Data quality assessment in context: A cognitive perspective

Stephanie Watts [a,*], G. Shankaranarayanan [b], Adir Even [c]

[a] Information Systems Department, Boston University School of Management, Boston, MA, USA
[b] Technology, Operations, and Information Management, Babson College, Babson Park, MA, USA
[c] Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, Beer-Sheva, Israel

ARTICLE INFO

ABSTRACT

In organizations today, the risk of poor information quality is becoming increasingly high as larger and more complex information resources are being collected and managed. To mitigate this risk, decision makers assess the quality of the information provided by their IS systems in order to make effective decisions based on it. To do so, they may rely on quality metadata: objective quality measurements tagged by data managers onto the information used by decision makers. Decision makers may also gauge information quality on their own, subjectively and contextually assessing the usefulness of the information for solving the specific task at hand. Although information quality has been defined as fitness for use, models of information quality assessment have thus far tended to ignore the impact of contextual quality on information use and decision outcomes. Contextual assessments can be as important as objective quality indicators because they can affect which information gets used for decision making tasks. This research offers a theoretical model for understanding users' contextual information quality assessment processes. The model is grounded in dual-process theories of human cognition, which enable simultaneous evaluation of both objective and contextual information quality attributes. Findings of an exploratory laboratory experiment suggest that the theoretical model provides an avenue for understanding contextual aspects of information quality assessment in concert with objective ones. The model offers guidance for the design of information environments that can improve performance by integrating both objective and subjective aspect of users' quality assessments.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Organizational data is a critical resource that supports business processes and managerial decision making. Advances in information technology have enabled organizations to collect and store more data than ever before. This data is processed in a variety of different and complex ways to generate information that serves as input to organizational decision tasks. As data volumes increase, so does the complexity of managing it and the risks of poor data quality. Poor quality data can be detrimental to system usability and hinder operational performance, leading to flawed decisions [27]. It can also damage organizational reputation, heighten risk exposure, and cause significant capital losses [28]. While international figures are difficult to determine, data quality problems currently cost U.S. businesses over $600 billion annually [1]. Data quality is hence an important area of concern to both practitioners and researchers. Data quality researchers have used the terms "data quality" and "information quality", interchangeably. We choose to use "information quality" in this paper because we investigate the quality of "inputs to decision tasks", which typically are the informational outputs of processed data.

Information quality may be measured along many dimensions such as accuracy, completeness, timeliness, relevance, believability, and others [43]. Some of these dimensions (e.g., accuracy and completeness) lend themselves to objective measurement – measurement of quality that is intrinsic to the information itself, regardless of the context in which it is used. Researchers have proposed that these objective quality measurements be linked to the information used in the decision task, in order to provide decision-makers with this additional information. Such measurements have been referred to as data tags [42], data quality information [7,12], and quality metadata [31,32]. In this paper we refer to these measurements as quality metadata. Research has shown that provision of quality metadata along with its associated information results in different decision outcomes than when the decision is made using the relevant information alone [9,12].

There are, however, dimensions of quality that cannot be objectively measured. For example, two widely acknowledged information quality dimensions, relevance and believability [43,44], tend to vary with the usage context. Information relevance generally depends on the task that it is applied to, since information that is highly relevant for one task may be irrelevant for another. Information believability is also difficult to evaluate objectively, since it often depends on the user's experience and personal preferences – for example, certain information that appears to be believable to a novice may be less

* Corresponding author.
E-mail addresses: swatts@bu.edu (S. Watts), gshankar@babson.edu (G. Shankaranarayanan), adireven@bgu.ac.il (A. Even).

believable to an expert. Fogg et al. examined the credibility of web sites and define credibility as a synonym for believability [13]. They support our view that believability is a subjective measure, and identify expertise as a key factor affecting believability. Prior research in information quality has not addressed these dimensions in depth (c.f. [22]), due in part to the theoretical and operational challenges of incorporating *contextual usage* into quality assessment. The context of use is defined by such factors as characteristics of the decision task and characteristics of the decision-maker.

