Judging the Veracity of Web Sites

Martha E. Crosby, Marie K. Iding and E. Barbara Klemm
University of Hawaii
Honolulu, HI 96822

Brent Auernheimer
California State University
Fresno, CA 93740

Abstract

People vary considerably in their competency levels depending on the web sites’ topic. Does this knowledge influence their ability to make judgements? In order to see if we could discover a relationship between judgments, confidence, and knowledge of the topic, we conducted a study using participants who already possessed considerable evaluation skills and were sophisticated users of the World Wide Web. We picked a topic that that some of our participants would know very well and some of the participants would know very little about. Participants were asked whether the web sites were considered objective and accurate, mistaken, or purposely misleading and give reasons for their decision. We found that the confidence ratings varied considerably with expertise but the judgments themselves did not. This gives educators hope that evaluation skills may be taught that are general enough that students can apply them to a variety of web sites.

An Experiment

How do you know when a web site is being truthful? The answer to this question is so important that universities explicitly train students in “information competence” [1]. In a previous study [2], we described the research of Kruger and Dunning [3] who investigated competency areas that included humor, grammar, and logic. These researchers explored college students' tendencies to erroneously overestimate their levels of competency, particularly with respect to areas where expertise is lacking. Ironically, the more expert college students tended to rate their own competence less highly than their less-capable peers. Kruger and Dunning argue that this could be due to the expert students mistakenly overestimating the competencies of their peers. In our previous study, we used this work as a foundation to raise questions about college students' possible overestimation of their competence levels regarding evaluating web sites. In that particular study, we found that educational psychology and computer science students self-ratings of competencies regarding content varied, with educational psychology students rating themselves higher than computer science students. With respect to rating information on the Web, the pattern was reversed. In this case, the computer science students rated themselves higher than educational psychology students. Students from both content areas had difficulty delineating possible vested interests of web site authors, and in articulating reasons for determining that a web site's information is valid/credible or inaccurate. Clearly, students’ vary in the accuracy of their competency levels. However, how does instruction affect accuracy judgments? And how much is necessarily content-specific?

Knowledge about a topic is a crucial component for accurate appraisal of a web site’s veracity. However, the reviewed research suggests that one’s confidence in their judgment does not necessarily mean that the judgment is accurate. The important point is: if content knowledge is the primary predictor of accurate judgement of web sites veracity, it may be difficult to teach general skills that can be applied to the judgment of a variety of web sites. In order to see if we could discover a relationship between the actual judgments, confidence, and knowledge of the topic, we conducted this study using participants who already possessed considerable evaluation skills and were sophisticated users of the web. The topic was chosen so that some of our participants would have knowledge of the content and some them would not. Participants consisted of 25 students enrolled in graduate courses in either computer science (CS) or communications and information sciences (CIS) and 10 faculty members or researchers in computer science and education related disciplines. All of the participants had considerable expertise in technology or computer science research.

For the previous study that served as a basis for the present research [2], students were asked to work in small groups to find web sites related to topics covered in class. The topics were relevant to the course content. Each group of students selected a web site that they determined to provide the most accurate and objective presentation of
the topic and a web site that exemplified misrepresentations about the topic. They explained why they selected the sites, what authors' vested interests might be, and addressed a question about how they know when web site content is wrong.

For the present study, CS and CIS graduate students and faculty members were provided with three web sites on the topic of "cleanroom" procedures for software development [4]. These web sites were selected from the previous study students' selections of accurate and inaccurate web sites (i.e., web sites containing misrepresentations). Interestingly, in the previous study, some of the same web sites were selected by different groups of students as being accurate and by others as being inaccurate. In the present study, participants rated their knowledge of the assigned topic. Also, the CS students had taken a class in this topic. For each of the three web sites, participants rated whether the web site was considered objective and accurate, mistaken, or purposely misleading and they were asked to explain their answers. In addition, on a scale of 1 (no confidence) to 5 (complete confidence) the participants rated their own confidence about their categorization, and addressed a question about what the authors' vested interests might be. In the final section of the questionnaire, respondents provided essay responses to the following question: "How do you know when a WWW site's content is wrong?"

Finally, on a scale of 1 (no confidence) to 5 (complete confidence), the participants rated their confidence in their decisions. The ratings were made in the following areas: (1) confidence in detecting misrepresentations on web sites in general, (2) competence in evaluating the validity of information on the Web in general and (3) competence in evaluating the validity of information about "cleanroom" software engineering.

The participants were divided into two groups based on their knowledge of "cleanroom" software engineering. Participants with a confidence score of 4 or 5 were called "topic experts," while participants with a confidence rating of 3 or lower were called "topic novices." Participants who did not provide a rating were eliminated from the analysis. An analysis of variance (ANOVA) was carried out using with the topic factor (expert or novice) as the independent variable. This analysis had no significant main effects for any of the actual "cleanroom" web site ratings. However, the analysis yielded significant effects on all of the confidence measures for each of the three web sites. For web site 1, $F(1,29) = 14.4, MSe = 5.8, p = .0007$; for web site 2 $F(1,29) = 4.3, MSe = 5.8, p < .05$ and for web site 3, $F(1,29) = 9.6, MSe = 7.35, p < .005$. There was also a significant effect for confidence in evaluating the validity of information on the Web in general, $F(1,29) = 6.7, MSe = 4.3, p = .01$. It is interesting to compare the remarks made by the topic experts to those of the topic novices for the same "cleanroom" web sites. For the first site an expert with a confidence rating of 5 judged the web site to be objective and accurate. The explanation for the decision was that the web site was "well organized, says the right things - the "cleanroom" party line. A novice with a confidence rating of 3 states "It is difficult to tell when a web site is misleading or mistaken, but this one appears to be objective and accurate. The web site has a professional look to it, their online library cites sources, and time was spent do the users can view their material..." For the second site, the expert claims that the web site is mistaken. The explanation states "misspellings, first person experience and doesn't really sound like he understands the terminology." A novice with a confidence rating of 2 also judges the web site as mistaken. The explanation states "Shipman is an application specialist...He appears to be expressing his interpretation of "cleanroom" technology, there is no assurance of accuracy." For the third web site, the expert states "He doesn't really talk about the "cleanroom," he just gives pointers." The novice, who has a confidence rating of 3, states "The main purpose of the web site is to sell the company's consulting services, and Dr. Pressman's books and videos. Because of this I can not say that their presentation of the "cleanroom" can be entirely objective." As these comments suggest, the confidence ratings varied considerably with expertise, but the judgments themselves and the rationale for making these judgments did not. This is an encouraging result. It suggests that general evaluation skills may be taught and that students can apply them to judging the veracity of web sites.

Acknowledgements
This research was supported in part by ONR grant no. N00014970578 awarded to M. E. Crosby and DARPA grant NBCH1020004 awarded to M. E. Crosby and M. K. Iding.

References

