2001

The Impact of Data Quality Tagging on Decision Outcomes

Graeme Shanks
The University of Melbourne, g.shanks@dis.unimelb.edu.au

The Impact of Data Quality Tagging on Decision Outcomes

Graeme Shanks

Department of Information Systems
The University of Melbourne, Melbourne, Australia
g.shanks@dis.unimelb.edu.au

Abstract

Data quality tags provide information about the quality of data in databases to decision makers. This paper reports an experiment that examines the impact of data quality tagging about data accuracy on decision outcomes. Two decision strategies were explored: additive and elimination by attributes. The inclusion of data quality tagging information was found to impact decision outcomes for the elimination by attributes strategy but not for the additive strategy and had no impact on group consensus. This knowledge will be valuable for designers of data warehouses and decision support systems.

Keywords

Data quality, data warehousing, decision support systems

INTRODUCTION

There is strong evidence that data quality problems are becoming increasingly prevalent in practice (Wand and Wang 1996), particularly in data warehousing, customer relationship management systems and the implementation of enterprise resource planning systems (Parr, Shanks and Darke 1999). This trend is important for these enterprise systems where data is obtained from multiple sources and used by people who are far removed from the original data collection and may have little understanding of the nuances regarding the meanings of data items (Tayi and Ballou 1998). Poor quality data in organisational information systems can have significant social and economic impacts (Strong et al. 1997). It is critical that organisations understand and manage data quality, and that procedures are in place to assure the quality of data.

Data in enterprise systems is often used by decision makers within their decision-making processes. Decisions must be made regardless of the quality of the data in databases, by either accepting the data as is or by avoiding using data that is suspected to be of poor quality. This paper examines the proposition that knowledge about the quality of the data may help decision processes and improve decision outcomes (Chengular-Smith et al. 1999).

Data quality tagging involves the storage of data about data quality within an organisation's databases. These data quality tags are then made available to decision makers when they use the data. Very little research has been conducted into the effectiveness of data quality tagging. Chengular-Smith et al. (1999) found that under some circumstances, for particular decision-making strategies, data quality tagging impacts decision outcomes. This paper builds on Chengular-Smith et al.'s work and examines the impact of data quality tagging on decision outcomes for different decision-making strategies using a database query facility. This work is significant for practitioners as determining and storing data quality tags is an expensive process and the impact of data quality tagging on decision outcomes must be clearly understood before any investment can be justified.

The paper is structured as follows. The next section of the paper discusses data quality and reviews previous research in data quality. The third section discusses decision-making strategies and explains why the additive and elimination by attributes models were selected for this study. The fourth section describes the research design for this study and discusses the research procedure. The following section discusses the results of the study, presents a number of implications for practitioners and researchers from the study and suggests areas for further research.

DATA QUALITY

It is important to consider both the intrinsic characteristics of data itself and the assessments of users of the data when defining data quality (Strong et al. 1997). This means that data must be usable and useful to the users of the data and support their effective work practices. I therefore adopt the definition of quality as "fitness for purpose".

Much of the previous research in data quality uses lists of desirable data quality dimensions as a means of thinking about data quality. Early research in data quality focused on the accuracy of hard accounting data in financial systems but this has been extended to include other data quality dimensions for both hard and soft data (Chengular-Smith et al., 1999). The dimensions used to understand data quality typically include accuracy, reliability, importance, consistency, precision, timeliness, understandability, conciseness and usefulness (Ballou and Pazer, 1985; Wand and Wang, 1996). Although these dimensions provide a useful means of thinking about data quality, they are often overlapping, vaguely defined, ambiguous and not soundly based in theory.

A number of frameworks have been proposed that organise and structure important concepts in data quality, and provide a more complete perspective on data quality. Wang and Strong (1996) used a survey of the expert opinions of practitioners to identify and cluster data quality dimensions into four categories: intrinsic, contextual, representational and accessibility. Other frameworks are theoretically based: the framework of Kahn et al. (1997) is based on product and service quality theory; the framework of Wand and Wang (1996) is based on Bunge's ontology; and the framework of Shanks and Darke (1998) is based on semiotic theory and Bunge's ontology. In this paper I use the framework of Shanks and Darke as it is soundly based in theory and includes both intrinsic and extrinsic aspects of data quality.

