THE ASSESSMENT OF WEB SITE QUALITY

(Research in Progress)

David Kopcso
Babson College
Babson Park, MA
02457
kopcso@babson.edu

Leo Pipino
University of
Massachusetts
Lowell
Lowell, MA 01854
Leo_Pipino@uml.edu

William Rybolt
Babson College
Babson Park, MA
02457
rybolt@babson.edu

ABSTRACT

The emergence of the World Wide Web has exacerbated the challenge faced by organizations in achieving and maintaining a state of high information quality. Since the web is the point of contact between transacting parties, it is not surprising that issues of quality of web sites arise. In this paper, we make use of a model of a web-based information system based on the classic Mason and Mitroff model and describe an exploratory experiment investigating the interactions of the variables in the model and their relationship to a standard set of information quality dimensions.

INTRODUCTION

Organizations continue to face the challenge of achieving and maintaining a state of high information quality. This challenge becomes increasingly formidable as the multimedia environment of the World Wide Web is used as a source of information and as a vehicle for external and internal business-to-business and business-to-consumer transactions. Since the web site is the point of contact between the transacting parties, it is not surprising that there is a great interest in assessing the quality of web sites and the quality of information on these web sites.

In theory, we recognize the distinction between quality of web site and quality of information on the web site. If one takes the limited view that quality of information equates to accuracy or correctness of information, then the distinction may be easily drawn. If, however, one views information quality as a multidimensional construct, as has been advocated elsewhere [Wang and Strong, 1996] and as we do in this paper, then the distinction between web site quality and the quality of information on the site becomes blurred. Just about everything on the web site can be considered information. For example, links to other pages and other sites are information. Likewise advertisements on the site are information. In reality, the information on the web site and the web site itself are so closely related that they almost become synonymous. The quality of the information on the web site is a function of such things as its timeliness, its accessibility, its correctness, its believability, and the like. Indeed, when judging the quality of a web site, much of the judgment is based on quality of the information as evaluated along a set of multiple criteria. The quality of the web site, however, depends not only on the information content but also on such variables as the aesthetics and design of the site, the goal of the visitors, the particular personal characteristics of the users, and the interaction of all these factors.

The quality of the web site affects how the user interacts with the site and, in turn, how easily the user can achieve his/her objective for visiting the site. It affects how readily the user can obtain and use the information on the site. Thus, the quality of the web site affects how well the site meets the user's needs, how well the site is accepted by users, and how often it is visited. These are all factors that are of importance to affect the provider. Among the provider's many objectives are those of meeting a user's needs and retaining the user as a customer or visitor.

Assessing web site quality is guided by such questions as: What factors affect a user's perceptions of a web site? How do these factors interact? Are the dimensions that have been found useful in assessing data quality for corporate, production databases and data warehouses applicable to the web environment? How do the perceptions match some standard measures? Are there standard sets of measures? How do the factors of aesthetics and web site design affect usability of information at the site and, by extension, the quality of information on the site?

In this ongoing research we adapt the classic Mason and Mitroff characterization of an information system (IS) to a web environment. Based on their model, we conducted a pilot experiment intended as a first step in examining the prior questions. The experiment was intended as a feasibility test for developing and conducting extended experiments in the future. As such it was exploratory in nature.

In this paper, we present a brief review of what has been done in the area of web site quality. We then review the Mason and Mitroff model, present our adaptation of the model to the web environment, and discuss the set of variables on which web quality assessments are dependent. We conclude with a presentation of the initial experiment and some descriptive statistical results obtained from the experiment.

BACKGROUND

The question of quality of web sites has been addressed from a number of perspectives. Among them are:

(a) ad hoc prescriptions for good web site design emphasizing the layout, the aesthetics, and the like. Examples of such prescriptive works are the many popular web design textbooks, such as Niederst [1999] and Nielsen [2000]. These usually are based on extensive personal experience of the individuals offering the advice.

