions assessed user experience with the Web: years of experience using the Internet, hours spent online, and number of purchases over the Web.

We again pilot tested the new version of our study to verify that the questionnaire was clear and easy to complete.

Throughout these four stages of question development, we also explored different formats for gathering information online. After various prototypes and user tests, we settled on a format that used a 7-point Likert-type scale for each of the 51 questions on Web credibility (see Figure 1). This format allowed people to selected a response from "-3," representing that the item made a site "Much less believable," to "+3," indicating the item made a site "Much more believable." For the demographic questions, we used drop-down menus and open-ended response boxes.

During our iterative design process, we discovered a key advantage to administering questionnaires online: the order of the questions can be randomized. Researchers have long worried about how the order of questions can affect a person's response. Specifically, an early question can change how people react to subsequent questions, skewing the data. We saw how the dynamic capabilities of the Web can avoid this problem. So we programmed our study to list the 51 questions on credibility in a random order for each participant, avoiding potential systematic biases in the data.

Figure 1: A sample of our online questionnaire, showing the format.

Recruiting Study Participants

Ideally, a study like this would include a representative sample of all Web users throughout the world. We found no

viable way to recruit this type of sample. Instead, we recruited study participants in two ways: through the cooperation of charitable groups and through a news media organization in Finland.

In working with the charity groups (such as the American Brain Tumor Society), we donated $10 to the organization for each person who completed the study on their behalf, up to a maximum donation of $5000. This turned out to be an effective means of gathering data quickly.

Another source of participants was through a Finnish news media organization that placed a link from their Web site to our study site. We intentionally involved participants from two countries so we could perform cross-cultural analyses of the data.

Running the Study

We ran the main study during December of 1999. Participants began the study by going to the Web site, www.webresearch.org, which they could do on their own time from their home or workplace. In completing the online study, each participant answered the questions by drawing on his or her cumulative experience using the Web.

The data collection went faster than expected, requiring less than one week, thanks to the power and reach of the Web. As we exceeded 1400 participants, we concluded the data gathering portion of the research and began the analysis.

STATISTICAL ANALYSIS

During the course of running the study, 1441 people completed the questionnaire. Of those, we discarded 19 questionnaires because they did not contain complete demographic data or contained less than 47 answers to the 51 content questions. We discarded another 12 questionnaires because the respondents did not come from either Finland or the U.S.

Overall Demographics

Who participated in this study? Table 1 shows the demographic information for the 1410 participants. The average age for our respondents was just over 33 years. The sample was 44% female, and 42% of the total sample was from the U.S. The other notable feature is that our sample was fairly experienced with the Web: more than 3 years on the Internet, almost 14 hours spent online each week, and most people having made at least one online purchase.

Analysis Plan

The data analysis had three parts: calculating the overall mean responses, creating the composite scales, and comparing responses based on demographics. Determining the overall means was a straightforward mathematical process, while creating composite scales of credibility and comparing demographic subgroups required more effort.

Table 1: Summary of participants' demographic information

Demographic Item	Overall results
Age (median)	33 years
Gender	44% female, 56% male
Country	42% U.S., 58% Finland
Education level	"some university"
Income (mean)	$31,459
Years on the Internet (mean)	3.7 years
Average number of purchases online (mean)	"1 to 5 purchases"
Average number of hours spent online a week (mean)	13.9 hours/week

Creating Composite Scales

To determine the underlying dimensions in our study, a correlation matrix was computed and a factor analysis performed. Using a scree-test, we determined that there were four factors (Eigen values > 1.73) accounting for 15.1% of the total variance. We used Varimax rotation to determine which questions loaded on each of the factors. Closer investigation indicated that three factors could be sensibly divided into two scales each in order to make further sense of the data. We calculated the internal reliability of each scale using Cronbach's alpha. We named the resulting seven scales as follows:

- Real-World Feel
- Ease of Use
- Expertise
- Trustworthiness
- Tailoring
- Commercial Implications
- Amateurism

Comparing Demographic Subgroups

In our analysis, we compared subgroups of our entire sample. Specifically, we compared subgroups defined by—

- Age
- Gender
- Country of origin
- Education level
- Income level
- Experience with the Web

A Bonferroni correction was applied and statistical significance was defined as $p < 0.00016$.

