Association for Information Systems
AIS Electronic Library (AISeL)

AMCIS 2005 Proceedings

Americas Conference on Information Systems (AMCIS)

2005

An Experimental Study of the Effects of Representational Data Quality on Decision Performance

Wonjin Jung
Claremont Graduate University, jungw@cgu.edu

Lorne Olfman
Claremont Graduate University, lorne.olfman@cgu.edu

Terry Ryan
Claremont Graduate University, terry.ryan@cgu.edu

Yong-Tae Park
California State University - Fullerton, ypark@fullerton.edu

An Experimental Study of the Effects of Representational Data Quality on Decision Performance

Wonjin Jung
Claremont Graduate University
jungw@cgu.edu

Lorne Olfman
Claremont Graduate University
Lorne.Olfman@cgu.edu

Terry Ryan
Claremont Graduate University
Terry.Ryan@cgu.edu

Yong-Tae Park
California State University, Fullerton
ypark@fullerton.edu

ABSTRACT

The effects of information quality and the importance of information have been reported in the Information Systems literature. However, little has been learned about the impact of data quality (DQ) on decision performance. Representational DQ means that data must be interpretable, easy to understand, and represented concisely and consistently. This study explores the effects of representational DQ and task complexity on decision performance by conducting a laboratory experiment. Based on two levels of representational DQ and two levels of task complexity, this study had a 2 x 2 factorial design. The dependent variables were problem-solving accuracy and time. The results demonstrated that the effects of representational DQ on decision performance were significant. The findings suggest that decision makers can expect to improve their decision performance by enhancing representational DQ. This research extends a body of research examining the effects of factors that can be tied to human decision-making performance.

Keywords

Representational data quality, task complexity, decision performance.

INTRODUCTION

According to the knowledge management literature, data is a prerequisite for information and information can be created from its raw data. Tuomi (1999, p. 103) states: "The generally accepted view sees data as simple facts that become information as they are combined into meaningful structures, which subsequently become knowledge as meaningful information is put into a context and when it can be used to make predictions." While the effects of information quality and the importance of information have been studied in the IS literature, little has been learned about the impact of data quality (DQ) on decision performance. Thus, the purpose of this study is to empirically examine the relationship between data quality and decision performance. To examine the relationship, it is necessary to understand what data quality means to data users. Wang and Strong (1996) developed a hierarchical framework that captures the aspects of data quality that are important to data users. The data quality categories are: intrinsic, contextual, representational, and accessibility. In their study, representational DQ emphasizes that data must be presented in such a way that they are interpretable, easy to understand, and represented concisely and consistently.

Graphical information representation research is of interest to many disciplines, such as Statistics, Psychology, Education, Engineering, Management, and Information Systems (Tan and Benbasat, 1990). Traditionally, problem solvers have relied on graphical information representations in improving decision quality (Smelcer and Carmel, 1997). They make faster, more accurate decisions when their information is presented in a format that best matches the characteristics of their task (Benbasat and Dexter, 1985; DeSanctis, 1984; Jarvenpaa and Dickson, 1988; Jarvenpaa, Dickson, and DeSanctis, 1985). However, it is not well understood whether graphical data representations affect problem-solving performance in decision-making settings. Hence, it would be worth investigating the effect of representational DQ on decision performance. In addition, much of the information representation research did not manipulate task complexity, nor did it show performance differences between individuals based on the level of task complexity. Therefore, this study, which is part of a larger study examining the effects of various aspects of data quality and task complexity on decision performance, empirically explores how representational DQ and task complexity simultaneously affect decision performance.

THEORETICAL BACKGROUND AND HYPOTHESIS DEVELOPMENT

Information Quality

In the IS literature, information quality is one of two major dimensions for evaluating the success of information systems (DeLone and McLean, 1992, 2003). Iivari and Koskela (1987) used various information quality criteria to measure users' information satisfaction. Their items included relevance, comprehensiveness, recentness, accuracy, credibility, convenience, timeliness, interpretability, and adaptability. More recently, IS researchers examined the relationship between information quality and individual performance (Etezadi-Amoli and Farhoomand, 1996; Teo and Wong, 1998; Wixom and Watson, 2001). Their studies provided strong support for the effects of information quality on individual performance.