A Framework for Understanding Data Quality

The framework of Shanks and Darke consists of a number of components: data quality goals; data quality dimensions for each of the goals; the stakeholders responsible for producing, maintaining, and consuming data; measures for each of the data quality dimensions; and improvement strategies. In this study I use data quality goals and measures for the data consumer stakeholder.

The four semiotic levels defined by Stamper (1992) are used to define the data quality goals. These are the syntactic, semantic, pragmatic, and social levels. Although these levels are separated for analytical convenience, they are closely interrelated and build on each other. The four levels are briefly discussed below and data quality goals are established for each. Data quality measures or relevant goals are discussed later.

Syntactic data quality concerns the form of data rather than its meaning. The goal for syntactic data quality is consistency where data values for particular data elements in the data warehouse use a consistent symbolic representation (Ballou et al. 1996, Wang et al. 1995).

Semantic data quality concerns the meaning of data. The goals for semantic quality are completeness, accuracy and currency (Tayi and Ballou 1998, Wang et al. 1995). Completeness is concerned with the extent to which there is a one-to-one correspondence between data and things in the real world system. Accuracy is concerned with how well data represent states of the real world. Currency is concerned with how up-to-date the data is (this is different to timeliness which depends on how the data is being used).

Pragmatic data quality concerns the usage of data. The goals for pragmatic quality are usability and usefulness (Kahn et al., 1997). Usability is the degree to which each stakeholder is able to effectively access and use the symbols. Usefulness is the degree to which stakeholders are supported by the symbols in accomplishing their tasks within the social context of the organisation. Desirable characteristics relating to pragmatic data quality include timeliness, understandability, conciseness, accessibility and reputation of the data source.

Social data quality concerns the shared understanding of the meaning of symbols. The goals for social data quality are an understanding of different stakeholder viewpoints and an awareness of any biases and other cultural and political issues involved (Shanks and Corbitt, 1999).

Data Quality Metadata

Data quality metadata is data about the quality of data within an organisation's databases. Data quality tagging is the process of measuring an aspect of data quality and storing it as metadata. Since multiple stakeholders share data in databases for many purposes, both the pragmatic and social semiotic levels of data quality are unsuitable candidates for data quality tagging. Data quality goals at these levels depend on the stakeholder, the task at hand and the context of the task.

Data quality goals at the syntactic and semantic semiotic levels however are universal and apply to all stakeholders regardless of the task at hand and the context of the task. Potential data quality metadata therefore concerns the consistency, completeness, currency and accuracy of the data. Each of these goals for data quality is of significant concern in practice. For simplicity, in this study I focus on one of these goals, data accuracy. The other data quality goals will be the focus of ongoing research.

Typical consistency problems involve differences in coding schemes and data types with multiple data sources. The data set used in this experiment has been designed to have perfect syntactic data quality. I also assume that the data set is complete, that is, there is a record in the data set for each thing in the real world system, and there is a value for each attribute of each record. In addition, I am not including currency in this study, and each record is assumed to be up-to-date.

Data accuracy, the extent to which data represents states of the real world. Although the value of each attribute may exist for each record in a data set, that is the data set is complete, it may not be accurate. For a data set to be accurate each data value must correspond precisely to the state in the real world of the thing that is represented. For example, an employee record may be complete in that each attribute has a value, but if, say, the salary value does not match the actual salary value then the record is inaccurate.

An important issue with data quality metadata is granularity, or the level of detail at which the data quality tags apply. Tags could apply at several levels of detail including the individual data item level, the data field (or attribute) level or the level of a file or relational database table. I adopt the approach of Chengular-Smith et al. (1999) and use the data field level. This level is a compromise between the expensive and complex option of individual data item with the simple and less useful file level. It could also be supported readily in data dictionaries by enhancing metadata information about data items.