(b) Traditional academic research has also addressed design issues [Zhang et al. 2000, Dran et al. 1999, Siau 1999, Gehrke and Turban 1999]. For example Zhang et al. have studied what features are important in different domains [Zhang, vonDran, Blake, and Pipithsuksunt, 2000]

(c) Research that concentrates on the end users' perceptions and evaluations of the quality of a web site. This work attempts to infer what the user "means" by quality based on the user's verbal responses and behavior. For example, Loiacono and Taylor [1999] reported initial attempts to develop a comprehensive web site quality measure by using exploratory groups to generate criteria that consumers use to assess web site quality. There also exists research on issues that are related to quality, such as, usability of web sites [Benbunan-Fich 1999, Myerschough 1999], and consumer satisfaction and web-site design [Lam and Lee 1999].

(d) Initiatives that attempt to determine the applicability of data quality research in the traditional database and data warehouse contexts to a web environment [Klein 1999, Klein 2000, Katerattanakul and Siau 1999]. In particular, these authors make use of the Wang and Strong dimensions [Wang and Strong, 1996].

(e) Commercial sites, such as *gomez.com*, that rate other sites.

(f) A number of marketing perspectives that focus on service quality, on classical marketing metrics that measure such variables as the number of visits, the number of clicks, site stickiness and the like, and on why and how individuals use the web.

The research and practices referenced above are of practical value and contribute to the general knowledge of web site quality. What is missing, however, is a solid theoretical foundation. Generalizing from the particular or using ad hoc prescriptions can mask the underlying dynamics and interaction of variables that affect the perception of a site's quality, the usefulness of a site, and achievement of user satisfaction.

Indeed, in much of the research on web site quality, an explicit definition of what is meant by quality is not given. In this paper, *we define a web site of quality as one that is fit for use*. This definition, equating quality with fitness for use, is in keeping with a large body of current research on data and information quality [Wang and Strong 1996; Strong, Lee, and Wang, 1997]. Our use of the term "fitness for use" is intended to be quite broad. Different users will assess fitness for use differently. Different users will define it differently. Users have different goals or objectives in visiting a web site. Users bring different modes of cognitive processing on their visits. The design and information presentation of different web sites vary widely. A web site's fitness for use and how its quality is assessed is dependent on a number of such interacting factors. These factors are the same factors that were used by Mason and Mitroff [1973] to characterize an information system in their classic paper. This seminal definition is as applicable today as it was when it was first enunciated. We adapt it to serve as our template for research on quality of web sites. The definition provides a basis on which to perform empirical research.

THE BASIC MODEL

Mason and Mitroff [1973] defined an information system as consisting of.

> *"at least one PERSON of a certain PSYCHOLOGICAL TYPE who faces a PROBLEM within some ORGANIZATIONAL CONTEXT for which he needs EVIDENCE to arrive at a solution (i.e. to select some course of action) and that the evidence is made available to him through some MODE OF PRESENTATION."*

The above characterizations had a strong influence on IS research. In the IS literature, a number of empirical studies investigating the effects and interactions of the original Mason and Mitroff factors on learning, performance, decision making, and processing strategies, to name a few, have been reported [Lusk and Kersnick 1979, Lucas and Nielsen 1980, Lucas 1981, Remus 1984, Remus 1987, Benbasat and Dexter 1985, Jarvenpaa 1989]. Recently, research on the affect of multimedia (a mode of presentation) in alleviating first impression bias was reported [Lim, Benbasat, and Ward 2000].

Although some may argue that older IS concepts no longer apply to the web, we do not agree. The World Wide Web is simply a highly distributed and highly complex information system. A web site is one component of an overall information system. We grant that the context in which this information system resides has changed since Mason and Mitroff first characterized an information system. Their basic definition, however,

remains valid and can easily be adapted to today's web environment. We adapt and extend their definition as follows:

> *A web-based information system consists of the PRESENTATION of INFORMATION in a specific way by a web site PROVIDER who launches and supports the web site for a PURPOSE. The web site is visited by a web site USER, of specific PSYCHOLOGICAL TYPE, who uses the web site to accomplish a specific TASK.*

We briefly elaborate on each of the relevant variables. The TASK can entail any one of a number of reasons for visiting a web site. It reflects the user's goal or purpose for visiting the site. For example, one task type might be to make a *purchase* at a specific web site. Another task type can simply be to *browse*, to access the web site for specific information. A wide variety of objectives constitute the concept of TASK.