To understand the contextual effects of information quality, it is important to account for factors pertaining to information *in use*. Hence information quality is often defined as "fitness for use" [28,38]. In this view, factors such as the relevance of the information to the task, the ability of the user to understand it, and the clarity of the task, all affect the usability of that information. From this usage perspective, quality assessment tends to be contextual - information that is of acceptable quality for one decision context may be perceived to be of poor quality for another decision context, even by the same individual. Research has posited that another type of metadata, process metadata (also referred to as data lineage), influences contextual quality assessment [21,30], but has not addressed how users may utilize this relatively new form of metadata to assess contextual information quality. Process metadata is an abstracted and often visualized description of information processing stages such as acquisition, transformation, storage and delivery [31]. Shankaranarayanan et al. show that providing end-users with process metadata in addition to quality metadata can enhance end-users assessment of quality, their capability to perform a decision task efficiently, and thus their objective decision outcomes [32]. What is not evident is how end-users use quality and process metadata for assessing the quality of data resources and what the roles of user and task characteristics are in this process.

Understanding how end-users utilize metadata as input for assessing data quality during the decision-making process has important implications for the design of decision environments. Metadata provides additional information over and above the information needed for the decision task, and therefore can cause information overload and impact decision performance. Understanding how users use metadata, and how task and user characteristics influence this process, will help to design decision environments that provide the right set of information (i.e., whether to include metadata or not) to the right set of users (i.e., defined by user characteristics) for the right task (i.e., defined by task characteristics). In this way, this study seeks to further our knowledge of the role of the end-user in the information quality assessment process. After all, users that deem information to be of poor quality are unlikely to weight it heavily in their decision making tasks even if it is objectively of high quality. In this study we build and then test a theoretical model of contextualized information quality assessment that takes characteristics of both the user and the task into account. This model is built on a widely accepted body of cognitive theory referred to as dual-process information processing [11,25] that has been applied to a number of other online contexts, such as email and the Web (e.g., [10,15,16,20,37,39]). This body of theory is ideally suited for understanding the problem of assessing information in context, since it explains how individual and task factors interact with the processing of new information. From this model we develop hypotheses about information quality assessment, and we investigate these using a laboratory study of 51 masters-level information-systems students.

This research represents an important extension to the data quality management (DQM) literature, which has primarily focused on objective characteristics of the information itself, without taking into account additional factors reflecting how people perceive and use the information in context. Researchers in DQM have proposed analytical tools for measuring quality along objective quality dimensions, developed tools for monitoring and cleansing information, and examined the scope of information management responsibilities (e.g., [4,15,18,24,26,40,43]). While these approaches have made solid

inroads towards increasing information quality, they typically address quality from a technical or organizational perspective, insufficiently exploring the implications of the individual information user. Because users are responsible for gauging information quality and accounting for it in their decision processes, their quality assessments can significantly impact decision outcomes and performance. This study offers important insights into how individuals assess quality and how these assessments can affect their decision performance.

This research makes several important contributions. It examines information quality at an individual level by developing a theoretical model of how user characteristics (e.g., expertise) and task characteristics (e.g., ambiguity) affect information processing during the execution of a decision task. The model emphasizes the importance of *both* the objective and contextual aspects of information quality assessment for improved support of decision making. By simultaneously observing both objective and contextual quality assessment, it examines how decision makers process these two different types of information. A dual-process theoretic explanation of our findings suggests that both systematic processing of information and heuristic processing of quality metadata affect decision outcomes, and that the relative strength of these two processing modes depends on characteristics of the user in context. In the remainder of this paper, we first provide relevant background on information quality assessment and the dual-process theories, since this body of theory underlies our model of individual quality assessment. We then elucidate the model and the hypotheses that are based on it. Next we describe an exploratory laboratory study undertaken to assess these hypotheses and the validity of the theoretical model. We close by discussing the implications of this work for research and practice.

2. Background and research model

2.1. Information quality assessment

Information quality is a multidimensional concept [43], having both objective aspects that do not vary across tasks and users (e.g., accuracy and consistency) and contextual aspects, related to the perceptions of decision makers who use the information. Contextual assessments depend on the requirements and characteristics of the task at hand, and characteristics of the decision maker such as experience. While there is some disagreement in the literature regarding the fine-line that differentiates the two (e.g., [26,43]), in this paper we adopt the view that objective measurements of quality are derived purely from the dataset itself, while contextual assessments are a function of assessments of quality as moderated by two extraneous factors - characteristics of the decision-maker and characteristics of the decision task. Here we examine contextual perceptions of information quality and show that their influence on decision making is moderated by two characteristics of the decision-maker and the task – expertise and ambiguity, respectively. We use cognitive theory to explain the mechanisms underlying these moderation effects, in order to illustrate that their influence on quality assessment is context dependent.