Categorical comparisons were made between genders and between countries of origin.

To compare age, education, experience, and income, we divided the sample into thirds. The middle third of the sample was eliminated and the high and low thirds were compared against each other using t-tests. For example, the youngest third of the sample was compared against the oldest third, while the results from the middle third were unused. We chose this method instead of using a general linear model as it facilitates calculating Cohen's effect sizes.

Our study assessed "experience on the Web" in three ways. To get one overall measure, we combined the percentile ranking of each subject across the three questions and analyzed the data as explained in the previous paragraph.

RESULTS

Composite Scales – Overall Results

This section contains Tables 2 through 8, which list the questionnaire items that comprise each of the seven scales. These tables also give the overall means for each item. For example, Table 2 shows that the "Real-World Feel" scale is made up of five items from the questionnaire. One item in this scale is "The site provides a quick response to your customer services questions." Taking into account all 1410 respondents, we found the overall mean for this particular item was 2.02 (again, the possible range was –3 to +3).

(Note: The items listed in each table are worded exactly as they appeared in the study. Also, the value for Cronbach's alpha listed in each table heading is a measure of how well the items represent a unified construct.)

Table 2: Real-World Feel Scale (Cronbach's alpha = 0.66)

Items in the **REAL-WORLD FEEL** scale	Mean
The site provides a quick response to your customer service questions.	2.02
The site lists the organization's physical address.	1.86
The site gives a contact phone number.	1.71
The site gives a contact email address.	1.53
The site shows photos of the organization's members.	0.69

As Table 2 shows, all five items were reported to increase the credibility of a Web site. The strong positive responses on four of the five items made this scale the most dominant in increasing credibility perceptions, as shown in Figure 2.

The next most influential scale was the "Ease of Use" scale, shown in Table 3. This scale was made up of five items, with three describing the ease of use construct and two describing the opposite end of the spectrum.

Table 3: Ease of Use Scale (Cronbach's alpha = 0.67)

Items in the **EASE OF USE** scale	Mean
The site lets you search past content (i.e. archives).	1.57
The site looks professionally designed.	1.55
The site is arranged in a way that makes sense to you.	1.48
The site takes a long time to download.	-0.94
The site is difficult to navigate.	-1.30

One notable item in Table 3 is the item on navigation. The negative value of –1.30 shows how much navigation problems were reported to hurt Web site credibility.

Table 4 shows the seven items that comprise the "Expertise" scale. Each of these items, such as listing author credentials and providing citations and references, was viewed as a positive contributor to a Web site's credibility.

Table 4: Expertise Scale (Cronbach's alpha = 0.63)

Items in the **EXPERTISE** scale	Mean
The site is by a news organization that is well respected outside of the Internet.	1.91
The site lists authors' credentials for each article.	1.49
The site has articles that list citations and references.	1.49
The site has few news stories but gives detailed information for each.	1.10
The site says it is the official site for a specific topic	0.85
The site has ratings or reviews of its content.	0.79
The site displays an award it has won.	0.45

The "Trustworthiness" scale is made up of eight items, as shown in Table 5. The scores for all eight items are positive, ranging from 0.58 to 1.29.

Table 5: Trustworthiness Scale (Cronbach's alpha = 0.57)

Items in the **TRUSTWORTHINESS** scale	Mean
The site is linked to by a site you think is believable.	1.29
The site states its policy on content.	1.26
The site links to outside materials and sources.	1.25
The site provides links to its competitors sites.	1.11
The site was recommended to you by a friend.	1.07
The site represents a nonprofit organization.	0.93
The site lists well-known corporate customers.	0.62
The URL for the site ends with ".org"	0.58

Table 6 lists the four items that make up the "Tailoring" scale. This scale has the lowest Cronbach's alpha in the study, indicating it is likely the least unified construct of the seven. In addition, this scale has the smallest positive impact on perceptions of credibility, as shown in Figure 2.

