Association for Information Systems
AIS Electronic Library (AISeL)

PACIS 2003 Proceedings

Pacific Asia Conference on Information Systems (PACIS)

December 2003

The Categorical Information Quality Framework (CIQF): A Critical Assessment and Replication Study

Michael Gendron
University of Melbourne

Graeme Shanks
University of Melbourne

Follow this and additional works at: http://aisel.aisnet.org/pacis2003

The Categorical Information Quality Framework (CIQF): A Critical Assessment and Replication Study

Michael Gendron* and Graeme Shanks**

** Department of Information Systems
University of Melbourne

Abstract

The CIQF has been used widely by researchers and practitioners for several years. In this paper we have identified a number of methodological issues in the process used to create the framework, in particular the predominant use of graduate students in data collection, the sample size for the statistical analysis techniques used, researcher bias in the four groupings of information quality dimensions initially proposed, the use of descriptive statistics to test the proposed categories and the domain specificity of the framework. We propose a replication study that addresses these issues, using practitioners as subjects, a larger sample size to ensure a stable factor structure, factor analysis techniques to determine both the information quality dimension and the categories within the framework. The framework developed in our replication study will provide researchers with an empirically based and fully validated framework for future information quality research and practitioners with a sound basis for developing information quality metrics and information quality improvement programs.

Keywords

Data Quality, Information Quality, Factor Analysis

Introduction

Information quality problems are widespread in practice and have significant economic and social impacts. (English, 1999; Redman, 2001; Wand & Wang, 1996) Poor quality information leads to customer dissatisfaction, increased operational costs, less effective decision-making   and lowered employee satisfaction. These problems are particularly important for organisations implementing enterprise-wide, integrated information systems, and electronic commerce systems that involve sharing of information within and between organisations. These systems rely heavily on high quality information for their success. (English, 1999; Wang, Lee, Pipino, & Strong, 1998)

Organizations are increasingly treating information as an asset which must be managed and quality information must be viewed as a strategic goal. Managers must determine the business merits of maintaining quality information, decide if it supports management of the firm, and ultimately determine how information quality affects the bottom line. (Huang, Lee, & Wang, 1999) A number of different approaches to understanding, assessing and improving

information quality have been proposed. (Agmon & Ahituv, 1987; English, 1999; King & Epstien, 1983; Redman, 2001; Wang & Strong, 1996) In these approaches, information quality is understood as a multi-dimensional construct with various levels of stakeholder focus and specificity. In this paper the Wang and Strong (Wang & Strong, 1996) information quality framework is focused on. This framework focuses on information quality as perceived by the data consumer and has received much attention by researchers (D'Onofrio & Gendron, 2001; Gendron & D'Onofrio, 2000; Gendron & D'Onofrio, 2002; Haung, Lee, & Wang, 1999; Kahn, 1997; Kovac, 1997). Since the framework consists of four categories of information quality dimensions, we refer to it as the categorical information quality framework (CIQF). We present a critical analysis of the CIQF, identify a number of methodological issues arising from its development and describe a replication of the study that we are currently undertaking. Our goals in this replication are:

- testing the construct validity of the CIQF and thus validating the categories and dimensions to ensure they operationalize the construct *information quality*;

- examining the domain specificity of information quality.

The paper is structured as follows. First, information quality is defined and a critical review of existing work in information quality is given. The following section discusses the CIQF and the method by which it was developed. The next section presents a critical analysis of the CIQF and identifies a number of potential issues with the framework. The following section describes our replication study and explains how the replication study will address the issues identified with the CIQF. The paper concludes with a discussion of the impact of our work for practitioners and researchers.

Information Quality

One definition of quality is "fitness for purpose" and therefore information quality includes not only the intrinsic quality of the information itself but also relates to how the information will be used by stakeholders for various purposes in different contexts. Data has been defined as a collection of symbols that are brought together because they are considered relevant to some purposeful activity (Mingers, 1995). In contrast, information is data that is structured and organised so that it has meaning to stakeholders. As the CIQF focuses on the perceptions of data consumers, we use the term information quality rather than data quality.