A further issue is the representation of the data quality tags. The way the data quality tags are represented can affect decision behaviour and should be designed to promote effective decision-making. Data quality tags represent the measures of data quality goals and may be determined and represented in many different ways. The determination and representation of data quality tags is a complex issue beyond the scope of the present study. I follow the approach adopted by (Chengular-Smith et al. 1999) and consider interval and ordinal scales. Interval scale representation could be on a scale of 0 to 100, so that data scored as 80 will be more accurate than data scored as 60. Ordinal scales can use categories such as very good, good, average, poor etc. In this study I use interval scales as they provide more detailed information and are more fully utilized in simple decision tasks (Chengular-Smith et al. 1999). The 0 to 100 scale was selected as it is expressed as a percentage and is easy to use, however it does give a false impression of the precision of the accuracy measures.

Decision-making Strategies

This research focuses on decision-making as the process of choosing among multiple alternatives described by the same set of attributes. In addition the strategies adopted are built into the spreadsheet-like interface of a database system used in the decision-making task. There has been much research about decision-making strategies over several decades. I use the perspective of Payne et al. (1993) as it is highly normative and is well suited to the design of this study. Payne et al. (1993) suggest that a decision-making strategy is adopted on the basis of a cost-benefit analysis, based on making the best possible decision while minimizing the cognitive effort required in making the decision. The impact of data quality tagging on decision outcomes may depend on the decision-making strategy adopted.

Payne (1976) presents the four most important decision strategies: additive, conjunctive, additive difference and elimination by attributes. The additive strategy involves evaluating each alternative separately by assigning a value to each attribute and adding them to give an overall value for that alternative. The alternative with the highest overall value is chosen. The conjunctive strategy uses the principal of satisficing (Simon, 1957) to reduce cognitive effort and involves searching alternatives until an alternative is found with the value of each attribute exceeding some minimum standard value. The additive difference strategy involves comparing alternatives directly on each attribute and then adding the differences to reach a decision. The elimination by attributes strategy involves comparing alternatives by first selecting one attribute and then eliminating all alternatives that do not have the required value of that attribute. The process is repeated until only one alternative remains.

Decision-making strategies may be categorized as either alternative-based or attribute-based and compensatory or non-compensatory (Payne, 1993). In alternative-based approaches multiple attributes of a single alternative are considered before other alternatives are processed. In contrast, in attribute-based processing the values of several alternatives on a single attribute are processed before other attributes are processed. In compensatory approaches trade-offs are made between attributes and a good value of one attribute can compensate for bad value in other attributes. In contrast, in non-compensatory approaches a bad value on an important attribute will ensure that an alternative would never be chosen.

In this study, the additive (A) and elimination by attributes (EBA) decision-making strategies have been used. The A strategy is an alternative-based and compensatory strategy whereas the EBA is attribute-based and non-compensatory. These two decision-making strategies therefore have contrasting properties and provide a useful comparison.

RESEARCH DESIGN

An experiment is used to examine the impact of data quality tagging on decision outcomes for different decision-making strategies using a database query facility. The research model is shown in Figure 1. The independent variables are decision strategy (additive and elimination by attributes) and data quality tagging (accuracy and none). The dependent variables are the decision outcome (complacency and consensus), the time taken for the decision and confidence in the decision outcome.