The context-dependent nature of information quality has been widely acknowledged in the form of the important information quality attributes of relevance and believability [15,43,44]. Information relevance refers to the extent to which the information is applicable and useful for the decision task that the information is targeted for [15,26]. Information believability is the extent to which information is considered to be true or credible [15,26]. Believability may reflect an individual's assessment of the credibility of the source, the comparison of the information value to a commonly accepted standard, and prior experience [26]. Clearly relevance and believability are important attributes of information quality, ones that affect how practitioners utilize information. Yet they are difficult to measure because they are contextual – information that is relevant and

believable to one decision maker may be irrelevant and unbelievable to another, or even to the same decision maker for a different task. For example, research in geographical information systems (GIS) has addressed the contextual nature of assessing quality [14]. These authors interviewed experts to understand the range of acceptable values for cadastral information in gauging the quality of topographical maps, and found the range to be broad [17]. Thus researchers are beginning to investigate the contextual nature of information quality, but have not examined its implications for decision making, nor have they utilized robust theoretical models of the decision process such as the cognitive one we develop here. The influence of contextual quality assessment on decision-making has been under-researched, in part due to lack of a theoretical basis for doing so. In this way this study answers the call for increased use of validated theoretical models in decision support research [2].

2.2. Dual-process theories of human information processing

To address this gap in the literature, we sought a theory that would take contextual and individual factors into account in order to explain information quality assessment. Dual-process theories of human information processing describe how individuals process newly received information. This body of theory is widely accepted among cognitive psychologists [35]. Cognitive theories are increasingly being utilized for understanding human-computer interaction [5,20,29], but have yet to be applied to the problem of understanding information quality assessment. The dual-process theories originated from individual-level, laboratory-based psychology research, and have been successfully applied to many domains for understanding how people process received information [9,10,15,16,35,37,41]. The dual-process approach encompasses a family of theories, all of which examine both the information content of received information and factors in the surrounding context. Here we focus on the heuristic-systematic model (HSM) [11] as a representative and well-established variant of this theory base. According to the HSM, an individual processes received information in two ways – heuristically and systematically, as follows: When faced with new information such as quality metadata, individuals apply pre-existing frames and heuristics to process it efficiently, and/or they undertake the relatively greater cognitive effort required to systematically analyze it. For example, during heuristic processing, people may utilize simple decision rules such as "credibility implies correctness" [6] to assess content validity. Alternatively, they may disregard the source entirely and assess the validity of the received content on the basis of its inherent merit, independent of its associated context. Individuals are continuously undertaking one or both of these two types of cognitive processes as they go about interpreting new information. Heuristics provide a very important means for dealing with the vast quantities of information we face daily. Due to bounded rationality, and because of the cognitive effort involved, individuals are not able to systematically process all the informational stimuli in their environment. For example, we utilize heuristics to determine which emails to delete without reading them and which advertisements to attend to in detail.

An important aspect of this theory for our purposes is that it explains the mechanisms underlying how people make tradeoffs among these two processing modes. The greater the cognitive resources available, and/or the greater the motivation to process, the greater the likelihood that a particular individual will undertake the additional cognitive effort to systematically process the newly received information, as opposed to relying on heuristics. For example, a person that is an expert on the topic is more likely to undertake systematic processing than one who is a novice [25]. Similarly, high levels of information ambiguity tend to motivate people to undertake the systematic processing necessary to resolve this ambiguity. Lack of expertise on the subject, limitations in terms of time and energy, and distractions and disruptions, increase the tendency of people to rely on heuristic processing. Thus the HSM helps

to explain why individuals react differently to identical information and the specific role of expertise and task ambiguity in this process. When the user is a human and not a machine, this individual-level theory of information processing offers insights into how the user processes new information during decision tasks.

