Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 1999 Proceedings

Americas Conference on Information Systems (AMCIS)

December 1999

WebQual: A Web quality instrument

Eleanor Loiacono
University of Georgia

Follow this and additional works at: http://aisel.aisnet.org/amcis1999

WebQual: A Web quality instrument

Eleanor T. Loiacono
eloiacon@blaze.cba.uga.edu
Terry College of Business
Department of Management
University of Georgia
Athens, GA 30602
(706) 542-4653

Introduction

There does not exist a comprehensive measure to assess the quality of a Web site. To date many companies have based Web design on trial-and-error, gut-instinct, and feedback from customers. A more effective approach is to develop an instrument to measure Web site quality and to compare alternative designs. Both academic and popular trade publications have emphasized a variety of measures. Previous studies have attempted to measure the quality of a Web site via less direct measures such as the number of "hits" (Berthon, Pitt et al. 1996). The purpose of this research is: **1)** to develop a multiple-item instrument for measuring Web site quality (called WEBQUAL) and **2)** to report norms for some classes of Web sites.

Quality

"Quality in a product or service is not what you put into it. It is what the client or customer gets out of it." These words of Peter Drucker capture the essence of quality. The true definition of quality is a combination of two previously independent interpretations (Herbig and Genestre 1996): quality is consistent conformance to customer expectation (Crosby 1979) and "fitness for purpose" (Juran 1988). It is partially objective and partially subjective. The product or service must possess certain characteristics and be judged by customers to serve them in a way they want it to. There is no better judge of quality, therefore, than customers themselves.

Perceptions of Web site quality may include a number of dimensions. Facets, such as speed of download and use of multimedia, have been discussed in terms of viewer perceptions. Customers without the benefit of high-speed modems may experience long download times, which can cause dissatisfaction. This may lead to a decrease in perceived quality (Lightner 1996). The effective use of multimedia on Web sites is also a means to improve customer satisfaction and increase the perception of quality (Merritt 1996). However, because increasing levels of multimedia result in increased download times, the use of relevant and not excessive pictures, charts, and audio is necessary to maintain consumer satisfaction and perceived quality. Thus, Web site design involves a number of tradeoffs. Without a valid instrument, the extent to which these design tradeoffs effect the quality of a Web site cannot be determined.

WebQual Development

Prior to any data collection, the accuracy and validity of the instrument capturing that data must be demonstrated. It is for this reason that the suggested eight-step procedure for developing better measures by Churchill (1979) is followed.

1 Specify Domain of Construct

1.1 Stage One, Part One—Consumers: A series of exploratory research groups were conducted to generate criteria used by consumers in assessing Web site quality. Four groups of 13 to 20 people each, ages 18 to 25, where asked to participate on a voluntary basis. As an incentive, ten-dollar gift certificates to a local music store were awarded to four randomly selected participants. Each group reviewed one of four different types of Web sites. Two of these groups reviewed Web sites offering products (books and compact discs), while the other two groups reviewed service providers (airline and hotel reservations). After each participant ranked the ten Web sites using a Q-sort scale of –2 to +2, they were asked to provide reasons for their rankings (Stewart 1989).

Statistical analysis of the ten book Web sites included in the study revealed two distinct "factors" (clusters). Differences in the average ranking between the factors for the same Web site indicate that differences in the criteria used to assess quality exist. Three sites differed significantly in terms of their average ranking between the two factors. This prompted a further review of the written notes taken by subjects on each company during the experiment. Notes on one site in particular, which ranked lowest on Factor 1, clearly revealed a *dislike* for its appearance and services provided. Subjects referred to the

lack of hypertext links, email, colors, and visuals as support. Overall, subjects considered it "boring." One subject noted, however, that though it was boring, it was easily accessed. (The subject's Q-sort matched the Factor 2 Q-sort closely). This indicates that Factor 2 is related to functionality as an important determinant of Web site quality. Factor 1 considers such features as the layout, ease of navigation, and "extras" representative of higher quality Web sites.

1.2 Stage One, Part Two— Business Assessment of Consumer: Along with reviewing data gathered from consumers directly, MBA student teams (5 graduate students per team) participated in a class project for a Fortune 500 company that examined the Web strategies in 13 different industries. Each team was assigned a different industry and analyzed over ten companies in that industry. They based their research on four basic categories developed by the Fortune 500 company; technical accessibility, information, graphics, and personal biases, but were free to incorporate additional factors they deemed important. Overall, the findings of the graduate student teams were consistent with the factors gathered from the initial study. Visual appeal, navigability, and value added services to the customer indicated higher quality sites. On the other hand, clutter, lack of continuity, and ineffective links suggested a site of lower quality.