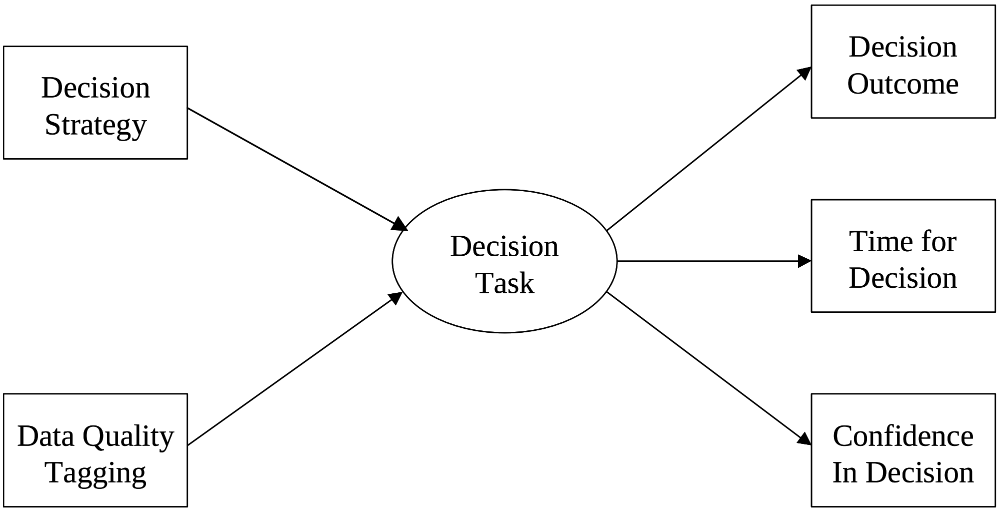


Figure 1: Research Model

The Decision Task

The decision task was selected in order for the results to be valid and generalisable and for comparison with the previous study of Chengular-Smith et al. (1999). The task was adapted from the apartment selection task used by Payne (1976) and then adapted by Chengular-Smith et al. (1999). The task involved selecting an apartment from a number of alternatives based on a number of attributes, including rent, commuting time, parking space and number of rooms.

The tasks used in previous studies used a small number of alternatives, in all cases less than ten. As one of the motivations for this study was the increasing usage of data from data warehouses by decision makers, the number of alternatives was increased greatly and stored in a computer database. Access to the data was provided by a spreadsheet-like interface.

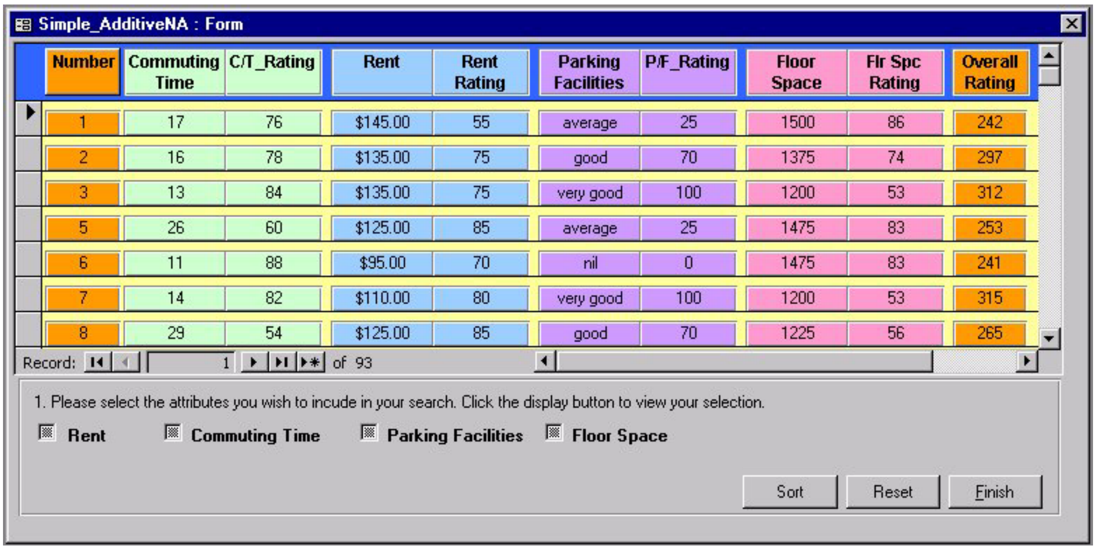
Independent Variables

There are two independent variables in the study: decision strategy and data quality tagging.

