

**FACTORS AFFECTING THE ASSESSMENT OF WEB SITE QUALITY
[RESEARCH IN PROGRESS]**

David Kopcso

Mathematics and Science Division, Babson College, Babson Park, MA, USA
Tel.: 1 (781) 239-4409
kopcso@babson.edu

Leo Pipino

College of Management, University of Massachusetts Lowell, Lowell, MA 01854, USA
Tel.: 1 (978) 934-2759, Fax: 1 (978) 934-4034
Leo_Pipino@uml.edu

William Rybolt

Mathematics and Science Division, Babson College, Babson Park, MA, USA
Tel.: 1 (781) 239-4242
rybolt@babson.edu

ABSTRACT

In the research-in-progress reported here, we adapted the classic Mason and Mitroff characterization of an information system to a web environment. The resulting definition was used to design a set of pilot experiments to examine the relationship between quality dimensions and the variables of the modified Mason and Mitroff definition. The experiments were intended as feasibility tests whose purpose was to form the basis for a comprehensive, controlled future experiment.

1. INTRODUCTION

Use of the World Wide Web as a source of information and as a vehicle for business to business and business to consumer transactions compels a close look at the quality of web sites and the quality of information on web sites. Although the quality of a web site is often assessed in terms of its aesthetics and design and/or traditional dimensions, such as ease of use, accessibility, and interpretability, we suggest that these ratings are dependent on a set of more fundamental variables.

The question of quality of web sites has been addressed from a number of perspectives. Among them are: (a) works that present prescriptions for good web site design [19], [20], (b) academic research that has addressed design issues [21], [8], [23], (c) work that focuses on the end users' perceptions and evaluations of the

quality of a web site [4], [6], [16], [17]. Of note in this area is the work on the WebQual instrument [1], [2], [3], which is intended to assess the user's perception of the quality of the web site. This instrument has been evolving and is currently in its fourth version, WebQual 4.0 [3]. (d) work that examines communication effectiveness and benefits delivered by commercial sites, (e) investigations into the applicability of data quality research in the traditional database and data warehouse contexts to a web environment [9], [12], [13], and (f) commercial sites, such as *gomez.com*, that rate other sites.

What is missing from the above initiatives is a theoretical foundation or framework that captures both the key variables and the underlying dynamics and interactions of variables that affect the perception of a site's quality. *We define a web site of quality as one that is fit for use.* We intend the term "fitness for use" to be quite broad. Different users will assess fitness for use differently. Fitness for use can be related to classic characterization of an information system given by Mason and Mitroff [18]. The user's assessment of fitness for use and web site quality will be dependent on the variables identified by Mason and Mitroff in their definition. In this research we adapt the Mason and Mitroff characterization to a web environment and use it as a guide for our empirical investigation of web site quality.

2. THE BASIC MODEL

Mason and Mitroff [18] 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."

Although the context in which information systems reside may have changed over time, Mason and Mitroff's characterization remains valid and is easily adapted to today's web environment. We characterize a web-based information system 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.

The TASK reflects the user's goal or purpose for visiting the site. In this paper, we are interested in the user's perspective. One would hope that the user's view of the task is compatible with the provider's purpose(s) for launching and maintaining the site. The mode of PRESENTATION represents the manner in which the information is presented. PSYCHOLOGICAL TYPE is intended to reflect the complex personal characteristics, including cognitive processing style, of the user. We would expect interactions among the variables TASK, PRESENTATION, USER'S PSYCHOLOGICAL TYPE and the INFORMATION that is being sought, conveyed, or used.

The definition above includes the site PROVIDER as well as the USER. It is important to distinguish between their roles. Although the research-in-progress reported in this paper has focused on the user perceptions, the provider is nonetheless important. Clearly the provider would like to have a match between his or her perceptions and the user's perceptions. With the trend towards web site customization, this match is of great import. As a first step, however, we limited our research to the user's perceptions of quality.

Note that the model proposed can accommodate the different perspectives of the research cited earlier. For example, the WebQual 4.0 instrument [3] contains questions that focus on three of the major dimensions: site quality, information quality, and interaction quality. These roughly correspond to the factors: presentation, information, and cognitive style of user, respectively. Interaction between user and site/provider will be dependent on cognitive style of user.