The mode of PRESENTATION reflects how the information is presented. It is under this broad category that such important considerations as the layout, the design, and the aesthetics of the web site are captured. It is within this category that the newer technologies and multimedia environment manifest themselves.

The use of the term PSYCHOLOGICAL TYPE is intended to capture, in a summary fashion, the complex personal characteristics that are brought to the site by a user. These will include an individual's personal aesthetics and the user's experience with and attitude toward the technology. Most importantly, this factor will also include the user's cognitive processing style and problem solving approaches. From the perspective of IS practice, obtaining these styles may prove difficult, but not infeasible.

One expects interactions among the variables TASK, PRESENTATION, USER'S PSYCHOLOGICAL TYPE and the INFORMATION that is being sought, conveyed, or used. For example, if the site is intended to sell products to the user and the user visits the site to purchase products, then the mode of presentation should be compatible with this goal. What constitutes a compatible mode, however, will depend on the user's processing style and the information being used to accomplish the purchase.

In our characterization we have included the site PROVIDER as well as the USER in order to distinguish between their roles. Each will have a perception of the site's quality. By understanding the basis of the user's perception, the provider can modify his or her approach and adapt the web site to the intended consumer. This knowledge becomes of great importance as the thrust towards customization of web sites increases. Clearly the provider would like to have a match between the provider's perceptions and the user's perceptions. The research-in-progress described in this paper limits its focus to the end user and to the factors on which the user's perceptions are dependent. It does not examine the provider's perceptions. Comparisons of the two assessments must remain for future research.

ASSESSMENT OF QUALITY OF WEB SITE

In the traditional database environment, assessments of the quality of information are, generally, obtained as an evaluation of specific quality criteria or dimensions. Attempts are made to assess such criteria as the correctness, consistency, timeliness, accessibility, and ease of use of the information. Wang and Strong [1996] and Kahn, Strong, and Wang [1997] presented a

reasonable set of 16 dimensions that can be used to assess information quality. They are shown in Table 1.

Table 1: Data Quality Dimensions

Dimensions	Definitions
Accessibility	the extent to which data is available, or easily and quickly retrievable
Appropriate Amount of Data	the extent to which the volume of data is appropriate for the task at hand
Believability	the extent to which data is regarded as true and credible
Completeness	the extent to which data is not missing and is of sufficient breadth and depth for the task at hand
Concise Representation	the extent to which data is compactly represented
Consistent Representation	the extent to which data is presented in the same format
Ease of Manipulation	the extent to which data is easy to manipulate and apply to different tasks
Free-of-Error	the extent to which data is correct and reliable
Interpretability	the extent to which data is in appropriate languages, symbols, and units, and the definitions are clear
Objectivity	the extent to which data is unbiased, unprejudiced, and impartial
Relevancy	the extent to which data is applicable and helpful for the task at hand
Reputation	the extent to which data is highly regarded in terms of its source or content
Security	the extent to which access to data is restricted appropriately to maintain its security
Timeliness	the extent to which the data is sufficiently up-to-date for the task at hand
Understandability	the extent to which data is easily comprehended
Value-Added	the extent to which data is beneficial and provides advantages from its use

These 16 dimensions, or subsets thereof, have been successfully applied in a number of formal data quality improvement programs [Funk 1999, Davidson 1999, Kovac 1997]. These same criteria, broadly interpreted, can be used to assess the quality of a web site. We have asserted that a user's perception of web site quality is a function of four variables: task. presentation, information content, and user's psychological type. The user assessment of the quality of the web site, however, will typically be articulated in terms of one or more of 16 dimensions. For example, if the information obtained is not current for the user's needs it will be judged as not timely or if the user has difficulty navigating to obtain the information the user

desires, the site may be rated low in accessibility of information. To our knowledge, however, no studies to define and examine the underlying variables that affect the assessments of these dimensions by the user of the data have been reported in the literature.