Decision Strategy

Two decision strategies are used, additive (A) and elimination by attribute (EBA). The A strategy is supported by a spreadsheet-like interface that displays a scrolling window of alternatives with the values of all attributes displayed for both the simple and complex decision tasks. Decision makers are able to select those attributes they wish to include in the decision-making process. The interface automatically calculates the sum of the individual attribute values and sorts the alternatives into descending order. (Note that each attribute is shown in both a realistic unit of measure and also as a percentage rating: this facilitates the summation and sorting of alternatives) Decision makers can reset the display and select a different combination of attributes if they wish. The interface for the A decision strategy is shown in Figure 2.

The EBA strategy is also supported by a spreadsheet-like interface that displays a scrolling window of alternatives with the values of all attributes displayed for both the simple and complex decision tasks. Decision makers are able to select those attributes they wish to include in the decision-making process. The interface automatically sorts the alternatives in the order of the selected attributes. Decision makers can reset the display and select a different combination and sequence of attributes if they wish. The interface for the EBA decision strategy is shown in Figure 3.



Simple_AdditiveNA : Form

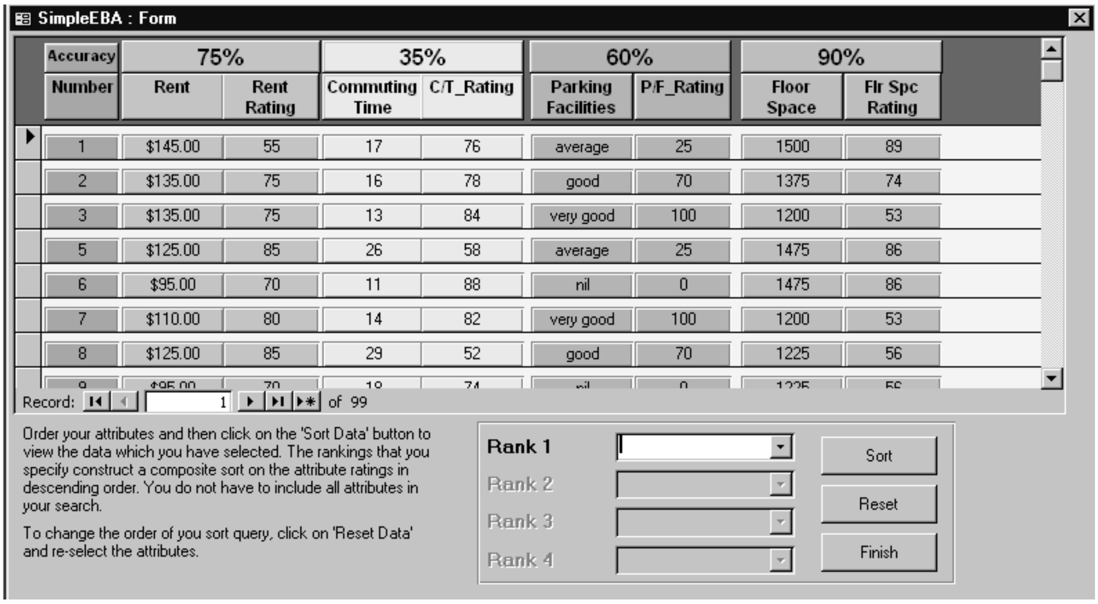
Number	Commuting Time	C/T_Rating	Rent	Rent Rating	Parking Facilities	P/F_Rating	Floor Space	Flr Spc Rating	Overall Rating
1	17	76	$145.00	55	average	25	1500	86	242
2	16	78	$135.00	75	good	70	1375	74	297
3	13	84	$135.00	75	very good	100	1200	53	312
5	26	60	$125.00	85	average	25	1475	83	253
6	11	88	$95.00	70	nil	0	1475	83	241
7	14	82	$110.00	80	very good	100	1200	53	315
8	29	54	$125.00	85	good	70	1225	56	265

