Evidence Based Library and Information Practice

Consumer Health Information Websites with High Visual Design Ratings Likely to Be also Highly Rated for Perceived Credibility

A Review of:
Robins, D., Holmes, J., & Stansbury, M. (2010). Consumer health information on the web: The relationship of visual design and perceptions of credibility. *Journal of the American Society for Information Science and Technology, 61*(1), 13-19.

Reviewed by:
Kate Kelly
Chief Librarian
Royal College of Surgeons in Ireland
Lower Mercer Street, Dublin 2. Ireland
Email: katekelly@rcsi.ie

Objective – To answer two research questions: 1) What is the relationship between the visual design of a consumer health information web site and perceptions of the credibility of information found on it? 2) Is there a relationship between brand recognition, visual design preference, and credibility judgments?

Design – Qualitative (correlation of rating of response to stimulus) and quantitative (credibility coding of participant comments)

Setting – Not stated; assumed to be academic institutions in the United States.

Subjects – Thirty-four participants over the age of 35 (34 for statistical power and age over 35 on the hypothesis that this age group is

most likely to seek health information on the Internet).

Methods – Screen shots of 31 consumer health information sites chosen from the results of a Google search using the term "consumer health information" were converted to slide format and shown to participants. The 31 sites included 12 of the top ranked consumer health information sites derived from three sources: the Consumer and Patient Health Information Section (CAPHIS) of the Medical Library Association (MLA), the MLA itself, and Consumer Reports. Participants were read and shown a script explaining the process prior to being asked to view and rate the 31 sites. Participants were first shown a blank slide with a crosshair to focus attention. Then a

stimulus slide was shown for 2.8 seconds. A blank black screen was shown while they determined their rating. Participants were first asked to rate the visual design and aesthetics of the 31 web sites using a rating scale of -4 to -1 for negative judgments and +1 to +4 for positive judgments. Then they were asked to remember why they had made positive or negative ratings and why some web sites were preferred to others. The process was repeated with the slides re-ordered, and participants were asked to rate the credibility of the sites using the same rating scales. Upon completion, participants were asked to recall their reasons for positive or negative credibility ratings. All ratings were converted to positive numbers and a scale of 1-8 was used to present results. A variety of statistical analyses were carried out on the data, including correlation, means ratings, and rankings. In addition, all solicited comments on credibility were coded using Fogg's four types of credibility (surface, earned, presumed, and reputed) in order to try to understand why participants rated the credibility of sites as they did.

Main Results – For the first question, concerning the relationship between visual design preferences and perceived credibility, the results are complicated. A statistically significant correlation was reported between visual design preference and perceived credibility in 8 of the 31 sites (26%). In these instances where visual design is rated highly, so is credibility. When visual design ratings were ranked highest to lowest, credibility ratings followed the same pattern. Similarly, when credibility ratings were ranked highest to lowest, visual, design ratings followed. A t-test confirmed that sites perceived to have higher credibility were also perceived to have better visual design. Furthermore, when design and credibility ratings were compared to site traffic rankings, as measured by Alexa (http://www.alexa.com), the trend was for both visual design and credibility ratings to decline as the site traffic ranking declined. This finding was also confirmed by a t-test. While there is not an exact relationship, the tendency is for sites with higher visual design

ratings to also receive higher ratings for perceived credibility.

On the second question, concerning the relationship between brand recognition and visual design and perceived credibility judgments, the results suggest a possible influence of brand name. This relationship is not clear, and as visual designs were always presented and rated first, there is possibly a co-founder. The analysis of participant comments found that participants performed credibility judgments in a very short time using a variety of criteria, including visual design, source of the site, reputation of the site, and prior use. There were negative reactions to the use of advertisements, drug and insurance company sponsorship, and dot com sites, as well as some suspicion that non-US consumer health information sites were less trustworthy.

Conclusions – Visual design judgments bore a statistically significant similarity to credibility ratings. Sites with recognizable brands were highly rated for both credibility and visual design, but this relationship was not statistically significant. The relationship is complicated and more research is needed on what visual design cues are important to credibility judgments.

Commentary

As an exploratory study, this is a very interesting and highly readable piece of research. It is, however, undermined by a lack of detail about the research participants and questions about the statistics.

Irrespective of whether the sample size of 34 gives statistical power, the information that the participants were chosen by age alone tells us nothing about their gender (studies indicates that females are more likely to seek health information on the Internet); their experience of searching for health information; comfort level with using the Internet; exactly how old or how IT/information literate they were; their ethnic background, or if English

was everyone's first language. Was it a convenience sample, or a random sample? The hypothesis that people over the age of 35 are more likely to seek health information on the Internet is debatable. For instance, the Pew Report 2009 indicated that 72% of 18-29 year olds sought health information on the Internet (Fox & Jones, 2009).

With regard to the statistics, the authors state that the statistical power is 0.80, but provide no details of how this was calculated. While 0.80 is a desired level of power, the minimum sample size for relationship testing with a single predictor using correlation or regression analysis is 50 cases (Harris, 1985; Green, 1991). According to Cohen and Cohen (1975), with an effect size of 0.30, 124 participants are needed to maintain 80% power. The sample size in this study does not meet the minimum requirement for statistical power of 50 cases, let alone the sample size required for statistical power of 0.80.

Readers should also note that only 8 out of 31 (26%) websites tested showed a significant correlation between visual design and credibility. That means that for 74% of the websites tested, there was no correlation between visual design and credibility. This would appear to invalidate the authors' conclusion that design is an important attribute of website credibility. Without details, the validity and generalizability of the findings has to be open to question.

While the research methodology is described in reasonable detail, it is unclear whether participants were shown slides in a group or individually; and similarly, whether comments were solicited individually or as part of a group setting. These details would be useful for replicating the study.

The literature review, however, is excellent and places the research in context in terms of an overall credibility framework, credibility on the web, and credibility of health information on the web. The exploration of credibility ratings is instructive, as is previous research by Robins and Holmes describing how

impressions of web sites are formed within 2.4-3.2 seconds. This is clearly a precursor to the research under review (Robins & Holmes, 2008).

Health sciences librarians are likely to find the list and analysis of the 31 web sites very interesting. It includes expected sites like MedlinePlus, but also drug companies such as Novartis and authoritative non-US sites such as the Royal Australian College of General Practitioners.

The authors discuss fully the limitations of this study and identify several areas for further research. While it is too early to say with any certainty, it is clear that this line of research could yield results which may influence website design, health information literacy, and the teaching of critical appraisal and evaluation of health internet sites. For those with responsibilities in these areas, this study should be essential reading, even with the limitations noted. It may be interesting for non-US readers to consider replicating the study to test the generalizability of these results within their own national and cultural contexts.

References

Cohen, J., & Cohen, P. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Fox, S., & Jones, S. (2009). Social life of health information. *Pew Research Centre*. Retrieved 12 May 2010 from http://www.pewinternet.org/Reports/2009/8-The-Social-Life-of-Health-Information.aspx

Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research, 26*(3), 499-510.

Harris, R. J. (1985). *A primer of multivariate statistics* (2nd ed.). New York: Academic Press.

Robins, D., & Holmes, J. (2008). Aesthetics and credibility in web site design. *Information Processing and Management, 44*(1), 386-399. Retrieved 13 May, 2010 from www.sciencedirect.com doi:10.1016/j.ipm.2007.02.003