Table 6: Tailoring Scale (Cronbach's alpha = 0.44)

Items in the **TAILORING** scale	Mean
The site sends emails confirming transactions you make.	1.41
The site selects news stories according to your preferences.	0.57
The site recognizes that you have been there before.	0.37
The site requires you to register or log in.	0.07

One notable item from Table 6 is the nearly negligible impact of requiring users to "register or log in" (mean = 0.07). This low value shows this element was reported to have almost no effect on the credibility perception of a Web site.

With Table 7, the impact of the scales on credibility perceptions changes from being overall positive to overall negative. Table 7 lists the eight items that make up the "Commercial Implications" scale, with most of the items reducing Web credibility. The overall effect of this scale, as shown later in Figure 2, is to reduce credibility perceptions somewhat.

Table 7: Commercial Implications (Cronbach's alpha = 0.65)

Items in the **COMMERCIAL IMPLICATIONS** scale	Mean
The site is advertised on the radio or on billboards.	0.57
The site has ads that match the topic you are reading about.	0.21
The site is designed for e-commerce transactions.	0.17
The site has a commercial purpose (as opposed to academic purpose).	-0.63
The site requires a paid subscription to gain access.	-0.71
The site has one or more ads on each page.	-0.77
The site automatically pops up new windows with ads.	-1.56
The site makes it hard to distinguish ads from content.	-2.08

One noteworthy item in Table 7 is the final item, where a "site makes it hard to distinguish ads from content." This item received the lowest score in the entire study, with a mean of –2.08.

Table 8 shows the ten items that comprise the "Amateurism" scale. This scale has two items that represent the opposite of amateurism (site is "updated" and site is "in more than one language"), with the remaining eight items being viewed as manifestations of amateurism.

Table 8: The Amateurism Scale (Cronbach's alpha = 0.64)

Items in the **AMATEURISM** scale	Mean
The site has been updated since your last visit.	1.55
The site offers information in more than one language.	1.04
The site is small (e.g. less than 5 pages).	-0.28
The site is hosted by a third party (e.g. AOL, Geocities).	-0.44
The site's domain name does not match the company's name.	-1.06
The site has a typographical error.	-1.28
The site is sometimes unexpectedly unavailable.	-1.28
The site has a link that doesn't work.	-1.45
The site links to a site you think is not credible.	-1.53
The site is rarely updated with new content.	-1.67

Of the 51 items we tested in this study, four items did not fit well into any of the scales listed above. For the sake of completeness, we list those items in Table 9.

Table 9: Individual items not part of a scale

Items not used in any scale	Mean
The site represents an organization you respect.	1.93
The site provides financial news at no charge.	0.53
The site contains information that doesn't match what you think.	-0.77
The site has lots of news stories without giving detailed information.	-0.89

To show the relative impact of each credibility scale, we created Figure 2 with a standardized vertical axis. (The

vertical axis was designed to have a theoretical maximum score of 100. For a credibililty scale—such as the "expertise scale"—to reach this maximum value, each individual item within the scale would need to receive a perfect mean score of "+3.0" in the study data.)

Figure 2 shows how much the first five scales increase perceptions of Web credibility. As a whole, the items in the "Real-World Feel" scale have the largest impact on increasing credibility, followed by items in the "Ease of Use" scale. The "Expertise," "Trustworthiness," and "Tailoring" scales also show positive effects.

Figure 2 also shows that how much the combined items in two scales, "Commercial Implications" and "Amateurism," decrease perceptions of Web credibility.

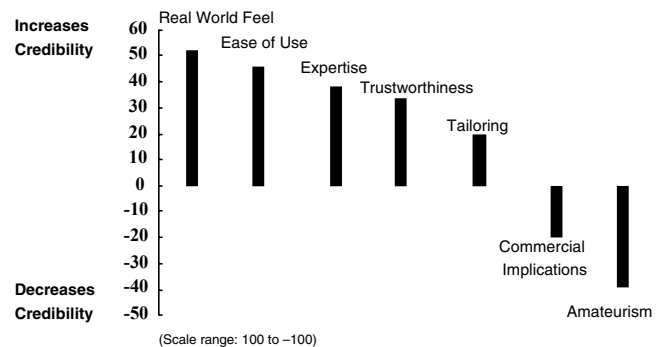## How the Factors Impact Web Credibility

Figure 2: The seven scales and their effects on perceived credibility.