Record: 1 of 93

1. Please select the attributes you wish to incude in your search. Click the display button to view your selection.

▨ Rent ▨ Commuting Time ▨ Parking Facilities ▨ Floor Space

Figure 2: Interface for Additive Decision Strategy



SimpleEBA : Form

Accuracy	75%		35%		60%		90%	
Number	Rent	Rent Rating	Commuting Time	C/T_Rating	Parking Facilities	P/F_Rating	Floor Space	Flr Spc Rating
1	$145.00	55	17	76	average	25	1500	89
2	$135.00	75	16	78	good	70	1375	74
3	$135.00	75	13	84	very good	100	1200	53
5	$125.00	85	26	58	average	25	1475	86
6	$95.00	70	11	88	nil	0	1475	86
7	$110.00	80	14	82	very good	100	1200	53
8	$125.00	85	29	52	good	70	1225	56
9	$95.00	70	18	74	nil	0	1225	56

Record: 1 of 99

Order your attributes and then click on the 'Sort Data' button to view the data which you have selected. The rankings that you specify construct a composite sort on the attribute ratings in descending order. You do not have to include all attributes in your search.

To change the order of you sort query, click on 'Reset Data' and re-select the attributes.

Figure 3: Interface for Elimination by Attributes Decision Strategy (including accuracy data quality tagging information)

Data Quality Tagging

Two values were used for data quality tagging: an interval scale from 0 to 100 and no data quality information. The scale does not represent an absolute measure of data accuracy but is used to indicate the relative quality of each attribute in relation to other attributes. Data quality tagging may be seen in Figure 3 above, where each data attribute is allocated a data accuracy value. It can be seen that the commuting time attribute has particularly poor data accuracy.

Dependent Variables

There are three dependent variables in the study: decision outcome, time for decision and confidence in decision.

Decision Outcome

The decision task was designed so that without data quality information there was a preferred solution and therefore an objective evaluation measure. The inclusion of data quality information made the selection of an alternative a subjective assessment as there was no longer a preferred solution. Two aspects of the decision outcome were explored: *decision complacency* and *decision consensus* (adapted from Chengular-Smith et al. 1999). Decision complacency refers to the degree to which decision makers ignore data quality information.

Decision consensus refers to the degree to which decision makers converge on a decision when presented with data quality information.

Time Taken for Decision

The inclusion of data quality tagging information may affect the time taken by each decision maker to reach a decision. Therefore the time taken by decision makers to reach a decision was recorded.

Confidence in decision

The inclusion of data quality tagging information may affect the confidence the decision maker has in her decisions. A five point likert scale was used to record the confidence of decision makers in their decision outcomes.

The Experimental Procedure

The participants used in the experiment were undergraduate students in the Department of Information Systems at the University of Melbourne. They were randomly assigned to four groups. Each group was assigned to the task using one of the decision-making strategies, with and without data quality information.

A database system with a spreadsheet-like interface was developed for each of the four groups. Each database system was loaded with 100 alternative apartments with one of the apartments being the preferred solution (without data quality information). A set of instructions and an answer sheet were also developed for each of the groups. The task involved selecting the preferred apartment, recording the start and finish times, nominating a confidence level of the decision, and providing a brief explanation of the reason for their decision. All experimental materials were trialed with several undergraduate students and minor modifications were made to the interface and the explanation document.

The experimental procedure involved a brief explanation of the particular decision-making strategy that was to be followed in the decision task, and then a demonstration of how to use the computer system. All participants had previously used similar types of system and had no difficulty using the interface.

Hypotheses and Data Analysis

The experiment was designed to explore the impact of data quality tagging information on decision outcomes. A number of hypotheses were developed based on decision complacency, decision consensus, time for decision-making and confidence in decision outcome. For brevity of presentation, only alternative hypotheses are shown.

It is expected that information about data quality will affect decision outcomes. If decision-makers ignore data quality information then the preferred apartment should not change. Hypotheses concerning *decision complacency* are:

H1A Including data quality information changes the number of times the originally preferred apartment continues to be the preferred apartment for the A strategy.

H1B Including data quality information changes the number of times the originally preferred apartment continues to be the preferred apartment for the EBA strategy.