The existing work on information quality may be categorised into three types: expert opinion, theoretical studies and empirical studies. Much of the work on information quality is expert opinion, and is concerned with the intrinsic quality of data in databases. It largely consists of lists of desirable information quality dimensions (Wand & Wang, 1996). These lists typically include dimensions such as completeness, accuracy, reliability, consistency, timeliness, precision and conciseness. Over 150 of these information quality dimensions have been identified in the literature (Wand & Wang, 1996). However, these dimensions are often overlapping, vaguely defined, ambiguous and not soundly based in theory (Shanks & Darke, 1998).

Theoretical work on information quality has been based in ontology (Wand & Wang, 1996), semiotics (Shanks & Darke, 1998) and service quality theory (Kahn, 1997). The framework which anchors information quality dimensions in the ontological foundations (Wand &

Wang, 1996) defined by Bunge (Weber, 1997) is limited to intrinsic information quality. The semiotic framework (Shanks & Darke, 1998) is based on the four levels defined by Stamper (Stamper, 1992) is comprehensive including both intrinsic and extrinsic information quality dimensions. It has not been used in practice yet. The service quality framework (Kahn, 1997) is based on concepts from total quality management and service quality theory that categorised information quality dimensions as either sound, useful, usable or effective.

The third type of approach to information quality is based on empirical studies. The CIQF is the best-known example of this type of study (Wang & Strong, 1996). An extensive survey of data consumers was used to identify information quality dimensions and then organised them into four categories. It is this study that we focus on in this paper. A detailed description of the framework and the method by which it was developed follows in the next section.

The CIQF

Wang and Strong created the CIQF using multiple data collection and analysis methods within a five-step process. Each of these five steps is described below.

Attribute List Generation (Step-1)

The first step in the creation of the CIQF was to generate a list of information quality attributes as potential candidates for further study. Two subject pools were drawn; 25 data consumers working in industry and 112 graduate students from an MBA program at a large north-eastern U.S. university. Subjects were asked to do two things: 1) generate a list of attributes that first came to mind when they thought of information quality (from the perspective of data consumers), and 2) given a list of cues (32 potential information quality attributes), add to the list.

The industry subjects self-administered the two surveys and were also interviewed, while the graduate students only self-administered the two surveys. This process resulted in 179 information quality attributes.

Attribute Rating (Step-2)

Next, a survey was created using the 179 attributes as items. This survey was pre-tested, during which many items were eliminated leaving 118. 1500 subjects were randomly drawn from over 3200 alumni at a north-eastern U.S. University. The effective response rate was 25%. Using the 118 attributes as items on the survey, subjects were asked to rate them for importance from their own experience as data consumers. A scale from 1 to 9 was used where 1 was anchored as *extremely important* and 9 was anchored as *not important*.

The results of the attribute rating survey were used in an exploratory factor analysis yielding 20 factors (dimensions) of information quality. The response-to-variable ratio was 3, which is less than the recommended ratio of 5. As a result the factor structure may be unstable. This adds additional support for the proposed replication study. (Comrey, 1978; Streiner, 1994; Wang & Strong, 1996)

Preliminary Conceptual Framework (Step-3)

The researchers next created a preliminary conceptual framework, which consisted of groupings of the 20 dimensions. They created 4 groupings of dimensions based on their understanding of information quality which they felt appropriately captured similarities between the dimensions. The original framework is shown in Table 1.

First Sorting Study (Step-4)

The first sorting study was conducted to test the preliminary conceptual framework. A pool of 30 subjects from evening MBA classes at another north eastern U.S. university were selected to participate in the sorting first (Step-4) and second (Step-5) sorting studies. 18 of these subjects were randomly selected to participate in the first sorting study (Step-4). Subjects were given 20 3x5 index cards with the dimensions printed on them. They were asked to sort the cards into 3, 4, or 5 piles and to label each pile.