2.3. Metadata and systematic processing

In order to understand the relative influence of systematic and heuristic processing on information quality assessment, we incorporate metadata. Metadata is abstracted information about data that provides end-users with additional details about the dataset, beyond the dataset itself. Among the various metadata types, quality metadata and process metadata (also referred to as "lineage" in the GIS literature) are most relevant to the assessment of information quality [31]. Quality metadata provides measurements of objective quality attributes such as accuracy and completeness. Process metadata provides information about the stages that the data went through on its journey to reach the user, such as how it was acquired, transformed, stored and delivered [31].

Utilizing process metadata requires systematic analysis on the part of the user. Provision of process metadata, in addition to quality metadata, helps us understand users' relative levels of heuristic and systematic processing of quality information. Here we provide process metadata to the decision maker in the form of an information-product map (IPMAP). The IPMAP is a visual representation of process metadata that tracks data sources, data processing at each of the various stages (including assumptions and associated business rules), and how the data has been aggregated to create the final information [33]. In addition to showing the processing resources used in the manufacture of the information product, the IPMAP also shows the flow of data elements and the sequence in which they are processed. In order to gauge the information reflected in quality and process metadata and use it as an additional input to the decision-making process, end-users must invest additional time and cognitive resources beyond that required for processing and interpreting the information itself. For the purposes of this research, users who incorporate metadata into their quality assessment process are undertaking the additional cognitive effort typical of systematic processing. On the other hand, users are more likely to process newly received information heuristically when they do not have the motivation or ability to process all information relevant to it, in this case, its metadata. Users are more likely to incorporate metadata to fully assess quality when they engage in systematic processing, above and beyond the minimal information processing that they undertake heuristically.

2.4. Expertise

According to the HSM and other dual-process theories, both systematic and heuristic processing influence how people assess received information [6,25]. However, the extent to which users engage in either type of information processing is affected by their level of expertise. As discussed above, the greater the expertise of the information recipients in the relevant domain, the more likely it is that they will undertake systematic processing. Therefore, systematic processing is more influential on the outcome of assessments made by experts than assessments made by novices. Applying this HSM precept to the quality assessment process suggests that user expertise in the task domain should affect how that user processes quality information. When users are experts in the given task, they are more able and hence more likely to undertake systematic processing of received information than are novices, for the same task. According to the HSM, information recipients will engage in systematic processing when heuristic processing is deemed insufficient [26]. Thus in the context of quality assessment, experts are likely to augment their heuristic processing of received information with

systematic processing of the associated metadata. By contrast, novices are likely to base their quality assessments on heuristic processing alone without undertaking additional systematic processing of metadata. Thus we expect to find that higher levels of user expertise (for the decision task at hand) will result in higher levels of systematic processing, as reflected in performance outcomes. Therefore:

H1. For users with high levels of expertise (for the task at hand), systematic processing will be more strongly associated with performance outcomes than heuristic processing.

Since novices do not have the requisite expertise to process new information systematically, they tend to rely on heuristic processing when they are faced with new information. Thus:

H2. For users with low levels of expertise (for the task at hand), systematic processing will be less strongly associated with performance outcomes than heuristic processing.

According to the HSM, we should not expect to see main effects of user expertise on performance, since expertise serves to moderate the processing modes rather than affecting performance directly [26]. The following hypothesis addresses this counter-intuitive situation:

H3. There will be no significant direct association between expertise and performance outcomes.

Hypotheses one and two investigate characteristics of the information user. We are also able to use the HSM to take into account factors relating to the nature of the task.

2.5. Task ambiguity

Here we focus on ambiguity, a task characteristic that is strongly linked with complexity and has been widely studied under the dual-process theories. Ambiguous task environments are those in which no single outcome is inherently correct, and multiple potentially acceptable solutions exist [8]. Ambiguous tasks tend to motivate greater cognitive effort on the part of the person undertaking them as they seek to understand the task thoroughly so that an optimal result can be obtained [19]. Perceptions of task ambiguity may vary between users – a task that is highly ambiguous for one user may not be ambiguous to another (e.g., one who practiced a similar task before). According to the HSM, users who perceive a task to be more ambiguous are more likely to engage in systematic processing rather than heuristic processing while performing that task [11]. We therefore theorize that:

H4. For tasks with high levels of perceived ambiguity, systematic processing will be more strongly associated with performance outcomes than heuristic processing.

Conversely, when the task is clear and unambiguous, it is less likely to motivate the additional effort necessary to engage in systematic processing. Thus for unambiguous tasks, we would expect to see greater reliance on heuristic processing:

H5. For tasks with low levels of perceived ambiguity, systematic processing will be less strongly associated with performance outcomes than heuristic processing.