Representational Information

Tan and Benbasat (1990, p. 417) state: "There is now common agreement in the Information Systems (IS) graphics research literature that the quality of a given representation depends on the characteristics of the task to which it is to be applied (Benbasat and Dexter, 1985; DeSanctis, 1984; Javenpaa and Dickson, 1988; Jarvenpaa et al., 1985)." Vessey (1991) also suggests that a decision-maker's task processing would be more efficient and effective when there is a cognitive fit (match) between the information emphasized in the representation type and that required by the task type. That is, the theory of cognitive fit focuses on the effect of a match between problem representation and task on problem-solving performance: spatial tasks need spatial representations; symbolic tasks need symbolic information. Specifically, while tables emphasize symbolic information and lead to better performance for the task of reading specific data values, graphs emphasize spatial information and lead to better performance for most elementary spatial tasks, including summarizing data, showing trends, comparing points and patterns, and showing deviations (Jarvenpaa and Dickson, 1988; Vessay, 1991). Previous research on the graphical representation of information developed a sound taxonomy for classifying tasks: elementary tasks or decision activities (Newell and Simon, 1972). Elementary tasks include basic perceptual cognitive information processes (e.g., retrieval of a data value or comparison of two data values). On the other hand, decision activities include higher mental processes such as judgment, integration of information, and inference (e.g., forecasting).

Chandra and Krovi (1999) extended the theory of cognitive fit to account for the congruence between external representation (e.g., information organization) and internal representation, and tested their extended model in an experimental setting with the two models of external representation (prepositional networks model from the cognition literature and object-oriented model from the systems literature). Chandra and Krovi state (1999, p. 273): "While the cognitive fit is an excellent framework for understanding the relationship between problem representation and problem-solving task, it does not explicitly account for specific internal representations and their effect on the efficiency and effectiveness of information retrieval." The logic in their model is that if an already existing knowledge structure (internal representation) is congruent with information organization, the problem solver is better able to match the latter to the internal knowledge, thereby leading to the better efficiency and effectiveness of information retrieval performance. Overall findings of their study provide some evidence that the retrieval process benefits when information organization is congruent with internal representation.

Similarly, research in cognition and human information processing suggests that designing for comprehension is an effective way to reduce a reader's mental efforts to understand the contents of a document (Thuring et al., 1995). The nature of the information retrieval process is likely to differ from managerial problem-solving. However, if the system presents data necessary to solve problems in such a way that they are organized, interpretable, and easy to understand (e.g., high-quality representational data), it would create a congruence between external (information organization) and internal representation. As such, it could be possible to infer that problem-solving performance can be improved due to the congruence leading to the better efficiency and effectiveness of retrieval process for the information necessary to make decisions.

Again, when there is a fit of representation and task type, each representation will lead to both quicker and more accurate problem-solving (Vessey, 1991). Likewise, when a person is given high-quality representational data that best match the experimental problem-solving tasks, a positive effect of the representational DQ on decision performance is expected. Based on the discussion above, the following hypotheses are proposed.

H1: Regardless of the levels of task complexity, subjects with high-quality representational data will require less problem-solving time than subjects with low-quality representational data.

H2: Regardless of the levels of task complexity, problem-solving with high-quality representational data will lead to an increase in problem-solving accuracy compared to problem-solving with low-quality representational data.

Task Complexity

Task complexity is defined as the degree of cognitive load or mental effort required to identify and/or solve a problem (Payne, 1976). According to Campbell's concept of task complexity (1988), tasks that increase information load and information diversity are considered as complex tasks. Wood (1986) suggests that complexity is a function of the number of acts that must be executed and the number of information cues that must be processed when performing a task. Thus, tasks are considered more complex as the number of acts and information cues increases. In an information retrieval context, task complexity increases as the number of potential solutions increases because decision makers must evaluate each potential solution if they want to get the most effective or accurate result (Newell and Simon, 1972). Rossano and Moak (1998) also suggest that mental workload increases as more data are evaluated and retained in working memory. Based on the discussion above, the following hypotheses are proposed.