Demographic Comparisons
Because we collected demographic data for each respondent, we were able to analyze what differences exist between subgroups of various characteristics.

Comparisons based on age
How does age change perceptions of Web credibility? In our efforts to answer this question, we found that the youngest third of our sample (age < 27.9 years) tended to be more critical of amateuism on a site, compared to the oldest third of our sample (age > 37.3 years). This means the younger respondents were more harsh on sites that had glitches like a single typo or a broken link.

The older respondents reacted more favorably to Web sites conveying markers of expertise and trustworthiness. The older respondents also reported slightly higher credibility evaluation for sites that used some type of tailoring. These differences shown in Table 10, are statistically significant ($p<0.00016$), though they may not be of great practical significance.

Table 10: Differences when comparing age subgroups
(*p*<0.00016)

	Mean responses from people **under 27.9 years**	Mean responses from people **over 37.3 years**
Amateurism Scale	-1.19	-1.06
Expertise Scale	1.13	1.23
Trustworthiness Scale	0.98	1.11
Tailoring Scale	0.56	0.74

Comparisons based on gender

How do men and women differ in perceptions of Web site credibility? The results from our data analysis are provocative, though not entirely conclusive. In comparing all responses from men with all responses from women, we found that men assigned less credibility overall. In other words, men answered the questions more negatively than women (men's mean response = 0.31; women's mean response = 0.42; p < 0.0001). On one hand, this could imply that when it comes to credibility, men are slightly more critical of Web sites than women. On the other hand, the lower response from men could indicate that men generally respond more critically when answering questionnaires.

As we examined how men and women compared on each of the seven scales, we found that men assigned lower credibility on five scales. Table 11 shows these results.

Table 11: Differences when comparing gender subgroups
(*p*<0.00016)

	Mean responses from **women**	Mean responses from **men**
Real-World Scale	1.65	1.5
Expertise Scale	1.24	1.08
Trustworthiness Scale	1.11	0.94
Tailoring Scale	0.73	0.51
Commercial Implications Scale	-0.55	-0.64

Comparisons based on country of origin

Because our respondents came from both Finland and the U.S., we were able to make cross-cultural comparisons in regard to perceptions of Web credibility. We found a number of differences, as outlined in Table 12. Compared to the U.S. participants, Finns reported lower credibility perceptions for Web sites that conveyed commercial implications and for those that showed marks of amateurism. Compared to Finns, respondents from the U.S. assigned more credibility to sites that used some type of tailoring technology and sites that conveyed expertise and trustworthiness, as described in our questionnaire. No other differences were significant.

Table 12: Differences when comparing country subgroups
(*p*<0.00016)

	Finland means	**U.S. means**
Amateurism scale	-1.21	-1.08
Expertise scale	1.03	1.32
Trustworthiness scale	0.87	1.21
Tailoring scale	0.42	0.85
Commercial Implications Scale	-0.68	-0.49

Comparisons based on education level

How does education level affect people's perception of Web credibility? To answer this question, we compared the most highly educated third (those who completed graduate school) of our sample with from the least educated third (those with no college education). This analysis showed the differences on only the trustworthiness scale. As shown in Table 13, compared to those with no college experience, respondents who completed graduate school assigned more credibility to Web sites that conveyed markers of trustworthiness. No other measures were significant.

Table 13: Differences when comparing education subgroups
(*p*<0.00016)

	Mean response from people with **no college experience**	Mean responses from people who **completed graduate school**
Trustworthiness Scale	0.857	1.03

Comparisons based on income level

In investigating differential responses based on income level, we found that the wealthiest third of our sample (income >$60,000 USD) differed from the least wealthy (income <$20,000 USD) on only one of the scales: tailoring. Specifically, compared to those with less income, the wealthier subjects assigned more credibility to sites that used tailoring. The results are shown in Table 14. Other comparisons were not significant.