And finally, as with expertise, according to the HSM the effect of task ambiguity on performance is indirect. It operates by altering the type of processing one engages in, not by affecting performance directly. Therefore we do not expect to find a significant association between task ambiguity and performance:

H6. There will be no significant association between task ambiguity and performance.

Fig. 1 below depicts the model and hypotheses discussed above. The arrow between the constructs "Systematic Processing" and "Performance Outcomes" and the arrow between "Heuristic Processing" and "Performance Outcomes" represent the main effects of systematic and heuristic processing on task performance. The moderating effects of "User Expertise" on each of these main effects are represented by the vertical arrows that terminate on each of the two main-effect arrows. Similarly, the vertical arrows from "Task Ambiguity" represent the moderating effects of this construct on the main effects.

3. Research method

The laboratory experiment described in this section is an exploratory attempt to empirically assess the theoretically-derived hypotheses above. We chose a laboratory-based method because HSM researchers rely on laboratory methodology almost exclusively. And, since the HSM has not been applied to understanding how people undertake information quality assessment, we sought to learn whether our model applies in a laboratory setting before we investigate its application in the field.

3.1. Participants, task and procedures

In a laboratory experiment, 51 masters-level information systems graduate students participated in a computer-based decision-making task. All of these students have organizational work experience. Participants were all assigned the same computer-supported, information-driven, decision making task of planning an advertising campaign. Given a fixed advertising budget, participants were asked to allocate it among multiple types of advertising media (Billboards, Magazines, Radio and TV) and multiple geographical locations, such that the maximum number of potential customers would be exposed to the campaign. To support this decision task, information was provided using a spreadsheet. It indicated the estimated number of people who would be exposed to each type of media within each geographical region. The information also provided estimations of exposure efficiency history, calculated as the average number of people exposed to the product per dollar spent on advertising. This task was chosen because it is information intensive, and also because our pilot study indicated it to be of medium ambiguity. In order to achieve variance in the measure of perceived task ambiguity, it was necessary to select a task that would be perceived by some as relatively straight forward, and by others to be somewhat ambiguous.

In addition to the actual information provided in the spreadsheet, participants were provided with two additional types of metadata regarding the quality of that information. The first reflected an assessment of the information along the standard quality metrics of accuracy, completeness, currency, consistency and relevance. This parsimonious set of indicators is standard in the information quality literature [42]. Levels of these quality dimensions were indicated with colors analogous to street light signals – red indicated low quality, yellow reflected medium quality, and green indicated high quality. The main screen of the spreadsheet interface is shown in Fig. 2. As these metrics were provided in visual traffic light format, they could be utilized quickly and heuristically, without systematic analysis.
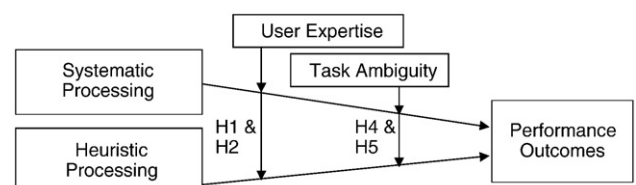

Fig. 1. Dual-process model of Information Quality Processing.

Fig. 2. Spreadsheet interface used for providing information.

Users were also provided with process metadata in the form of an information product map (IPMAP) shown in Fig. 3. The information product described is an exposure report prepared for an advertising director. According to the IPMAP, regional data is used to compute the exposure effectiveness for that region. This map provides more detail than the quality indicators provided in traffic signal colors. This process metadata was provided in order to offer participants more material to systematically process. However, since systematic processing entailed analysis of the entire set of information for the purpose of problem solving, we were not able to use the extent of IPMAP usage as a surrogate for systematic processing.

The Excel-based decision support tool allowed the participants to explore different budget allocations using a "what if" style to get immediate feedback on the overall expected level of exposure (E), and an indication of the decision quality based upon the underlying information (Q). The final performance score was used as the dependent measure, and calculated using a geometric-average of two scores: the first score consisted of the estimated number of people exposed to the advertisement campaign given the budget allocation – the final solution to the assigned task. The second score reflected the quality of this solution, based on the extent that high quality information

elements were used to calculate it. The final score (S) was calculated as a geometrical-average of exposure score E and the quality score Q (S = √E*Q). This combined score was used to measure performance outcomes, and hence reflects objective task outcomes rather than perceived ones. For statistical purposes the combined score was linearly rescaled by setting the maximum score obtained by the top scoring participant to one and linearly adjusting the rest to this scale.