H3: Regardless of the levels of representational data quality, subjects with a simple task will require less problem-solving time than subjects with a complex task.

H4: Regardless of the levels of representational data quality, subjects with a simple task will make more accurate decisions than subjects with a complex task.

RESEARCH METHODOLOGY

Experimental Design

To examine the effects of representational data quality DQ and task complexity, a laboratory experiment was conducted. Based on the two factors, representational DQ (high vs. low) and task complexity (simple vs. complex), a 2 x 2 factorial design was implemented to test the hypotheses. Two attributes of data quality (table and graph) were used to map to the data type. Each subject's decision performance was assessed based on predetermined measurement, and decision performance referred to solution time and the accuracy of problem-solving. A Web-based simple system to deliver the representational data to the subjects was developed using the latest version of Web programming languages, Hyper Text Markup Language (HTML) and Practical Extraction and Report Language (PERL).

Procedures

The subjects were assigned randomly to one of the four treatments. The experimental task and a set of data were given to them. The task for this study asked subjects to solve a decision problem. In order to help subjects understand the decision-making rules for the task, an example to simulate the decision-making rules was provided. After that, the subjects were provided with an answer sheet to record their solutions as they performed the task. Next, with the data set and the task, the subjects made decisions. Finally, this study observed the effects of the various treatments on decision performance.

Independent Variables

Two levels of representational DQ and two levels of task complexity were operationalized as independent variables.

Representational Data Quality

The first independent variable is representational data at two levels of quality, referred to as high and low. As discussed earlier, previous research on the graphical representation of information developed a sound taxonomy for classifying tasks: elementary tasks or decision activities (Newell and Simon, 1972). The decision tasks used here, with known solutions, were close to decision activities rather than elementary tasks or spatial tasks in terms of difficulty, requiring higher mental processes and managerial analysis such as judgment, integration of information, and inference. The experiment used two different representations (graph and table). While a table presents data as a series of discrete numbers, a graph presents data as a series of colors or patterns (Vessey, 1991). This study carefully constructed the table and graph representations to contain the same data. That is, the two types of data presentations provided equivalent values, except in the data presentation format. Since the experimental decision-making tasks for this study were close to decision activities rather than elementary or spatial tasks, based on the theory of cognitive fit (Vessey, 1991), it was believed that there was a cognitive fit between data in the form of tables and the experimental decision-making tasks. That is, tables were expected to facilitate the tasks' solution and to produce superior performance than graphs. Hence, tables were considered to provide more high-quality representational data than graphs for the tasks. As such, the level of representational DQ was manipulated by the table and graph representations (see Appendix A).

Task Complexity

The decision task created by Jarvenpaa (2003) was used for this laboratory experiment, with some minor adjustments. It asks subjects to select a site for the construction of a Chinese restaurant. While the complex task asked subjects to select a site from among five alternative sites in which to locate a Chinese restaurant, the simple task asked subjects to select a site from among three alternative sites. The complex task had five factors for each site, while the simple task had three factors for each site. The factors were very important in deciding where the restaurant should be located. The scores for the factors were predetermined.

Two levels of task complexity (high and low) were used for this study. The degree of task complexity was manipulated by the number of problems in the task. The task required simple arithmetic calculations based on the decision criteria (factors) and decision choices (alternative sites for the restaurant). Specifically, the simple task with 24 problems required subjects to sum scores over three years for each factor. After averaging the summed scores for each factor, subjects were asked to sum the average scores for each site. Finally, they were asked to select a site that overall performs the best from among three alternative sites.

The complex task with 80 problems required subjects to average scores over three years for each factor. In addition, subjects were asked to assign a weight for each factor. After that, they were asked to evaluate the sites by pair-wise comparison (always comparing two sites at a time) with the weighted scores and select the site that wins the last comparison by having the largest number of factors of higher weighted value. That is, subjects were requested to rank the sites according to the predefined decision rules and the weighted scores of each factor.

Dependent Variables

The dependent variable of this study is decision performance. Decision performance was operationalized as the accuracy of problem-solving and solution time. Problem-solving accuracy was measured by the number of correct answers from the correct solutions. Solution time was measured from the time when the subjects began working on the task until they recorded their solutions on their answer sheet and logged out of the system. Because time factors, pressure or constraints, affect decision-making (Ahituv et al., 1998; Austin, 2001), subjects were not informed of any time expectation for this experiment.