Table 14: Differences when comparing education subgroups
(*p*<0.00016)

	Mean response from those with **incomes <$20,000**	Mean response from those with **incomes >$60,000**
Tailoring Scale	0.55	0.72

Comparisons based on experience with the Web

Finally, we investigated whether experience with the Web caused people to evaluate Web credibility differently. We hypothesized it would. But the data tell a different story. The most experienced third of our sample did not differ from the least experienced third on any of the 7 scales of credibility.

Overall assessment of demographic comparisons

As shown above, the data showed statistically significant differences when comparing how different subgroups assessed the credibility of Web sites. However, these differences were typically small; they are likely not to be of great practical significance. Additional research, with more focused and additional measures, can help provide a firm conclusion.

At this point, one can reasonably hypothesize that most people—regardless of age, gender, or other demographic factors—assess Web site credibility in similar ways. Although real differences do exist, it's more striking to see how many things were *not* different, suggesting that the various demographic groups shared similar approaches to evaluating Web credibility.

DESIGN IMPLICATIONS

We have explained the method and rational of this research so readers can understand how these findings apply—or do not apply—to their particular situations in designing Web experiences. Below we outline some design implications from this research. We present them as proposed guidelines for those seeking to create highly credible Web sites. Each guideline comes from one of the seven scales that emerged from the study data. (One could also extract more design implications from the items listed in Tables 2-9 and 11-14.)

Guideline #1: Design Web sites to convey the "real world" aspect of the organization.

According to our study results, the most effective way to enhance the credibility of a Web site is to include elements that highlight the brick-and-mortar nature of the organization it represents. We examined Web elements such as a listing a physical address and showing employee photographs. Many other possibilities exist that were not included in this study. To create a site with maximum credibility, designers should highlight features that communicate the legitimacy and accessibility of the organization.

Guideline #2: Make Web sites easy to use.

In the HCI community we have long emphasized ease of use, so this guideline is no surprise. However, this study adds another important reason for making Web sites usable: it will enhance the site's credibility. In this study, people awarded a Web site credibility points for being usable (e.g., "The site is arranged in a way that makes sense to you"), and they deducted credibility points for ease-of-use problems (e.g., "the site is difficult to navigate"). While this information should not change how we, as HCI professionals, design user experiences for the Web, it does add a compelling new reason for investing time and money in usable design—it makes a site more credible. Going beyond the data, one could reasonably conclude that a simple, usable Web site would be perceived as more credible than a site that has extravagant features but is lacking in usability.

Guideline #3: Include markers of expertise.

Expertise is a key component in credibility, and the data in this study support the idea that Web sites which convey expertise can gain credibility in users' eyes. Important "expertise" elements in this study included listing an author's credentials and including citations and references. It's likely that many other elements also exist. In our opinion, many Web sites today miss opportunities to legitimately convey expertise to their users.

Guideline #4: Include markers of trustworthiness.

Trustworthiness is another key component in credibility. As with expertise, this study suggests that Web site elements that convey trustworthiness will lead to increased perceptions of credibility. In this research we tested how people assessed specific "trustworthiness" elements: linking to outside materials and sources, stating a policy on content, and so on. Of course, other markers of trustworthiness exist. We propose that Web site designers who concentrate on conveying the honest, unbiased nature of their Web site will end up with a more credible—and therefore more effective—Web site. This suggestion may be opposed to other forces in an organization that seek a more promotional approach to Web design.

Guideline #5: Tailor the user experience.

Although not as vital as the previous suggestions, tailoring does make a difference. Our study shows that tailoring the user experience on a Web site leads to increased perceptions of Web credibility. For example, people think a site is more credible when it acknowledges that the individual has visited it before. To be sure, tailoring and personalization can take place in many ways. Tailoring extends even to the type of ads shown on the page: ads that match what the user is seeking seem to increase the perception of Web site credibility.

Guideline #6. Avoid overly commercial elements on a Web site.

Although most Web sites, especially large Web sites, exist for commercial purposes, our study suggests that users penalize sites that have an aggressively commercial flavor. Web pages that mix ads with content to the point of confusing readers will be perceived as not credible. In this study, mixing ads and content received the most negative response of all. But it is important to note that ads don't always reduce credibility. In this study and elsewhere [14] quantitative research shows that banner ads done well can enhance the perceived credibility of a site. It seems reasonable that, as with other elements of people's lives, we accept commercialization to an extent but become wary when it is overdone.