A 70% hit ratio between subject sorting and the preliminary conceptual framework was attained. Adjustments to the preliminary conceptual framework were made based on this sorting study: completeness was moved from grouping one to grouping two, and traceability, variety of data sources, ease of operation, flexibility, and cost effectiveness were eliminated. The adjustments can be seen in Table 1.

Table 1 - CIQF Groupings and Factors

Categories	Preliminary Conceptual Framework	Step-4 Adjustments
1) Accuracy of Data	Believability	Believability
	Accuracy	Accuracy
	Objectivity	Objectivity
	Completeness	*Moved to grouping 2*
	Traceability	
	Reputation	Reputation
	Variety of data sources	
2) Relevancy of Data	Value-added	Value-added
	Relevancy	Relevancy
	Timeliness	Timeliness
	Ease of operation	
	Appropriate amount of data	Appropriate amount of data
	Flexibility	
		Completeness
3) Representation of data	Interpretability	Interpretability
	Ease of understanding	Ease of understanding
	Representational consistency	Representational consistency
	Concise representation	Concise representation
4) Accessibility of data	Accessibility	Accessibility
	Cost effectiveness	
	Access security	Access security

Second Sorting Study (Step-5)

A second sorting study was conducted to confirm the adjusted conceptual framework. The remaining 12 subjects from the sorting study subject pool participated in this step. Subjects were given 4 short sentences that describe the groupings in the conceptual framework that can

be seen in Table 2. Subjects were given 15 3x5 index cards with the 15 dimensions remaining after Step-4 and were asked to place the 15 cards into the category that best represented the dimension. An 81% hit ratio between subject sorting and the adjusted conceptual framework was attained.

Table 2 –Descriptions

Grouping 1	The extent to which data values are in conformance with the actual or true values.
Grouping 2	The extent to which data are applicable to, or pertain to, the task of the data user.
Grouping 3	The extent to which data are presented in an intelligible and clear manner.
Grouping 4	The extent to which data are available or obtainable.

The Final CIQF

The final CIQF consists of 4 categories and 15 dimensions of information quality which can be seen in Table 3. The dimensions were reviewed by the researchers and a more descriptive label was created for each of the four categories. The process used to develop the CIQF is critically assessed the in next section of this paper.

Table 3 – Final Categories and Dimensions

Target Category	Final Categorical Labels	Dimension
Accuracy	**Intrinsic Data Quality**	Believability
		Accuracy
		Objectivity
		Reputation
Relevancy	**Contextual Data Quality**	Value-added
		Relevancy
		Timeliness
		Completeness
		Appropriate amount of data
Representation	**Representational Data Quality**	Interpretability
		Ease of understanding
		Representational consistency
		Concise representation
Accessibility	**Accessibility Data Quality**	Accessibility
		Access security

A Critical Assessment of the CIQF

Although the CIQF has been widely used by researchers and practitioners for several years, a number of methodological issues may be identified in the process used to create the framework. These issues are discussed below and provide the basis for our replication study described later in the paper.

General Critique

Throughout the CIQF study graduate students and alumni from two eastern U.S. universities acted as subjects. It was believed that those subjects had work experience as data consumers although no indication of how that was determined is given in the original work. This presents at least two concerns: 1) the way subjects were selected is unclear and makes us uncomfortable with their use (i.e. how was it determined subjects were data consumers prior to subject selection), and 2), since all graduate students and alumni had similar training at the

same two universities there could be resultant artefacts in data. These issues may limit the external validity of the study and the generalizibility of the results.

The framework for information quality has been shown to be domain specific. For example, within the healthcare industry it has been suggested that some of the dimensions which were eliminated early in the original study appeared as significant in follow up studies. (Gendron & D'Onofrio, 2001)

There was an excluded cohort in the original study – those who are not data consumers. Data collected from the excluded cohort may be helpful to the understanding of the hypothesized dimensions (Ellenberg, 1994). While this is out of the scope of our proposed replication study it should be dealt with in future studies.