RESEARCH FINDINGS

A total of 40 undergraduate students from various academic programs at California State University, Fullerton, and California State Polytechnic University, Pomona, participated in the experiment. Of the participants, 60 percent were male, and 60 percent were younger than age 25. The average age of participants was 24.6 years. The number of years in college was 2.8 years. Two-thirds of the participants were majoring in business administration. Problem-solving accuracy and time were each analyzed with two-way ANOVAs. The tests were carried out at a 95% confidence level. The descriptive statistics for the dependent variables are summarized in Table 1.

Measures	**Treatment Conditions**			
	Simple Task		Complex Task	
	High Rep. DQ	Low Rep. DQ	High Rep. DQ	Low Rep. DQ
Solution Accuracy: (a higher score implies greater accuracy)				
Mean	97.917	44.583	97.917	55.250
Std. Dev.	4.0493	25.3106	1.6536	24.3228
n	10	10	10	10
Solution Time: (minutes: seconds)				
Mean	0:07:59	0:12:18	0:22:22	0:25:12
Std. Dev.	0:02:40	0:04:29	0:06:59	0:04:21
n	10	10	10	10

Table 1. Descriptive Statistics for Problem-Solving Performance

The results of the two-way ANOVA for time showed that the main effects of representational DQ ($p = .002$) and task complexity ($p = .000$) were significant (see Table 2). The results indicated that the simple task was solved more quickly than the complex task. Therefore, H3 was supported. Also consistent with expectations, subjects using high representational DQ took less time than subjects using low representational DQ. That means, problem-solving time with high representational DQ was significantly shorter than with low representational DQ. Therefore, H1 was supported.

Source	Type III Sum of Squares	Mean Square	F	Sig.
Corrected Model	19767641.737(a)	6589213.912	44.718	.000
Intercept	132295964.112	132295964.112	897.832	.000
COMP	18018663.612	18018663.612	122.284	.000
REP	1555425.313	1555425.313	10.556	.002
COMP * REP	193552.812	193552.812	1.314	.255
Error	11198639.150	147350.515		
Total	163262245.000			
Corrected Total	30966280.887			

(a) R Squared = .638 (Adjusted R Squared = .624)

Table 2. ANOVA Table for Two-Way Analysis of Problem-Solving Time: Representational DQ by Task Complexity

The ANOVA on problem-solving accuracy found a significant main effect for representational DQ (p = .000, see Table 3). Subjects using high representational DQ made more accurate decisions than subjects using low representational DQ. Therefore, H2 was supported. Surprisingly, the results of ANOVA for problem-solving accuracy showed that there was no significant main effect of task complexity for problem-solving accuracy (p = .385, see Table 3). Subjects completing the complex task had comparable problem-solving accuracy to those completing the simple task. That means, the subjects assigned to the complex task were evidently able to handle additional task complexity without significant detriment to problem-solving accuracy. Thus, H4 was rejected. However, it is interesting to note that the main effect of task complexity was significant for problem-solving time. These confounding results suggest that because there was no time constraint, that is, there was no specific allotment of time for making a decision, subjects used as much time as they needed to complete the complex task while keeping problem-solving accuracy as high as possible. Therefore, these unexpected results may imply the existence of accuracy-time trade-offs only in the effect of task complexity.

Source	Type III Sum of Squares	Mean Square	F	Sig.
Corrected Model	30471.753(a)	10157.251	16.022	.000
Intercept	271833.472	271833.472	428.799	.000
COMP	483.472	483.472	.763	.385
REP	29548.828	29548.828	46.611	.000
COMP * REP	439.453	439.453	.693	.408
Error	48179.497	633.941		
Total	350484.722			
Corrected Total	78651.250			

(a) R Squared = .387 (Adjusted R Squared = .363)

Table 3. ANOVA Table for Two-Way Analysis of Problem-Solving Accuracy: Representational DQ by Task Complexity