Guideline #7. Avoid the pitfalls of amateurism

Most Web designers seek a professional outcome in their work. This study suggests organizations that care about credibility should be ever vigilant—and perhaps obsessive—to avoid small glitches in their Web sites. These "small" glitches seem to have a large impact on Web credibility perceptions. Even one typographical error or a single broken

link is damaging. While designers may face pressures to create dazzling technical features on Web sites, failing to correct small errors undermines that work.

CONTRIBUTIONS TO RESEARCH

Although this study has its shortcomings (most notably the self-reported data and some modest Cronbach's alpha values), the overall contributions from this research are significant, both in terms of research methods and increased understanding.

Contributions to Research Methods

This study has contributed methodologically. While using questionnaires is not a new method, we believe our research efforts generated methods for and insights into HCI research online, including response scale design, randomization of items, recruitment of study participants through charity group collaborations, and rapid data collection. Future researchers can employ or build upon these things, saving time and effort. Future researchers can also benefit from the concise set of demographic questions we developed through our research design process.

Increased Understanding

Finally, to our knowledge this is the first public study on web credibility of this magnitude. The scope and focus makes it unique. Even though some findings in this research were predictable, others findings were new and provocative. In either case, we believe this study serves a valuable function for HCI researchers by establishing baseline credibility values for the 51 web site elements it investigated. The study also outlines seven scales that affect credibility perceptions. Furthermore, the findings come from a large, cross-cultural sample and are supported by statistical analyses.

Because this study has raised issues for further inquiry, we now hope that sharing this knowledge and method will stimulate additional research in this domain, which is an area that will become increasingly important as the Web permeates the lives of people around the world.

ACKNOWLEDGEMENTS

REFERENCES

1. Buller, D.B. & Burgoon, J.K. (1996). Interpersonal Deception Theory. *Communication Theory*, 6(3) 203-242.

2. Cheskin Research, *Trust in the Wired Americas*, 2000. See www.cheskin.com/think/studies/trust2.html.

3. Cheskin Research, S.A.S., *eCommerce Trust Study.* 1999. See www.studioarchetype.com/headlines/etrust_frameset.html.

4. Critchfield, R. *Credibility and Web Site Design*, 1998. See www.warner.edu/critchfield/hci/critchfield.html.

5. Fogg, B.J., Lee, E., & Marshall, J. (forthcoming). Interactive Technology and Persuasion. In J. P. Dillard and M. Pfau (Eds.), *The Handbook of Persuasion: Theory and Practice*. Thousand Oaks, CA: Sage.

6. Fogg, B.J. & Tseng, H. The Elements of Computer Credibility. *Proceedings of the CHI99 Conference on Human Factors and Computing Systems*, 1999 (pp. 80-87), ACM Press.

7. Gatignon, H., & Robertson, T.S. (1991). Innovative Decision Processes. In T.S. Robertson & H.H. Kassarjian (Eds.), *Handbook of Consumer Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

8. Morkes, J. and Nielsen, J., *Concise, SCANNABLE and Objective: How to Write for the Web,* 1997. See www.useit.com/papers/webwriting/writing.html

9. Nielsen, J., *Trust or Bust: Communicating Trustworthiness in Web Design*, 1999. See www.useit.com/alertbox/990307.html.

10. Petty, R. and Cacioppo, J. (1981). *Attitudes and Persuasion: Classic and Contemporary Approaches*. Dubuque, IA: Brown.

11. Self, C.S. (1996). Credibility. In M. Salwen & D. Stacks (Eds.)*, An Integrated Approach to Communication Theory and Research*. Mahway, NJ: Erlbaum.

12. Stiff, J. (1994). *Persuasive Communication*. New York: Guilford.

13. Wilkinson, G.L., Bennett, L.T., & Oliver, K.M. 1997. Evaluating the Quality of Internet Information Sources. See itech1.coe.uga.edu/faculty/gwilkinson/criteria.html.

14. Kim, N., & Fogg, B.J. (1999). World Wide Web Credibility: What effects do advertisements and typos have on the perceived credibility of Web page information? (Unpublished thesis, Stanford University).