It is expected that information about data quality will not prevent decision makers from reaching consensus about decisions. If data quality does not prevent consensus then the selected apartment should not change. Note that the selected apartment does not necessarily need to be the preferred apartment. Hypotheses concerning *decision consensus* are:

H2A Including data quality information changes the number of times a selected apartment continues to be a selected apartment for the A strategy.

H2B Including data quality information changes the number of times the selected apartment continues to be a selected apartment for the EBA strategy.

It is expected that information about data quality will increase the time decision makers take to make decisions. Hypotheses concerning *time to make decisions* are:

H3A Including data quality information increases the time taken to make a decision for the A strategy.

H3B Including data quality information increases the time taken to make a decision for the EBA strategy.

It is expected that information about data quality will increase the confidence decision makers have in the decisions they make. Hypotheses concerning *confidence in decisions taken* are:

H4A Including data quality information increases the confidence in decisions for the A strategy.

H4B Including data quality information increases the confidence in decisions for the EBA strategy.

RESULTS AND DISCUSSION

Hypothesis Testing

Results for each of the hypotheses tested are summarised in Tables 1 and 2.

	Decision Strategy	
	Additive	Elimination by Attributes
Outcome Complacency	χ^2 = 1.364	χ^2 = 13.132
	p = 0.077 (H1A)	**p = 0.000* (H1B)**
Outcome Consensus	χ^2 = 1.364	χ^2 = 4.286
	p = 0.077 (H2A)	p = 0.609 (H2B)

Table 1: Summary of Hypothesis Testing Results

	Decision Strategy			
	Additive		Elimination by Attributes	
	No DQ tag (mean/stdev)	DQ tag (mean/stdev)	No DQ tag (mean/stdev)	DQ tag (mean/stdev)
Time (minutes)	6.133 (.764)	9.133 (.764)	7.000 (.764)	6.733 (.764)
	p = 0.007* (H3A)		p = 0.806 (H3B)	
Confidence (5 point likert scale)	3.600 (.183)	4.067 (.183)	3.800 (0.183)	3.667 (0.183)
	p = 0.077 (H4A)		p = 0.609 (H4B)	

Table 2: Summary of Hypothesis Testing Results

Decision *complacency* refers to the degree to which decision makers ignore data quality information. Table 1 shows a significant difference for the EBA decision-making strategy (p=0.000) but not for the A strategy. Comments made by participants in explaining their decisions indicated that many chose to ignore the commuting time attribute due to its poor quality. In the EBA strategy, the sequence of attributes considered in the decision is highly significant for the decision outcomes. If the attribute commuting time is not included in the attributes selected for query due to its poor quality, the decision is almost certain to change. In contrast, the compensatory nature of the A strategy means that the impact of one poor quality attribute is reduced by the inclusion of other attributes in the overall ranking of apartments. This result differs from the study of Chengular-Smith et al. (1999). The difference may be explained by the use in this study of a database of 100 alternative apartments and a decision strategy built into a query interface. In the database environment, decision strategy becomes a more dominant factor in the decision process and outcome.

Decision *consensus* refers to the degree to which decision makers converge on a decision when presented with data quality information (although not necessarily the preferred apartment). Table 1 shows that a consensus was reached in the decision outcome for both the A and EBA decision strategies. Therefore, although the inclusion of data quality information may impact decision outcomes, it does not impair the ability of a group to reach consensus on a decision, and indicates the individuals process data quality information in similar ways.

The *time* taken to make a decision was significantly increased when data quality information was used in the A decision strategy but not for the EBA strategy. Comments from participants indicated that the decision to ignore the poor quality attribute in the sort sequence for retrieval was made easily and quickly. However, in the A strategy the decision is less obvious as other attributes can compensate for the attribute with poor data quality. The decision process therefore becomes more complex.

The *confidence* in the decision outcome has not changed significantly for either decision strategy when data quality information is included. It should be noted that the average confidence level was high for each of the four groups involved in the study. Therefore, although the inclusion of data quality information may impact decision outcomes, it does not effect the confidence decision makers have in their decisions.