For problem-solving time, the interaction between representational DQ and task complexity was not significant (p = .255, see Table 2), indicating these two variables do not jointly affect problem-solving time. Table 4 shows that a comparison involving high and low representational DQ in the effect of the complex task showed no significant mean difference on problem-solving time. This suggests that high representational DQ did not provide benefits to problem-solving time in the effect of the complex task. That is, subjects using the low quality representational data for the complex task apparently did not take additional time to translate the data in graph format into the precise numeric data it represents. Based on these results, it could be possible to infer that when a complex task was given with low quality representational data, problem solvers spent most time primarily on understanding and solving the decision task, rather than on translating the data presented in the graph into the precise numeric data.

On the other hand, the effect of representational DQ on problem-solving time was significant in the effect of the simple task. Subjects using the low quality representational data for the simple task took more time than subjects using the high quality representational data for the simple task. This is likely due to the fact that at a lower level of task complexity, subjects using the low quality representational data might take more effort (as measured by time) to translate the graph data into the precise numeric data it represents to generate good solutions, instead of taking effort to understand the task. In summary, it appears that the insignificant interaction effect between representational DQ and task complexity resulted from the insignificant mean difference between high and low representational DQ in the effect of the complex task, indicating these two variables do not jointly affect problem-solving time.

COMPLEXITY	REP. DQ	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Simple Task	High	10:22	03:24	08:25	12:20
	Low	16:39	06:54	14:42	18:37
Complex Task	High	27:50	07:33	25:53	29:47
	Low	30:50	06:52	28:53	32:48

Table 4. Table for Mean Values of Problem-Solving Time: Representational DQ by Task Complexity

For problem-solving accuracy, the interaction between representational DQ and task complexity was not significant (p = .408, see Table 3), indicating these two variables do not jointly affect problem-solving accuracy. This is likely due to the insignificant main effect of task complexity on problem-solving accuracy. There was a significant mean difference between high (72.708) and low (38.958) representational DQ in the effect of the simple task (see Table 5). There was also a significant mean difference between high (82.313) and low (39.188) representational DQ in the effect of the complex task. It appears that these significant differences resulted from the significant main effect of representational DQ on problem-solving accuracy.

However, the difference between the representational DQ effect in the simple task effect (e.g., 72.708 – 38.958 =33.75) and the representational DQ effect in the complex task effect (e.g., 82.313 – 39.188 = 43.125) was not significant (e.g., 33.75 – 43.125 = -9.375). It appears that this insignificant mean difference resulted from the insignificant main effect of task complexity. Thus, even though the main effect of representational DQ was significant on problem-solving accuracy, it was not significant enough to outweigh the insignificant main effect of task complexity on problem-solving accuracy, resulting in the insignificant interaction effect between representational DQ and task complexity.

COMP	REP. DQ	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Simple Task	High	72.708	5.630	61.495	83.921
	Low	38.958	5.630	27.745	50.171
Complex Task	High	82.313	5.630	71.099	93.526
	Low	39.188	5.630	27.974	50.401

Table 5. Table for Mean Values of Problem-Solving Accuracy: Representational DQ by Task Complexity

Hypotheses	Statistics		Evaluation
H1	F = 10.556	P = .002	Supported
H2	F = 46.611	P = .000	Supported
H3	F = 122.284	P = .000	Supported
H4	F = .763	P = .385	Rejected

Table 6. Summary of Hypotheses Testing

CONCLUSIONS AND DISCUSSION

When high representational DQ (e.g., tables) was used to solve the task used in this experiment (e.g., decision activities), problem-solving with cognitive fit resulted in increased problem-solving efficiency and effectiveness. However, when low representational DQ (e.g., graphs) was used for the task, a mismatch occurred between the problem representation and the task, which required subjects to transform the data values derived from the problem representation (e.g., graphs) into the mental representation suitable for task solution, which in turn had a negative impact on problem-solving performance. Therefore, it seems clear that the results of this study, especially the main effect of representational DQ on problem-solving accuracy and time, are compatible with the cognitive fit theory. Dickson et al. (1986) compared tables and bar charts for their effects on readability, interpretation accuracy, and decision-making effectiveness in a financial/accounting context. No differences in interpretation accuracy or decision quality were observed for the two groups. For their study, Dickson et al. picked simple levels of task content and task complexity. However, this study used two levels of task complexity and examined the interactions between two levels of representational DQ and two levels of task complexity. In addition, the task used in this study differs from the task Dickson et al. used. These experimental differences might lead to the different results, which were interpreted by different rationales, which may make some contributions to IS research, especially decision performance studies.