Based on our adaptation of the Mason and Mitroff model, we propose that the end-user's perception of quality of web site is a function of the task to be performed, the information content available, the presentation of the information, the user's psychological type, in particular the user's cognitive processing style, and the interactions among these factors. This may be expressed in functional notation as

User's Perception of Quality =
f(Task, Presentation, Information Content, User's Psychological Type)

We should note that the assessment of overall quality of a web site is a function of the assessment of these 16 dimensions in the aggregate, or some variant combination of these, and possibly other dimensions. This leads one to the proposition that an aggregate, overall quality evaluation of a web site will be transitively dependent on the four factors above and their interactions. The question of the relationships between overall quality, the dimensions, and the factors, however, remains part of a stream of future research.

THE EXPLORATORY EXPERIMENT

The pilot experiment reported in this paper was intended to examine the feasibility of conducting a comprehensive controlled study. We performed the experiment in order to uncover potential procedural obstacles and to obtain insight that would be useful in the design of a comprehensive future study. As such, it was limited in scope, used a relatively small sample size of subjects, and was exploratory in nature.

Fifty-five full-time, undergraduate students, enrolled in a web technology class, were used as subjects. The subjects were asked to visit the sites of two specific airlines, each of which serviced the route chosen for the experiment. The subjects were asked to find a flight itinerary between two points. The itinerary was to have the minimum number of plane changes between the departure and arrival city. A hierarchy of rules to break any ties was also given the subjects as part of their instructions.

The airlines used were Southwest Airlines and United Airlines. The web sites of these two airlines are quite different. The subjects were also asked to evaluate each of the sites along a subset of the sixteen quality dimensions as well as the aesthetics of the site. The subject's solutions and the time to task completion were recorded.

All subjects were asked to complete the Keirsey Temperament Sorter [Keirsey and Bates 1984, Keirsey 1998]. The results of this instrument were used to categorize the subjects into different groups. The instrument measures four personality dimensions: Extroversion and Introversion (E/I), Intuition and Sensing (N/S), Thinking and Feeling (T/F), and Judgment and Perception (J/P) and is based on Jungian personality typology. It is related to the classic Myers and Briggs assessment instrument.

PRELIMINARY RESULTS

Table 2 shows the distribution of subjects along the four cognitive preferences: Extroversion/Introversion (E/I), Intuition/Sensing (N/S), Thinking/Feeling (T/F), and Judgment/Perception (J/P).

Table 2: Distribution of Cognitive Types

Cognitive Type	Number of Subjects
E/I	39/16
N/S	28/27
T/F	25/30
J/P	36/19

Although not perfectly uniform, a reasonable distribution of types was obtained.

Figure 1 shows a frequency count plotted against time to complete the task at each site. Although exact times were captured for each subject, to generate Figure 1 intervals of one minute were used. It is clear that the task was completed in much shorter times at the SW site in contrast to the UAL site

A statistical test was conducted to see if there was a statistical difference between the time to complete the task for the two sites for each individual. Formally, the null hypothesis is that the mean difference of time to completion is zero (times are the same) versus the alternative that the difference is not zero (times are not the same). The null can be rejected with a p-level of .0000. Since which site was visited first was randomly assigned to the groups, this result indicates that the intuitive conclusion that performance at the SW site was faster is statistically supported.