The end-user's perception of a specific quality dimension can be expressed as follows:

User's Perception of Specific Quality Dimension =

$$f(\text{Task, Presentation, Information Content, User's Psychological Type})$$

Typically, the assessment of quality is measured along a set of dimensions. In the traditional file and database environment, a number of dimensions have been suggested to measure data quality. For example, Wang and Strong [22] have suggested 16 dimensions. These dimensions, or subsets thereof, have been used in a number of data quality improvement programs [5], [7], [15]. We choose the following subset for our experiments: accessibility, completeness, consistency, ease of manipulation, interpretability, and understandability.

3. THE EXPLORATORY EXPERIMENTS

The experiments reported in this paper were intended as "proof of concept investigations". We performed the experiments 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, they were limited in scope, used a relatively small sample size of subjects, and were exploratory in nature.

The first experiment was conducted using undergraduate students, enrolled in a web technology course as subjects. Subsequent to the first experiment, a follow-up second experiment using graduate students enrolled in a web technology course was conducted. Preliminary results of the first experiment were first reported at the Conference on Information Quality [14]. In this section we summarize the results of the first trial, present results of the second trial, and compare the results.

In both experiments, 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 asked to evaluate each of the sites along the six quality dimensions listed earlier. Subjects were also asked to rate the use of space on the site. In the case of the second experiment, the subjects were also asked to give an overall rating of quality of the two sites.

All subjects completed the Keirsey Temperament Sorter [10], [11]. 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). It is based on Jungian personality typology and is related to the Myers and Briggs instrument.

4. PRELIMINARY RESULTS

Table 1 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) for the two experiments. Note that there were 55 usable data points in experiment 1 and 62 in experiment 2.

Cognitive Type	Experiment 1 Number of Subjects	Experiment 2 Number of Subjects
E/I	39/16	42/20
N/S	28/27	20/42
T/F	25/30	33/29
J/P	36/19	50/12

Table 1: Distribution of Cognitive Types - Experiment 1

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

One should note that, at the time of the experiments, the SW web site was relatively a straightforward design with a link specifically labeled "Schedule". This was in contrast to the UAL site, which had a busier home page and no clearly identified direct link to schedule information. We anticipated that the SW site would be easier to use than the UAL site for the assigned task.

On the average students took longer to complete the task on the United Airlines site than on the Southwest site. For example, in experiment 2 visitors to the SW site took on the average 4 minutes and 19 seconds to complete the task while visitors to the UAL took 5 minutes and 55 seconds to complete the task. Clearly, this measure of time to completion indicates that the SW site is superior to the UAL for the assigned task. The Presentation does matter.

We were interested in examining the relationships between rating of quality dimensions and cognitive style. Tests for independence between each of the dimensions and cognitive styles produced p-values shown in Table 2 for experiment 1 and Table 3 for experiment 2. Recall that subjects were also asked to assess the "efficient use of space" on the web site, one criterion that relates to aesthetics. For simplicity of discussion, we have included "Efficient Use of Space" in the tables. Note that the test for independence between cognitive style and the overall quality rating obtained in experiment 2 has been included below Table 3.

Airline	Southwest				United Airlines			
	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	.8293	.0198	.4822	.1708	.0703	.5956	.7168	.1256
Consistency	.2730	.4915	.4077	.9370	.2653	.5795	.0779	.3721
Completeness	.2313	.0001	.2629	.4143	.1598	.0221	.3637	.0150
Understandability	.3976	.1641	.4583	.4688	.2204	.1306	.2474	.1304
Efficient Use of Space	.2552	.0245	.4172	.5629	.0405	.7967	.9596	.8041

Table 2: Results of Tests of Independence: Dimensions vs. Cognitive Style (p-values) - Experiment 1

Airline	Southwest				United Airlines			
	E/I	N/S	T/F	J/P	E/I	N/S	T/F	J/P
Ease of manipulation	.5527	.1962	.9262	.3287	.9633	.2625	1.0000	.8021
Interpretability	1.000	.8973	.6685	.1903	1.0000	.0813	.7348	.0582
Accessibility	.4559	.6949	.7439	.6220	1.0000	.4688	.2502	.7987
Consistency	.2669	.7294	.5256	.9295	.5039	.8469	1.0000	.2152
Completeness	.4916	.0759	.7966	.4475	.5920	.4898	.4898	.1203
Understandability	.5533	.3641	1.0000	.1843	.6625	.0248	.0523	.2201
Efficient Use of Space	.6397	.4456	.7937	.6727	.4777	.6141	.2216	.2175