Data Quality Dimensions Creation Critique

The Wang and Strong study was an initial attempt at creating hypothetical dimensions and categories of information quality as perceived by data consumers. The hypothetical dimensions were created in Step-2 using factor analysis which is an appropriate use of that statistical technique (Mulaik, 1972). However, when employing factor analysis it is best if underlying latent factors are validated through multiple replications (Comrey, 1978); a replication is proposed below.

A criterion of factor analysis is a response-to-variable ratio greater than 5 (Streiner, 1994). When the ratio is less than 5, factors may be unstable. Stated differently, there should be at least 5 respondents for each item on a survey instrument.   The results of the Step-2 attribute rating survey yielded a response-to-variable ratio of 3 (118 items/355 respondents), which is less than the recommended ratio of 5. While this could indicate an unstable factor structure the Step-2 results yielded a theoretical set of dimensions against which to test.

Sorting Studies Critique

The use of a preliminary conceptual framework in sorting studies for determining the CIQF could have caused substantial researcher bias. The four groupings originally proposed by Wang and Strong are based on expert opinion rather than empirical analysis. The sorting study technique is accepted in the literature and congruent with the psychology of categorization (Gammack, 1987). These results need to be replicated.

Some dimensions (factors) and categories in the CIQF do not seem plausible. For example the conceptual framework contained the target category *relevancy* which in turn contains the dimensions *completeness* and *appropriate-amount-of-data*, (for more detail see Table 1). While Wang and Strong did rename the category *relevancy* to *contextual data quality* in the CIFQ, they used the term *relevancy* during their sorting studies. Using the word *relevancy* may have produced some bias and we are unclear how the dimensions relate to context.

The hypothesized preliminary categorical framework was based on the concept of fitness for purpose and the researcher's experience with data consumers (Wang & Strong, 1996) which were applied to sorting studies. Sorting study results are based on subjects own method of categorization, human intention, the basic utility of the categories, and results are closely related to individual problem solving heuristics (Stubbart, 1989). Results are largely subjective to individuals and thus may be difficult to interpret and replicate. However, replication of the sorting studies is necessary. It is also important that a deeper understanding

of the CIQF and its dimensions is obtained – this will be accomplished through empirical data collection and analysis.

The Proposed Replication Study

The proposed replication study includes a number of improvements:

- Graduate students will only be used to pilot test instruments. All other empirical data will be collected from practitioner data consumers;

- Larger sample sizes of empirical data will be collected from the industry sectors represented in the original Wang and Strong study to enhance generalizibility and give greater confidence in the validity of the results. The larger sample size of empirical data will increase the variable-to-response ratio and thus enhance the factor stability of Step-2;

- Analysis of industry sectors will be undertaken to determine the domain specificity of the information quality framework – specifically we will compare 2 samples – one stratified according to the industry sector response structure in the original study and a second of equal size within the healthcare sector;

- Visual Card Sorting (VCS) (Budhwar, 2000) studies will be performed which will allow us to confirm the categorical structure of the CIQF. These studies will be performed on two samples which will allow for the comparison of industry sectors and further study of domain specificity of the CIQF;

- The CIQF study will be extended through an additional empirical data collection (post Step-5). This will allow further exploration of the dimensions within the CIQF and its categorical structure. The VCS studies will show how subjects believe the dimensions cluster whereas the empirical data collection will give the ability to statistically analyse the relationships between the dimensions, determine how they co-vary, uncover underlying constructs (perhaps a different categorical structure) and create models that explain information quality;

- Web-based survey collection instruments will be used to collect responses from a large sample size more cost-effectively.

We have described the process to create the original CIQF as five steps. (A summary of our critique and replication study can be seen in Table 7.) Step-1 and the subsequent pre-test of instruments for Step-2 provide results which we are comfortable using for the replication. Therefore, we will start at Step-2, attribute rating and factor analysis. For sake of clarity we will describe the replication as three phases.