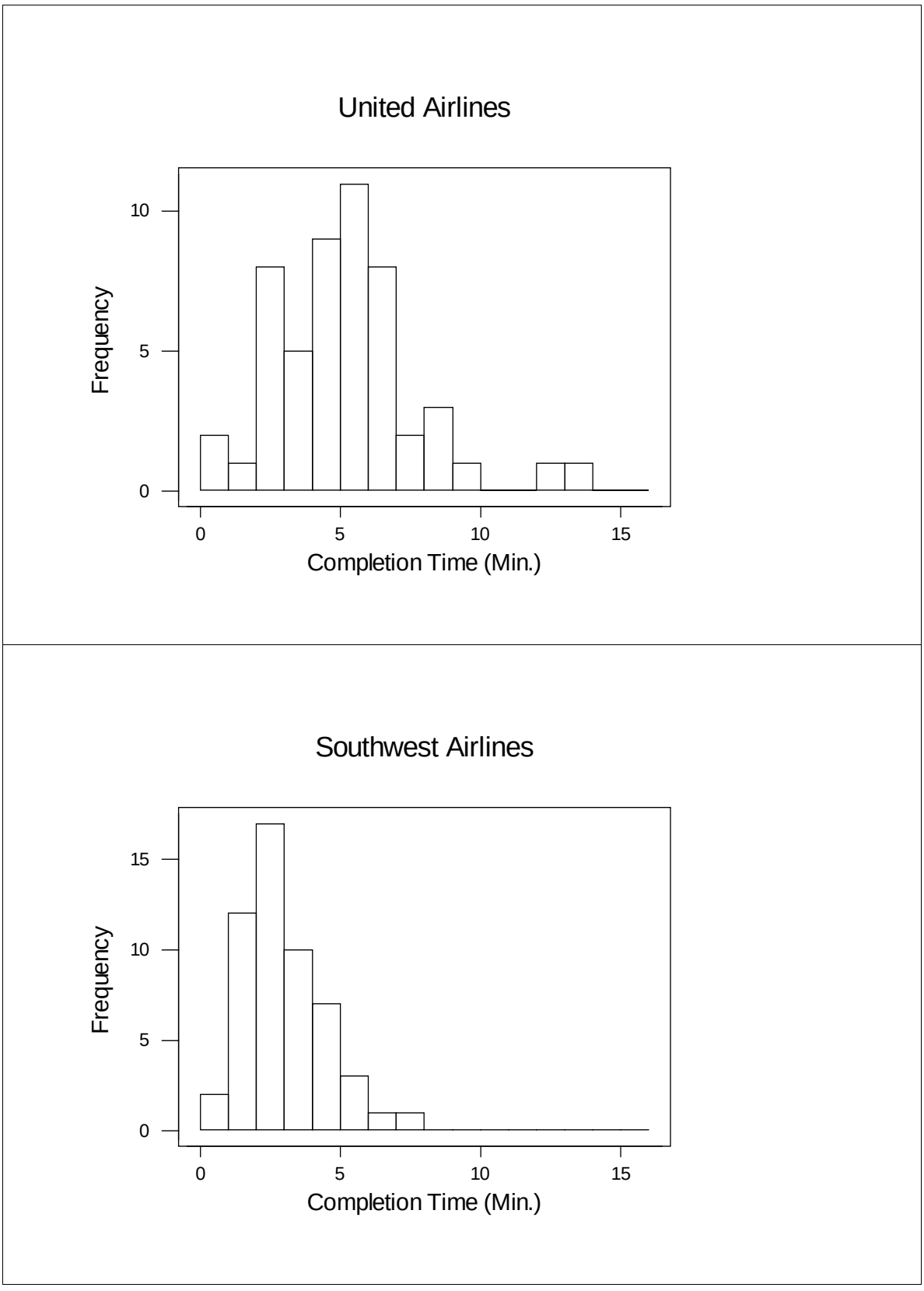

Figure 1:
Comparison of Time Interval to Complete Task - UAL Website and SW Website

Table 3:
Results of Tests of Independence: Dimensions vs. Cognitive Style (p-values)

Airline	Southwest				United Airlines			
Dimension	E/I	N/S	T/F	J/P	E/I	N/S	T/F	J/P
Ease of manipulation	.8293	.0198	.4822	.1708	.0703	.5956	.7168	.1256
Interpretability	.2548	.0046	.2851	.0507	.0038	.0265	.4606	.4813
Accessibility	.6425	.1175	.3419	.9397	.6452	.7741	1.000	.1865
Consistency	.2730	.4915	.4077	.9370	.2653	.5795	.0779	.3721
Completeness	.2313	.0001	.2629	.4143	.1598	.0221	.3637	.0150
Understandability	.3976	.1641	0.4583	0..4688	.2204	.1306	.2474	.1304
Efficient Use of Space	0.2552	0.024 5	0.4172	0.5629	0.0405	0.7967	0.9596	0.8041