Quality	.2308	.8668	.4943	.0658	.5713	.5192	.0326	.0237
---------	-------	-------	-------	-------	-------	-------	-------	-------

Table 3: Results of Tests of Independence: Dimensions vs. Cognitive Style (p-values) - Experiment 2

The results summarized in Table 2 indicate that there are dependency relationships between cognitive style and quality dimensions. There are 56 cells in Table 2. One would expect on the basis of chance only three of the cells to have p values less than $\alpha = 0.05$. Since nine of the cells have such values, this suggests that there is indeed a relationship between cognitive style and quality dimensions.

Of the same 56 cells in our second experiment, only one had a p value less than $\alpha = 0.05$. This result is contrary to our expectation, based on the first experiment, that we would obtain more dependency relationships.

We had naively expected that the intuitive/sensing dimension would have the greatest influence on the quality dimensions. It was encouraging, in experiment 1, to note that six of the nine cells involving statistically significant relationships fell in this category. The replication attempt of experiment 2, however, resulted in only 1 statistically significant instance along the intuitive/sensing dimension.

Additional analyses of the data did not yield any interesting or reportable results. Cluster analyses of the two experimental groups were performed. No meaningful or insightful results were obtained.

4. SUMMARY

We have observed relationships between psychological type and some of the quality dimensions. Although the second experiment yielded less encouraging results than experiment 1, further experimentation is warranted. The results of these "proof of concept" experiments will help to improve the design of a more elaborate study. It is clear from the results that the design (presentation) of the web site and ease of performing the task are related. We are investigating the use of a different instrument to measure psychological type (cognitive style). In addition, we are considering the use of the WebQual instrument [3], which has been validated, to measure the quality dimensions.

The web has numerous examples of sites that customize content to better fit the user's needs. We believe that the next major innovation will be to improve the user experience through the customization of the web site. To do so, the site designer will have to consider a presentation mode and a site design that is compatible with the objective of the visit (task) and the user's cognitive style.

REFERENCES

- [1] Barnes, S. and Vidgen, R., "WebQual: An Exploration of Web-site Quality," *Proceedings of the Eighth European Conference on Information Systems*, Vienna, July 2000.
- [2] Barnes, S. and Vigden, R., "Assessing the Quality of Auction Web Sites," *Proceedings of the Hawaii International Conference on Systems Sciences*, Maui, Hawaii, January 2001.
- [3] Barnes, S. and Vidgen, R., "An integrative approach to the Assessment of E-Commerce Quality," Available online at <http://www.bath.ac.uk/~mnsrtv/research/webquality.htm>.
- [4] Benbunan-Fich, R. "Methods for Evaluating the Usability of Web-Based Systems," *Proceedings of the Fifth Americas Conference on Information Systems*, 1999, 868-870.
- [5] 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.
- [6] Eighmey, J., "Profiling User Responses to Commercial Web Sites," *Journal of Advertising Research*, May-June 1997, 59-66.
- [7] 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.

- [8] 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.
- [9] 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.
- [10] Keirse, David and Bates, Marilyn. *Please Understand Me: Character and Temperament Types*, Prometheus Nemesis Book Co., 5th Edition, 1984.
- [11] Keirse, David. *Please Understand Me II: Temperament, Character, Intelligence*, Prometheus Nemesis Book Co., 1998.
- [12] 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.
- [13] 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.
- [14] Kopcsó, D., Pipino, L., and Rybolt, W. "The Assessment of Web Site Quality," *Proceedings of the 2000 Conference on Information Quality*, Cambridge, MA, October 2000.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] Mason, R.O. and Mitroff, I.I. "A Program for Research on Management Information Systems", *Management Science*, 26, 1973, 475-485.
- [19] Nielsen, Jakob. *Designing Web Usability*, New Riders Publishing, Indianapolis, Ind., 2000.
- [20] Niederst, Jennifer. *Web Design in a Nutshell*, O'Reilly & Associates, Sebastopol, CA, 1999.
- [21] 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.
- [22] 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.
- [23] 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.