Attribute Rating and Factor Analysis (Phase-1)

Phase-1 will continue with the 118 information quality attributes identified above. Two samples will be collected: 1) a reference sample composed of employees from the same industries represented in the original CIQF Step-2 sample and stratified according to its response structure and 2) a comparative sample from with an equal number of subjects from the healthcare sector. Subjects from a variety of departments and management levels who regularly use information to make decisions will be included. Subjects will be randomly

selected from a commercial database of businesses and sales leads available in the U.S. Based on a 10% response rate 7080 subjects are needed per sample; sample size datum can be seen in Table 4.

The references sample responses will be used in a confirmatory factor analyses to validate the dimensions in the CIQF. Additional analysis will be undertaken on the comparative sample to determine if their responses show the same latent factor structure.

Table 4 – Phase-1 Sample Size Datum

118	ITEMS	
5	RESPONSE-TO-VARIABLE RATIO	
20%	MISSING RESPONSE ADJUSTMENT	
708	MINIMUM NUMBER OF RESPONDENTS	
	Subjects	
Reference Sample	7080	
Comparative Sample	7080	

Sorting Studies (Phase-2)

Validation of the categories within the CIQF will be performed using VCS studies. Two VCS passes will occur – Pass-1 will ask respondents to sort dimensions into 3, 4, or 5 piles and Pass-2 will confirm the first pass results as in the CIQF study. Two samples of practitioners will be drawn, a reference sample and a comparative sample. Subjects in the 2 samples will be randomly assigned to VCS Pass-1 or Pass-2. Subjects will be drawn from the same mailing list database as Phase-1 subjects, mailed an invitation to participate in a dinner, presentation and the VCS studies. They will be requested to confirm their attendance. Both samples will include subjects whose companies are in proximity to the U.S. University sponsoring the replication. The samples will have the same subject stratification and structure as Phase-1, but based on the Phase-1 results the VCS studies may actually have different dimensions. Based on a 10% response rate we will need 300 subjects per sample; sample size datum can be seen in Table 5. Phase-2 will allow us to confirm the CIQF in the reference sample and potentially uncover a different category in the comparative sample.

Table 5 – Phase-2 Sample Size Datum

120	Number of Subjects Needed (30 per pass per sample)		
	Number of Subjects Assuming 10% Response Rate		
Reference Sample	Pass-1: 300	Pass-2: 300	600 overall
Comparative Sample	Pass-1: 300	Pass-2: 300	600 overall

Extending the CIQF (Phase-3)

In order to extend our understanding of the CIQF, empirical data collection is planned using the Phase-1 dimensions as items. Survey data will be collected from a reference and comparative sample. The samples will have the same structure as the in Phase-1 and will be drawn from the same database. This will give the ability to statistically analyse the relationships between the dimensions, determine how they co-vary, uncover underlying constructs (perhaps a different categorical structure between reference and comparative samples) and create models that explain information quality.

Since the results of Phase-1 are not yet available, the number of items that will appear on the Phase-3 survey can only be estimated. It could be assumed that the same 20 dimensions in the CIQF will emerge, however in order to allow for the maximum flexibility in final analysis a substantially larger quantity is estimated. Therefore, the number of items is increased to 30. Based on a 10% response rate, 2250 subjects are needed per sample; sample size datum can be seen in Table 6. Extending the CIQF, uncovering latent constructs through exploratory factor analysis and model creation will occur in Phase-3.

Table 6 – Phase-3 Sample Size Datum

30 ITEMS	
5 RESPONSE-TO-VARIABLE RATIO	
20% MISSING RESPONSE ADJUSTMENT	
225 MINIMUM NUMBER OF RESPONDENTS	

	Subjects
Reference Sample	2250
Comparative Sample	2250

Sample Creation and Use of Web-based Survey Instruments

The largest mailing-list database provider in the U.S. will provide the subject pools. They will provide a random subject pool by various demographics (e.g., job title, standard industry classification, major industry grouping, etc.). They will also send the mailing to the subjects which will contain a link directing them to a website to complete surveys for Phase-1 and -3. Phase-2 subjects will be randomly extracted from the same database, but will be contacted by mail and telephone to request their participation in the study. The members of the email list opt in, and thus are expecting email distributed by the list provider. The mailing-list provider has guaranteed the list members that their email addresses will not be sold. The list provider's database contains over 3 million managers within U.S. companies.