Participants were given an initial overview of the task by the experiment coordinator. Included in this overview were the details of how the final score was to be computed. Participants were then given 20 minutes to complete the task and were instructed to allocate the budget so as to maximize the final score. The spreadsheet interface provided participants with their composite performance score based on their current allocation and associated performance quality at all times during the task. Participants were not aware of the progress or score of the other participants. A hidden back-end process tracked the navigation and the time-spent on each screen. In order to motivate them to work to achieve an optimal decision outcome, the task was conducted as a competition in which cash prizes were offered to the four participants that obtained the highest final scores. Participants filled out a survey instrument after completion of the experimental
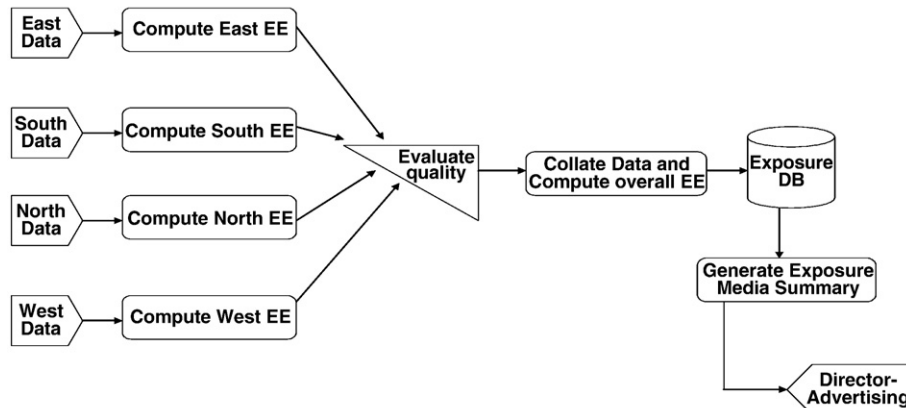

Fig. 3. Sample Information Product Map.

task. The survey consisted of measures of the model constructs discussed below. See Appendix A for the actual items used and their origins. Participants were assigned a secure code to ensure their anonymity and their demographics were reported as part of this survey. We pre-tested the task and the survey on a sample of ten graduate (doctoral) students. A majority of this group had over six years of industry experience in a managerial decision-making capacity.

The survey measured participants' perceptions of subjective aspects of the information – to what extent they found it to be valuable, informative, helpful, and useful. It also contained items to measure participants' self-assessed expertise levels and their perceptions of task ambiguity. These measures utilized previously validated scales. However, no previously validated measure of systematic processing was available, because this has generally been a manipulated variable in the laboratory. The items for this construct were developed by one author, revised by the others and then pre-tested on the sample of ten graduate students (the same set referred to above who did not participate in the actual experiment).

All model constructs except the dependent measure were measured perceptually. This was necessary since the theory describes an intra-psychic cognitive phenomenon. An example makes this clear: A person that receives high scores on a series of training tasks in a new domain is not likely to see themselves as an expert in the new domain. In such cases, perceptions of domain expertise rather than actual expertise levels will drive tradeoffs among systematic and heuristic information processing. For this reason, subjective measures of expertise are preferable to objective, external ones. This same logic holds true for task ambiguity. Regardless of how objectively ambiguous a task is, that task will be perceived as more or less ambiguous depending on the tendencies of the person faced with it, and it is these perceptions that drive information processing tradeoffs.

As mentioned earlier, we were not able to use IPMAP usage as a surrogate for systematic processing, since such processing encompassed more than assessment of the IPMAP. However, in our pilot test, those using the IPMAP to assess data quality took a significantly longer time (objectively measured using an online tool) to complete the task than those who did not. This gave us confidence that use of the IPMAP required systematic processing.

3.2. Preliminary measures and hypothesis testing

Participants consisted of 35 males and 16 females, reflecting a gender distribution typical of graduate programs in information systems. Participants' mean age was 27.6, with a standard deviation of 2.6. Their mean work experience was 4.8 years, with a standard deviation of 2.2. There were no significant differences among constructs due to gender, age or years of work experience.