Based on a review of knowledge management literature, this research assumed that data is a prerequisite for information and information can be created from its raw data. DeLone and McLean (1992, 2003) postulated that system quality and information quality singularly and jointly affect both system use and user satisfaction that are direct antecedents of “individual impact.” Thus, based on DeLone and McLean’s model and the assumption mentioned above, this study predicted that improved data quality would positively affect information quality, which affects both system use and user satisfaction, which in turn have an impact on user performance. The results of this study showed that the effect of representational DQ influences problem-solving efficiency and effectiveness. Thus, the findings of this study are consistent with the IS success model. However, what is lacking is a detailed model for describing how data (quality) is transformed into information (quality), the strength of the relationship between data (quality) and information (quality), and the strength of the relationship between information (quality), once transformed, and user performance. One area for future research would be to develop a model examining the transformation of data (quality) into information (quality). Figure 1 presents a model for extending the IS success model by recognizing and including the representational aspect of DQ into the model.

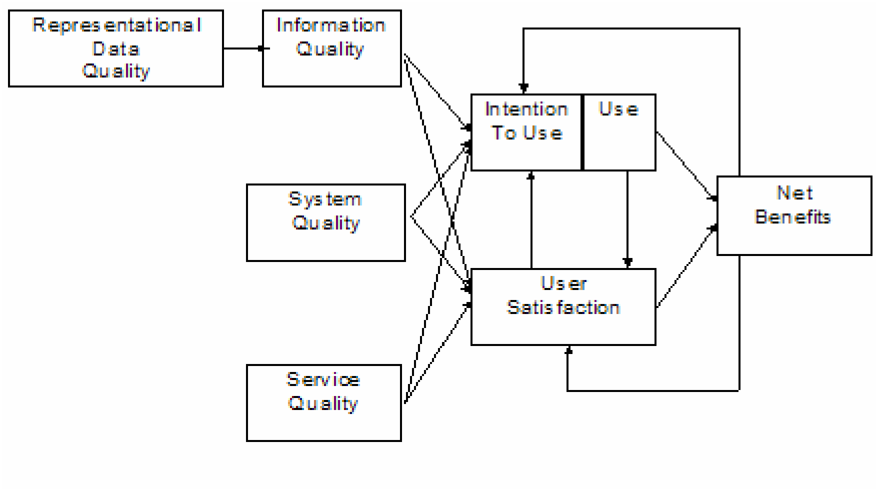


Figure 1. IS Success Model with Representational Data Quality

The observed main effect of representational DQ on problem-solving performance has practical implications for enhancing the efficiency and effectiveness of problem-solving. In order to improve users' ability to make decisions, systems designers and managers should not only make data available to users, but also enable users to access better (high-quality) data. To accomplish this, it is recommended that systems designers and managers support the task by providing users with high-quality representational data that matches the task. Various information technologies and systems are making possible the access of high-quality data across a firm. In recent years, data warehousing and the Internet are considered as the two key technologies that offer potential solutions for managing corporate data (Chen and Frolick, 2000). Large database management systems such as data warehouses have continued to be well ingrained into the business environment as one of the most important strategic initiatives in the information systems field (Watson, 2001) and a dedicated source of data to support decision-making applications (Gray and Watson, 1998). The Internet using extensible markup language (XML) makes it easy and less costly to access and present high-quality representational data from anywhere at anytime. Therefore, the integration of the two technologies (Web-based data warehousing) makes the processing, presenting, and distributing of high-quality representational data more efficient and economical.