Conclusions

This study has drawn together concepts from a semiotic-based theory on data quality and normative theories on decision-making to examine the impact of data quality tagging about data accuracy on decision outcomes. The inclusion of data quality tags in databases has been shown to impact decision outcomes for the elimination by attributes decision strategy but not for the additive decision strategy. At the same time consensus on decision outcomes and confidence in the decision outcomes were not adversely affected.

The results of this study have important implications for designers of data warehouses and decision support systems. Practitioners should carefully consider if the additional cost of determining and storing data quality tags is justified by the potential improvements in decision-making process and outcomes in their particular organisational context. Even when data quality tags are included in databases, it is clear that they are only useful when certain decision-making strategies are included in query facilities.

Limitations

The main limitations of this study are the use of senior undergraduate students as participants and the conduct of the study in a laboratory setting. Although senior undergraduate students are familiar with the decision task selected for the study, experienced practitioners would be more accustomed to the complexities of real world decision-making. Although the use of a laboratory setting increased the internal validity of the experiment, the decision process was consequently highly structured and not impacted by contextual issues.

Further Research

Further studies are planned involving more complex decision tasks, other kinds of data quality including consistency, currency and completeness, and using experienced decision-makers as participants. In addition, studies of decision-making process as well as outcomes are planned.

REFERENCES

Ballou, D.P. and Pazer, H.L. (1985) Modeling data and process quality multi-input multi-output information systems, *Management Science*, 31:2, 150-162.

Chengular-Smith, I.N., Ballou, D. and Pazer, H.L. (1999) The Impact of Data Quality Information on Decision Making: An Exploratory Analysis, *IEEE Transactions on Knowledge and Data Engineering*, 11:6, (November/December).

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

Kahn, B., Strong, D.M. and Wang, R.Y. (1997) A Model for Delivering Quality Information as Product and Service, *Proceedings of the 1997 Conference on Information Quality*, Boston: MIT, 80-94.

Parr, A., Shanks, G. and Darke, D. (1999) *"Identification of necessary factors for successful implementation of ERP systems"*, in O. Ngwenyama, L.D. Introna, M.D. Myers and J.I. DeCross (eds.) New Information Technologies in Organisational Processes, Boston: Kluwer Academic Publishers, pp. 99-119.

Payne, J.W. (1976) Task Complexity and Contigent Processing in Decision Making: An Information Search and Protocal Analysis, *Organisational Behaviour and Human Performance*, Vol. 16, 366-387.

Payne, J.W., Bettman, J.R. and Johnson, E.J. (1993) *The Adaptive Decision Maker*, Cambridge, Cambridge University Press.

Shanks, G. and Darke, P. (1998) Understanding Metadata and Data Quality in a Data Warehouse, *Australian Computer Journal* (November).

Shanks, G. and Corbitt, B. (1999) Understanding Data Quality: Social and Cultural Aspects, *Proc. Australasian Conference on Information Systems*, Wellington (December).

Simon, H (1957) *Models of Man*, Wiley, New York.

Stamper, R. (1992) Signs, Organisations, Norms and Information Systems, *Proc. 3rd Australian Conference on Information Systems*, Wollongong.

Strong, D.M., Lee, Y.W. and Wang, R.Y. (1997) Data Quality in Context, *Communications of the ACM*, Vol 40, No 5, 103-110.

Tayi and Ballou (1998) Examining Data Quality, *Communications of the ACM*, Vol 41. No 2, 54-57.

Wand, Y. and Wang, R. (1996) Anchoring Data Quality Dimensions in Ontological Foundations, *Communications of the ACM*, Vol 39, No 11, 86-95.

Wang, Y., Reddy, M.P. and Kon, H.B. (1995) Toward Data Quality: An Attribute-based Approach, *Decision Support Systems*, Vol 13(3,4), 349-372.

Weber, R. (1997) *Ontological Foundations of Information Systems*, Coopers and Lybrand, Melbourne.

COPYRIGHT