The model constructs were assessed for conformance to Ordinary Least Squares (OLS) assumptions. All independent constructs show acceptable internal reliability and discriminant validity. Cronbach's alphas for these constructs are within acceptable bounds (> 0.75) for exploratory research, while all Cronbach alphas between constructs were less than 0.5. Table 1 presents statistical descriptors for each construct.

Hypotheses were tested using OLS regressions in SPSS due to the small sample size. We report first on main-effect analyses, and then on moderation effects. Hypotheses 3 and 6 predict a lack of main performance effects of either expertise or task ambiguity, on the basis of HSM theory. For expertise, this was borne out: regressing performance onto expertise resulted in a non-significant model ($F = 2.13$, sig. $F = 0.15$, adj. R^2 of 0.02, d. f. $= 1$, 48). Note that this does not confirm the null hypothesis. Contrary to theory (H6) however, perceptions of task ambiguity were found to be significantly and negatively associated with performance: ($F = 6.31$, sig. $F = 0.015$, Adj. $R^2 = 0.096$, d. f. $= 1$, 48; $\beta = -0.34$).

3.3. Effect of expertise

Next we investigated the moderation effects hypotheses - H1, H2, H4, and H5, measuring them in two ways. First, a strong form of moderation was tested for by multiplying the individual indices of each item and seeking significance of the resultant moderation term above and beyond that generated by the two main effects. Where this method proved non-confirmatory (e.g. H1, H2 and H5), we used split sample analysis to explore the direction of moderation effects, in accordance with precepts of the theoretical model.

H1 and H2 investigate the extent that levels of user expertise affect the strength of association between the two processing types and performance. Using the strong form of moderation testing described above, we regressed performance onto perceived systematic processing, expertise and the interaction term. The interaction term was not significant, but the overall model explained 41% of performance variance ($F = 12.50$, sig. $f = 0.000$, d. f. $= 3,47$), 8% greater than the main effects model, representing a significant increase in variance explained above the main effects model (Sig. of Δ $F = 0.047$). For details see Table 2 below. We then split the sample at the median to see whether this would reveal a pattern of significance consistent with dual-process theory. For those participants reporting above median levels of expertise, perceived systematic processing accounted for 43% of the variance in performance ($F = 16.54$, sig. $F = 0.001$, d. f. $= 1$, 20). Participants reporting below median expertise had a weaker association between systematic processing and performance (Adj. $R^2 = .29$, $F = 10.58$, sig. $F = 0.004$; d. f. $= 1$, 22). The relative decrease in adjusted R^2 from higher to lower levels of expertise supports a weak form of interaction effect that is consistent with H1.

H2 investigates the extent that perceived expertise levels influence the strength of the relationship between heuristic information processing and performance. Regressing performance outcome onto heuristic processing, expertise, and an interaction term calculated as described for H1 above, the overall model was significant (adj. $R^2 = .15$, $F = 3.76$, sig. $F = 0.017$, d. f. $= 3$, 46) but explained less variance than did systematic processing, reported above. Similar to H1, the interaction term explained 8% of the additional variance. However unlike H1, this

Table 1
Means, Standard Deviations, and Pearson Correlation Coefficients among constructs.

		Cron-bach's Alpha	Mean	Std. Dev.	1	2	3	4	5
1	Heuristic Processing	$\alpha = .8398$	4.74	1.11	1.0				
2	Systematic Processing	$\alpha = .8021$	3.10	1.39	.359*	1.0			
3	Expertise	$\alpha = .8604$	2.69	1.44	−.091	.086	1.0		
4	Task Ambiguity	$\alpha = .7728$	4.05	1.35	−.360*	−.030	.219	1.0	
5	Score (objective performance outcome)	N/A	.550	0.06	.341*	.605**	−.204	−.338*	1.0

*Indicates $p < 0.05$ (two-tailed).
**Indicates $p < 0.01$ (two-tailed).

Table 2
Details of moderation analyses resulting in significant findings.

