

Chapter XXXV

Web Credibility Measurement

Rodney A. Reynolds

Azusa Pacific University, USA

(on leave from Pepperdine University, USA)

BACKGROUND

Several researchers (e.g., Carter & Greenberg, 1965; Flanagin, & Metzger, 2000; Fogg, 2002; Johnson & Kaye, 2004; Newhagen & Nass, 1989) discuss or mention the concept of media or Web credibility. The classic concept of credibility (typically attributed to Aristotle's *Rhetoric*) identifies credibility as a multidimensional perception on the part of the receiver that the source of a message has a moral character, practical wisdom, and a concern for the common good. Warnick (2004) points out that the "authorless" nature of the online environment complicates the use of traditional analyses of credibility.

The most common set of Web credibility scales cited in the research are the Flanagin and Metzger (2000) items. The five Flanagin and Metzger scale items address the believability, accuracy, trustworthiness, bias, and completeness of the information on the Web site. Other researchers have added other items such as fairness or depth of information. Flanagin and Metzger used a 7-point response format with anchors for each term (e.g., "Not At All Believable" to "Extremely Be-

lievable"). Other researchers have used a 5-point response format.

Fogg (2002) indicates a slightly different approach adopted at the Stanford University Web credibility project (Fogg cites Danielson, in press, as the most current overview on Web credibility). In the Stanford studies, the researchers asked participants to examine many different characteristics from a number of Web sites, and rate each characteristic on the degree that the characteristic contributed to the "believability" of the Web site. The Stanford approach basically replicated the Flanagin and Metzger (2000) Web credibility features, but extended the list of features to include items related to operation or navigation on the site (e.g., organization of site, access or contact, ease of use, updating, excesses in promotional content, and errors of any form).

RELIABILITY

Flanagin and Metzger (2000) report an average Web credibility reliability across four different types of information (news, entertainment, commercial, and reference) of $\alpha = .91$ (p. 523).

VALIDITY

There is relatively little information on the validity of the measurement of Web credibility. Flanagin and Metzger (2000) point out that the major limitation of their study is the nonrepresentative sample of Web users. The Stanford Web credibility project is certainly broader in the representation of the variety of Web pages. There is a clear need for a comprehensive and systematic assessment of the validity of Web credibility measurement.

RESULTS

In the typical research study on Web credibility, the researchers sum the items to obtain a single Web credibility score. Researchers often report comparisons of both items and totals across different types of media or information.

COMMENTARY

It seems questionable to use single items to measure what appear to be complex perceptions on each dimension of Web credibility. The argument is a foundational question about the ability to fit the verbal universe to the observational universe (Miller & Boster, 1989, p. 19). A single word can seldom sum up the breadth or depth of most concepts. The concept of "accuracy," for example, has at least five major synonyms that capture substantively different aspects worthy of attention. When respondents check off "Not At All Accurate" when rating a Web page, we might wonder if the concern is with incorrect information, incomplete information, or just the meticulousness in the information provided. On a more pragmatic note, it is difficult to assess internal consistency (reliability) of ratings on a dimension with a single rating scale for that dimension. Textbooks that cover the basics of measurement often illustrate the point with some reference to assessing

a person's skills based on a single performance. While Pedhazur and Schmelkin (1991, p. 101) warn that increasing scale length may artificially increase reliability estimates, we cannot forget Epstein's (1979) dictum about having an adequate number of questions to adequately represent the measured concept.

COST

There are various versions of the Web credibility scale items readily available in print.

LOCATION

Various versions of the instrument are available in print, but there is sufficient information in this chapter for most researchers to construct their own versions of the scale. The original description is in Flanagin and Metzger (2000). An electronic version of the instrument is available from the author of this profile, so long as the users take personal responsibility for protecting the rights of anyone who might claim to hold a copyright to the scales.

REFERENCES

- Carter, R. E., & Greenberg, B. S. (1965). Newspapers or television: Which do you believe? *Journalism Quarterly*, 42, 29-34.
- Danielson, D. R. (in press). Web credibility. To appear in C. Ghaoui (Ed.), *Encyclopedia of Human-Computer Interaction*. Hershey, PA: Idea Group.
- Epstein, S. (1979). The stability of behavior: On predicting most of the people most of the time. *Journal of Personality and Social Psychology*, 37, 1097-1126.

1 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/web-credibility-measurment/20246

Related Content

Development of a Computerized Adaptive Test for a Large-Scale Testing Program: A Case Study of the Development of the North Carolina Computerized Adaptive Testing System

Lori McLeod, Albert Bethke, Cheryl Hill, Pamela Van Dyk and Kelly Burling (2006). *Online Assessment and Measurement: Case Studies from Higher Education, K-12 and Corporate* (pp. 130-142).

www.irma-international.org/chapter/development-computerized-adaptive-test-large/27670

A New MIM Directional Coupler With Twin Bands for Photonic ICs

Kondaveeti Muralikrishna, ShafiShahsavari Mirza and Satbir Singh Dhula (2020). *International Journal of Electronics, Communications, and Measurement Engineering* (pp. 30-40).

www.irma-international.org/article/a-new-mim-directional-coupler-with-twin-bands-for-photonic-ics/258315

Developing an Assessment Plan for the Arts and Sciences Division of a Comprehensive Community College

Annette E. Smith and Stephanie J. Jones (2013). *Cases on Assessment and Evaluation in Education* (pp. 380-399).

www.irma-international.org/chapter/developing-assessment-plan-arts-sciences/69497

Bit-Wise Coding for Exploratory Analysis

T. O'Daniel (2007). *Handbook of Research on Electronic Surveys and Measurements* (pp. 287-292).

www.irma-international.org/chapter/bit-wise-coding-exploratory-analysis/20244

Enhanced Detectability Using Multi-Cycle Cyclostationary Detector in Cognitive Radio

Shweta Singh and Rahul Kumar (2022). *International Journal of Electronics, Communications, and Measurement Engineering* (pp. 1-14).

www.irma-international.org/article/enhanced-detectability-using-multi-cycle-cyclostationary-detector-in-cognitive-radio/297082