The proposed samples probably limit the gereralizibility of our study to those managers who have access to the Web and email. Since data consumers/managers usually use technology it makes sense that subjects who use the Web and email are more likely to be information consumers, and thus make good candidates for this study. They must have some level of technical ability to use those technologies and those technical abilities are inherent in people who use information in their job.

Table 7 – Study Comparison

General Critique: Only used graduate students and alumni from same two U.S. universities; CIQF has been shown to be domain specific; Excluded cohort was not examined.			**General Enhancements:** All subjects will be selected from a nationwide US database of practitioners; A reference and a comparative sample will be taken to allow for examination of domain specificity; Examination of excluded cohort is outside the scope of this study		
CIQF	**Procedure**	**Critique**	**Replication**	**Procedure**	**Enhancement**
Step-1	Potential data quality attributes were generated		Use Attributes Generated in Original Study for Replication		
Step-2	Survey using 118 items Exploratory factor analysis	Replication of hypothetical factors has never been undertaken; Response-to-variable ratio < 5	**Phase-1**	Survey using 118 items; Confirmatory factory analysis on reference and comparative sample	Larger sample size yielding more stable factor structure; Factor structure of both samples will be examined to better understand potential domain specificity of the dimensions
Step-3	Preliminary conceptual framework created from researchers prior experience with data consumers	Use of preliminary conceptual framework may have been a cause of bias	**Phase-2**	VCS studies will be performed using from both the reference and comparative samples	Only practitioners will be use for VCS studies; VCS studies will be performed using both the reference and comparative samples to examine their any potential differences in perceived categorical structures.
Step-4	Subjects asked to sort 20 dimensions into 3, 4, or 5 piles	Some of the dimensions and category groupings do not seem plausible; Card sorting is inherently subjective			
Step-5	Subjects asked to place 15 dimensions into appropriate category				
			Phase-3	Survey using dimensions from Phase-1 sent to subjects in a reference and comparative sample for exploratory factor analysis	Additional data collection to allow exploration of dimensions; examine co-variation; uncover any existing alternate categorical structure; create models of information quality; use of reference and comparative samples

Conclusion

The Wang and Strong CIQF has been used widely by researchers and practitioners for several years. In this paper we have identified a number of methodological issues in the process used to create the framework. A replication study has been proposed to validate the categories and dimensions of the framework to ensure they operationalize the construct *information quality* and examine the domain specificity of information quality.

The replication study has a number of important implications for both practitioners and researchers Practitioners will be provided with a framework in which all constructs are fully validated. The ability to further understand the differences between data consumers' definitions of data and information quality in different industry sectors will be enhanced. The framework will provide a solid basis for the development of information quality metrics and information quality improvement programs. Researchers will have an empirically based and validated framework for understanding information quality that may be used as the basis for further work in all aspects of information quality. Information quality dimensions and categories identified may be used in the development of information quality metrics for studies of the impact of information quality on decision processes and outcomes. (Shanks & Tansley, 2002) Further replication studies in different national and cultural settings will enable cross-cultural comparisons of information quality to be undertaken.

Information is being increasingly recognised by managers as an important and strategic asset that must be managed. A high quality framework to understand information quality is a basic requirement for successful management of information. The replication study described in this paper for the validation and refinement of the CIQF is a key step in this direction.

References

Agmon, N., & Ahituv, N. (1987). Assessing data reliability in an information system. *Journal of Management Information Systems, 4*(2), 34-44.