In addition, the results of this study may help organizations to justify their attempt of improving data quality and/or their investments in a certain information technology. Based on a recent industry report, the economic and social damage from poor-quality data costs billions of dollars (Redman, 1998). Wang and Strong (1996) also implied that poor data quality can have substantial negative social and economic impacts. Since few empirical studies of the impact of data quality and task complexity on decision-making performance have been investigated in the IS literature, many organizations might have made investments in expensive data management projects without theoretical foundation. Since the findings of this study showed that data quality, especially representational DQ, brought problem-solving effectiveness and efficiency improvements, improving data quality by investing in data management practices would appear to be beneficial for organizations' performance. Thus, the findings of this study may provide evidence that helps organizations to justify their efforts to improve data quality.

A number of limitations should be considered in terms of the methods used when interpreting the findings. It is almost impossible to control the influence of all potential extraneous variances by the nature of the experimental setting, the subject population, the subjects' capability and characteristics, the decision support applications, the task type, and the set of data used in this study. Furthermore, since data were collected from a small sample of 40 students and the subjects were undergraduate students, the findings of this study might not generalize to a broader population. Because a single empirical study is not sufficient to validate the findings, further research should address these limitations and apply the findings of this study in specific contexts, population, and decision support technology as a whole.

REFERENCES

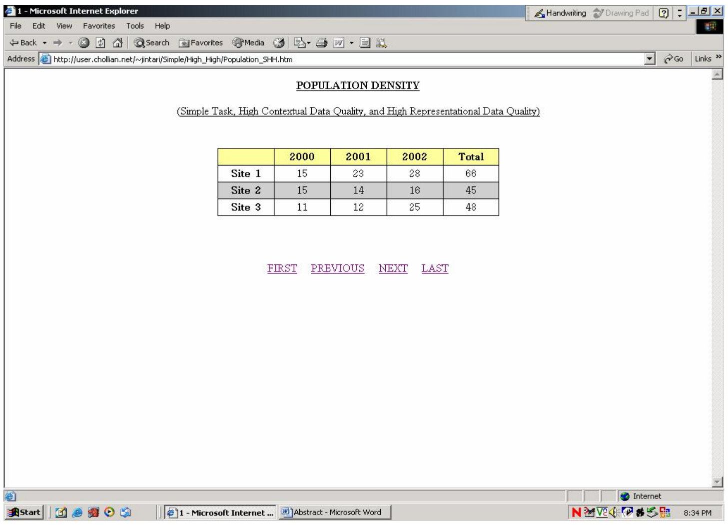
1. Ahituv, N., Igbaria, M., and Stella, A. (1998) The Effects of Time Pressure and Completeness of Information on Decision Making, *Journal of Management Information Systems*, 15,2, 153-172.
2. Austin, R.D. (2001) The Effects of Time Pressure on Quality in Software Development: An Agency Model, *Information Systems Research*, 12, 2, 195-207.
3. Benbasat, I. and Dexter, A.S. (1985) An Experimental Evaluation of Graphical and Color-Enhanced Information Presentation, *Management Science*, 31, 1348-1363.