One should note that, at the time of the experiment, the two web sites were chosen because of what the authors believe are sharp contrast in web site design and presentation. The SW web site was relatively a straightforward, uncluttered design without a large set of functions, and with a link specifically labeled "Schedule". This was in contrast to the UAL site, which had a denser home page and many more functions and options available. We expected that for the task assigned the SW site should be easier to use. Note that the commercial rating services have ranked the UAL site as superior to the SW site. For example see *gomez.com*. This is not unexpected when one considers the number of options that the UAL site makes available to the user. It also, however, highlights the importance of the task. If one is simply searching for flight information for a specific itinerary, then the simpler site may be superior. There is an interaction between task, presentation, and information. One size, one design does not fit all situations.

We are interested in examining the relationships between rating of quality dimensions and cognitive style. Recall that the presentation of the information was different at the two sites, the subjects had a range of cognitive profiles, and the task was held constant. For purposes of this feasibility experiment, ratings of the sites by the subjects along the following dimensions were obtained: accessibility, completeness, consistency, ease of manipulation, interpretability, and understandability. Tests for independence between each of the dimensions and cognitive styles produced p-values shown in Table 3. The results summarized in Table 3 indicate that there are dependency relationships between cognitive style and quality dimensions. The initial results suggest that further research is warranted.

The subjects were also asked to assess the "efficient use of space "on the web site, one criterion that relates to aesthetics. The results are included in Table 3. We anticipated that there would be a dependency between the space dimension and cognitive style. Again, the results indicate some dependence relationship between "effective use of space" and cognitive preference (See Table 3.)

CONCLUDING REMARKS

We have observed relationships between some of the quality dimensions and psychology types (cognitive style) of the individual users. The results are sufficiently encouraging to warrant further experimentation. The results of this "proof of concept" experiment will be quite helpful in designing a more elaborate study. It is clear from the results that the design (presentation) of the web site and the task are related. Also there are indications of dependencies between cognitive style, task, and web site design (presentation). It is possible that a different instrument to measure cognitive style would produce more discriminating results as might the use of a 7-point Likert Scale rather than the 4 point scale that was used. A larger sample of subjects also will be necessary.

In a future experiment, ratings along all sixteen dimensions as well as others of interest must be obtained. In the current multimedia environment, additional dimensions may exist that have significant affect on the perception of quality. This possibility will be subject of future research.

In addition, an overall rating of quality should be assessed in order to study the relationship of an overall rating of quality to the sixteen dimensions and the four basic variables of our model.

Finally, the differences in task and their effects must be examined. For the task we assigned the subjects, the more straightforward, direct approach allowed by the Southwest design, at the time of the experiment, appears more suitable than the United site. For a different task this may not be the case.

REFERENCES

Benbasat, I. and Dexter, A.S. "An Experimental Evaluation of Graphical and Color-Enhanced Information Presentation", *Management Science*, 31(11), 1985, 1348-1364.

Benbunan-Fich, R. "Methods for Evaluating the Usability of Web-Based Systems", *Proceedings of the Fifth Americas Conference on Information Systems*, 1999, 868-870.

Chun, A.T. and Davidson, B. "Implementing the Information Quality Survey: A Case Study at Cedars-Sinai Health System", *Proceedings of the 1999 Conference on Information Quality*, Cambridge, MA, October 1999, 266-284.

Funk, J., Lee, Y., and Wang, R. "Institutionalizing Data Quality Practice: The S.C. Johnson Wax Case", *Proceedings of the 1998 Conference on Information Quality*, Cambridge, MA, October 1998, 1-17.

Gehrke, D. and Turban, E. "Determinants of Successful Website Design: Relative Importance and Recommendations for Effectiveness", *Proceedings of Thirty-first Hawaii Conference on System Science*, January 1999.

Jarvenpaa, Sirkka L. "The Effect of Task Demands and Graphical Format on Information Processing Strategies", *Management Science*, 35(3), 1989, 285-303.

Kahn, B., Strong, D., and Wang, R. "A Model for Delivering Quality Information as Product and Service", *Proceedings of the 1997 Conference on Information Quality*, Cambridge, MA, October 1997, 80-94.