Budhwar, P. (2000). The use of visual card sorting techniques to study managers' belief structure - An international comparative study. *Journal of Managerial Psychology, 15*(5), 440-459.

Comrey, A. L. (1978). Common Methodological Problems In Factor Analytic Studies. *Journal of Consulting and Clinical Psychology, 46*, 648-659.

D'Onofrio, M. J., & Gendron, M. S. (2001). *Technology Assisted Research Methodologies: A historical perspective of technology-based data collection methods.* Paper presented at the Internet Global Summit: A Net Odyssey - Mobility and the Internet - The 11th Annual Internet Society Conference, Stockholm Sweden.

Ellenberg, J. H. (1994). Selection bias in observational and experimental studies. *Stat Med, 13*(5-7), 557-567.

English, L. (1999). *Improving Data Warehouse and Business Information Quality*. New York: Wiley Computer Publishing.

Gammack, J. G. (1987). *Modelling expert knowledge using cognitively compatible structures.* Paper presented at the Third International Expert Systems Conference, London.

Gendron, M. S., & D'Onofrio, M. J. (2000). *An Exploratory Study Investigating Data Quality in the Healthcare Industry: What are the Implications for Data Warehousing.* Paper presented at the Americas Conference on Information Systems, Long Beach, CA.

Gendron, M. S., & D'Onofrio, M. J. (2001). Data Quality in the Healthcare Industry. *Data Quality, 7*(1).

Gendron, M. S., & D'Onofrio, M. J. (2002, 9/2002). *A Model for Making Decisions Using Quality Attributes.* Paper presented at the DSIAge, Cork Ireland.

Haung, K.-T., Lee, Y. W., & Wang, R. Y. (1999). *Quality information and knowledge*. Upper Saddle River: Prentiss-Hall.

Huang, K.-T., Lee, Y. W., & Wang, R. Y. (1999). *Quality Information and Knowledge*. Upper Saddle River: Prentice Hall.

Kahn, B., Strong, D.M. and Wang, R.Y. (1997). Information Quality Benchmarks: Product and Service Performance. *Communications of the ACM, June*.

King, W. R., & Epstien, B. J. (1983). Assessing information system value: an experimental study. *Decision Sciences, 14*, 34-45.

Kovac, R., Lee, Yang and Leo Pipino. (1997). *Total Data Quality Management: The Case of IRI.* Paper presented at the Conference on Information Quality.

Mingers, J. C. (1995). Information and Meaning: Foundations for an Inter-subjective Account,. *Information Systems Journal, 5*, 285-306.

Mulaik, S. A. (1972). *The Foundations of Factor Analysis*. New York: McGraw Hill.

Redman, T. (2001). *Data Quality: the Field Guide*. New Jersey: Digital Press.

Shanks, G., & Darke, P. (1998). *A Framework for Understanding the Quality of Data in Data Warehouses.* Paper presented at the International Conference on Information Quality, MIT, Boston.

Shanks, G., & Tansley, E. (2002). *Data Quality Tagging and Decision Outcomes: An Experimental Study.* Paper presented at the IFIP Working Group 8.3 Conference on Decision Making and Decision Support in the Internet Age, Cork, Ireland.

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

Streiner, D. (1994). Figuring out factors: the use and misuse of factor analysis. *Canadain Journal of Psychiatry, 39*(3), 135-140.

Stubbart, C. I. (1989). Management cognition: a missing link in srategic management research. *Journal of Managemetn Studies, 26*(4), 325-347.

Wand, Y., & Wang, R. (1996). Anchoring data quality dimension in ontological foundations. *Communications of the ACM, 39*(11), 86-95.

Wang, R. Y., Lee, Y., Pipino, L., & Strong, D. M. (1998). Manage your information as a product. *Sloan Management Review, 39*(4), 95-105.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: what data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5-34.

Weber, R. (1997). *Ontological Foundations of Information Systems*. Melbourne: Coopers & Lybrand Accounting Research Methodology Monograph No. 4.