4. Campbell, D.J. (1988) Task Complexity: A Review and Analysis, *Academy of Management Review*, 13, 1, 40-52.
5. Chandra, A. and Krovi, R. (1999) Representational Congruence and Information Retrieval: Towards an Extended Model of Cognitive Fit, *Decision Support Systems*, 25, 271-288.
6. Chen, L. and Frolick, M.N. (2000) Web-based Data Warehousing: Fundamentals, Challenges, and Solutions, Information Systems Management, Spring, 80- 86.
7. DeLone, W.H. and McLean, E.R. (1992) Information Systems Success: The Quest for The Dependent Variable, *Information Systems Research*, 3, 1, 60-95.
8. DeLone, W.H. and McLean, E.R. (2003) The DeLone and McLean Model of Information Systems Success: A Ten-Year Update, *Journal of Management Information Systems*, 19, 4, 9-30.
9. DeSanctis, G. (1984) Computer Graphics as Decision Aids: Direction for Research, *Decision Sciences*, 15, 4, 463-487.
10. Dickson, G.W., DeSanctis, G., and McBride, D.J. (1986) Understanding The Effectiveness of Computer Graphics for Decision Support: A Cumulative Experimental Approach, *Communications of the ACM*, 29, 1, 40-47.
11. Etezadi-Amoli, J. and Farhoomand, A.F. (1996) A Structural Model of End User Computing Satisfaction and User Performance, *Information and Management*, 30, 2, 65-73.
12. Gray, P. and Watson, H.J. (1998) Decision Support in the Data Warehouse, Upper Saddle River, New Jersey: Prentice Hall.
13. Iivari, J. and Koskela, E. (1987) The PIOCO Model for Information Systems Design, *MIS Quarterly*, 11, 3, 401-419.
14. Jarvenpaa, S.L. (2003) Additive-Difference Task: Ying-Yang Corporation Site Selection, Indiana University Kelley School of Business, (available online at http://kelley.iu.edu/bwheeler/ISWorld/index.cfm; accessed Nov. 1, 2003).
15. Javenpaa, S.L. and Dickson, G.W. (1988) Graphics and Managerial Decision Making: Research Based Guidelines, *Communications of the ACM*, 31, 6, 764-774.
16. Jarvenpaa, S.L., Dickson, G.W., and DeSanctis, G. (1985) Methodological Issues in Experimental IS Research: Experiences and Recommendations, *MIS Quarterly*, 9, 2, 141-156.
17. Newell, A. and Simon, H.A. (1972) Human Problem Solving, Englewood, Cliffs, NJ: Pretice-Hall.
18. Park, J. and Kim, J. (2000) Effects of Contextual Navigation Aids on Browsing Diverse Web Systems, *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, 1, 6, 257-264.
19. Payne, J.W. (1976) Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis, *Organizational Behavior and Human Performance*, 16, 366-387.
20. Redman, T.C. (1998) The Impact of Poor Data Quality on the Typical Enterprises, *Communications of the ACM*, 1998 February.
21. Rossano, M.J. and Moak, J. (1998) Spatial Representation from Computer Models: Cognitive Load, Orientation Specificity and the Acquisition of Survey Knowledge, *British Journal of Psychology*, 89, 3, 481-497.
22. Smelcer, J.B. and Carmel, E. (1997) The Effectiveness of Different Representations for Managerial Problem Solving: Comparing Tables and Maps. *Decision Sciences*, 28, 2, Spring, 391-419.
23. Tan, J.K.H. and Benbasat, I. (1990) Processing of Graphical Information: A Decomposition Taxonomy to Match Data Extraction Tasks and Graphical Representations, *Information Systems Research*, 1, 4, December, 416-439.
24. Teo, T.S.H. and Wong, P.K. (1998) An Empirical Study of the Performance Impact of Computerization in the Retail Industry, *Omega – The International Journal of Management Science*, 26, 5, 611-621.
25. Thuring, M., Hannemann, J., and Haake, J.M. (1995) Hypermedia and Cognition: Designing for Comprehension, *Communications of the ACM*, 38, 8, August, 57-74.
26. Tuomi, I. (1999) Data Is More Than Knowledge: Implications of the Reversed Knowledge Hierarchy for Knowledge Management and Organizational Memory, *Journal of Management Information Systems*, 16, 3, 103-117.
27. Vessey, I. (1991) Cognitive Fit: A Theory Based Analysis of The Graphs Versus Tables Literature, *Decision Sciences*, 22, 219-241.
28. Wang, R.Y. and Strong, D.M. (1996) Beyond Accuracy: What Data Quality Means to Data Consumers, *Journal of Management Information Systems*, 12, 4, 5-34.
29. Watson, H.J. (2001) Recent Developments in Data Warehousing," *Communications of the Association for Information Systems*, 8, 1-25.

30. Wixom, B.H. and Watson, H.J. (2001) An Empirical Investigation of the Factors Affecting Data Warehousing Success, *MIS Quarterly*, 25, 1, 17-41.

31. Wood, R. (1986) Task Complexity: Definition of the Construct, *Organizational Behavior and Human Decision Processes*, 37, 60-82.

APPENDIX A



POPULATION DENSITY

(Simple Task, High Contextual Data Quality, and High Representational Data Quality)

	2000	2001	2002	Total
Site 1	15	23	28	66
Site 2	15	14	16	45
Site 3	11	12	25	48

FIRST PREVIOUS NEXT LAST