Katerattanakul, P. and Siau, K. "Measuring Information Quality of Web Sites: Development of an Instrument", *Proceedings of the Twentieth International Conference on Information Systems*, Charlotte, NC, 1999, 279-285.

Keirsey, David and Bates, Marilyn. *Please Understand Me: Character and Temperament Types*, Prometheus Nemesis Book Co., 5th Edition, 1984.

Keirsey, David. *Please Understand Me II: Temperament, Character, Intelligence*, Prometheus Nemesis Book Co., 1998.

Klein, B.D. "Perceptions of Information Quality: A Study of Internet and Traditional Text Sources", *Proceedings of the Fifth Americas Conference on Information Systems*, 1999, 618-620.

Klein, B.D. "Demographics, Experience, and Perceptions of Information Quality on the World Wide Web", *Proceedings of the Sixth Americas Conference on Information Systems*, 2000, 1062-1064.

Kovac, R., Lee, Y., and Pipino, L. "Total Data Quality Management: The Case of IRI", *Proceedings of the 1997 Conference on Information Quality*, Cambridge, MA, October 1997, 63-79.

Lam, J.C.Y. and Lee, M.K.O. "A Model of Internet Consumer Satisfaction: Focusing on the Web-site Design", *Proceedings of the Fifth Americas Conference on Information Systems*, 1999, 526-528.

Lim, K., Benbasat, I., and Ward, L., "The Role of Multimedia in Changing First Impression Bias", *Information Systems Research*, 11(2), June 2000, 115-136.

Loiacono, E.T. and Taylor, N.J. "Factors affecting Perceptions of Web Site Quality", *Proceedings of the Fifth Americas Conference on Information Systems*, 1999, 529-531.

Lucas, H.C. "An Experimental Investigation of the Use of Computer Based Graphics in Decision Making", *Management Science*, 27(7), 1981, 757-768.

Lucas, H.C. and Nielson, N. R. "The Impact of the Mode of Information Presentation on Learning and Performance", *Management Science*, 26(9), 1980, 982-993.

Lusk, E.J. and Kersnick, M. "The Effect of Cognitive Style and Report Format on Task Performance: The MIS Design Consequences", *Management Science*, 25(8), 1979, 787-798.

Mason, R.O. and Mitroff, I.I. “A Program for Research on Management Information Systems”, *Management Science*, 26, 1973, 475-485.

Myerscough, M. "Application Usage and Usefulness Assessment for Web-Based Intranet Systems (W-BIS) Applications", *Proceedings of the Fifth Americas Conference on Information Systems*, 1999, 880-882.

Nielsen, Jakob. *Designing Web Usability*, New Riders Publishing, Indianapolis, Ind., 2000.

Niederst, Jennifer. *Web Design in a Nutshell*, O'Reilly & Associates, Sebastopol, CA, 1999.

Remus, W. "An Empirical Investigation of the Impact of Graphical and Tabular Data Presentations on Decision Making", *Management Science*, 30(5), 1984, 533-542.

Remus, W. "A Study of Graphical and Tabular Displays and Their Interaction with Environmental Complexity", *Management Science*, 33(9), 1987, 1200-1204.

Siau, K. "Website Design: The Concepts of Informational and Computational Equivalence", *Proceedings of the Fifth Americas Conference on Information Systems*, 1999, 892-894.

Strong, D.M., Lee, Y.W., Wang, R.Y. "Data Quality in Context", *Communications of the ACM*, 40(5), 1997, 103-111.

von Dran, G.M., Zhang, P., and Small, R. "Quality Websites: An Application of the Kano Model to Website Design", *Proceedings of the Fifth Americas Conference on Information Systems*, 1999, 898-900.

Wang, R.Y. and Strong, D.M. "Beyond Accuracy: What Data Quality Means to Data Consumers", *Journal of Management Information Systems*, 12(4), 1996, 5-34.

Zhang, P., von Dran, G., Blake, P, and Pipithsuksunt, V. "A Comparison of the Most Important Website Features in Different Domains: An Empirical Study of User Perceptions", *Proceedings of the Sixth Americas Conference on Information Systems*, 2000, 1367- 1372.