1.3 Stage One, Part Three—Designers: In addition to collecting data on consumer identified quality factors, criteria deemed useful by Web designers were also considered. Sources for this stage were collected from individual designers through a telephone interview conducted with 9 Web site designers and developers. The interviews focused on the designer's concern for quality. Each interview was tape recorded, transcribed, and a list of key learnings developed. The number of interviews was determined by continuing to interview until the information given by an additional Web designer became repetitive. Reviews of popular press Web site evaluation columns, such as Deconstruct featured in *Internet Week,* were conducted to identify any further factors. The major criteria given are graphic designs, ease of use, content, and download time. These factors relate to those categories used by judges to classify the comments gathered by consumers in stage one.

1.4 Stage Two--Focus Groups: Focus groups are planned to further identify and define those factors that indicate Web quality. Subjects who differ in their ranking on Factors 1 (appearance and layout) and 2 (functionality and download speed) revealed earlier will be brought in for further discussion. The goal is to understand why these discrepancies exist and what they mean. From this more focused research, a finer-grained definition and refinement of the quality factors can be produced. Items from this list

will form the basis of an initial WEBQUAL instrument, which will be iteratively refined.

2 Generate Sample of Items

The purpose of this step is to take the information gathered in the initial exploration of the construct and develop a sample of items (Churchill 1979). The items generated from the initial explorative research in phases one and two tap into the nuances of Web quality. The items will then be further edited to fully capture the essence of Web quality.

3 Collect Data

Data will be collected using the sample of items generated in step two. Web users will be polled to ensure generalizability to the larger population.

4 Purify Measure

During this step, the initial sample of items are purified and refined. Two statistical methods are performed at this stage of instrument development: coefficient alpha and factor analysis. Since the sample items are drawn from the domain of a single construct, they should be highly correlated. Coefficient alpha, which measures the internal consistency of a set of items, is the most appropriate means of determining the quality of an instrument (Churchill 1979). Factor analysis, which determines the number of dimensions underlying a construct, is used to confirm or refute components.

5 Collect Data

Data will be collected from Web users using the final purified sample of items generated in step four.

6 Assess Reliability

If an instrument is not reliable, one cannot have confidence in the relation between variables (Kerlinger 1986). Cronbach's alpha is the most commonly used measure of reliability (Churchill 1979; Goodhue 1998). Low alphas (generally below .50 or .60), the sample of items does not truly capture the construct for which it was meant to measure and thus is not a quality measure.

7 Assess Validity

Face, content, construct, external, and criterion validity will be tested.

8 Develop Norms

The final step in the WEBQUAL development will be to develop norms. The raw score on a measuring instrument is not particularly informative (Churchill 1979). Without an understanding of what the actual "norm" is, incorrect conclusions might be drawn. Thus, after WEBQUAL is developed, the average scores for

Web sites in different industries and an overall Web site rating will be determined through further testing. Even before the setting of norms is complete, WEBQUAL can be used to compare Web sites.

References

Berthon, P. R., L. F. Pitt, and R.T. Watson. (1996). "Marketing communication on the World Wide Web: what managers need to know about conversion and efficiency." Business Horizons **39**(5): 2-10.

Churchill, G. A. (1979). "A paradigm for developing better measures of marketing constructs." Journal of Marketing Research **16**(1): 64-73.

Crosby, P. B. (1979). Quality is Free. New York, McGraw-Hill.

Goodhue, D. (1998). "Development and measurement validity of a task-technology fit instruments for user evaluations of information systems." Decision Sciences **29**(1): 105-138.

Herbig, P. and A. Genestre (1996). "An examination of the cross-cultural differences in service quality: the example of Mexico and the USA." Journal of Consumer Marketing **13**(3): 43-51.

Juran, J. (1988). "Managing for quality." Journal for Quality and Participation **11**(1): 8-12.

Kerlinger, F. N. (1986). Foundations of Behavioral Research. Fort Worth, Harcourt Brace Jovanevich.

Lightner, N. J. (1996). What is wrong with the World Wide Web--a diagnosis of some problems and prescription of some remedies. Ergonomics. **39**(8)**:** 995-1004.

Merritt, J. (1996). Tour the campus without leaving home. Business Week (n3498)**:** 157.

Stewart, D. W. (1989). "Q methodology." Journal of Marketing Research **26**(2): 249-250.