Independent Construct		Main effect impact on Score	Impact of Moderator	Overall moderated model: Main effect, moderator and interaction term			
H1	Systematic Processing	Adj. $R^2 = .35$ (d.f. = 1,48), F = 28.29*** $\beta = .61$	Expertise Δ Adj. $R^2 = .08$*	Adj. $R^2 = .41$ F = 12.5*** (d.f. = 3,47)	Sys.Proc. $\beta = .443$ t = 2.03* sig. = .05	Expertise $\beta = -.518$ t = -1.80 sig. = .078	Sys.Proc. x Expertise $\beta = .347$, t = .976 n/s
	Under above median expertise:	Adj. $R^2 = .43$ (d.f. = 1,20), F = 16.54*** $\beta = .67$					
	Under below median expertise:	Adj. $R^2 = .29$ (d.f. = 1,22) F = 10.58** $\beta = .57$					
H2	Hueristic Processing	Adj. $R^2 = .10$ (d.f. = 1,47) F = 6.30* $\beta = .34$	Expertise Δ Adj. $R^2 = .08$*	Adj. $R^2 = .15$ F = 3.76* (d.f. = 3,46)	Heur. Proc. $\beta = .732$ t = 2.77** sig. = .008	Expertise $\beta = .846$ t = 1.45 sig. = .153	Heur.Proc. x Expertise $\beta = -1.077$ t = -1.78 sig. = .082
	Under above median expertise:	Adj. $R^2 = .02$ (d.f. = 1,20) F = .503 $\beta = .157$					
	Under below median expertise:	Adj. $R^2 = .24$ (d.f. = 1,21) F = 8.04** $\beta = .53$					
H4	Systematic Processing	Adj. $R^2 = .35$ (d.f. = 1,48) F = 28.29*** $\beta = .61$ (same as above)	Ambiguity Δ Adj. $R^2 = .14$	Adj. $R^2 = .47$ F = 15.85*** (d.f. = 3,47)	Sys. Proc. $\beta = .094$ t = .316 sig. = .753	Ambiguity $\beta = -.661$ t = -3.07 sig. = .004	Sys.Proc. x Ambiguity $\beta = .626$ t = 1.80 sig. = .078
	Under above median ambiguity:	Adj. $R^2 = .52$ (d.f. = 1,19) F = 22.78*** $\beta = .74$					
	Under below median ambiguity:	Adj. $R^2 = .17$ (d.f. = 1,22) F = 5.83* $\beta = .46$	* = p<.05 ** = p<.01 *** = p<.001				

change was non-significant (Sig. of F change = 0.11), although this may be attributed to the smaller size of the below median sample. Expertise did appear to have moderating effects on this sample in the split sample analysis: the level of heuristic processing of participants reporting above median levels of expertise was not significantly associated with performance (sig. F = 0.486; d. f. = 1, 20), but was highly significantly associated for those reporting below median expertise (Adj. $R^2 = .24$, Sig. F = 0.010 (d. f. = 1, 21), F = 8.04, $\beta = 0.53$). Analysis of H2 suggests that novices utilize heuristic processing when they interpret novel data, consistent with theory. The moderating effect of expertise is less influential on systematic processing, however.

3.4. Task ambiguity

H4 and H5 investigate the extent that higher levels of task ambiguity increase the influence of systematic processing on performance. Using the strong form of moderation testing described above, we regressed performance onto perceived systematic processing, expertise and the interaction term. The interaction term was weakly significant (Sig. T = 0.08), and the overall model explained 47% of performance variance (F = 15.85, sig. F = 0.000, d. f. = 3, 47), 14% greater than the main effects model. This represents a significant increase in variance explained above the overall model (Sig. of Δ F = 0.003). Details are shown in Table 2 below. We then split the sample at the median to explore whether the directionality of this interaction effect is consistent with theory. For those participants reporting above median levels of task ambiguity, perceived systematic processing accounted for 52% of the variance in performance (F = 22.78, sig. F = 0.000, d. f. = 1,19; $\beta = .74$). Participants reporting below median task ambiguity had a much weaker association between systematic processing and performance (Adj. $R^2 = .17$, F = 5.83, sig. F = 0.025; d. f. = 1, 22; $\beta = .46$). Consistent with theory, task ambiguity does seem to moderate the relationship between systematic processing and performance. Also, the direct effect of task ambiguity is significantly negative (see H6 above), while the betas of the split sample analyses are positive. The effect of systematic processing in concert with ambiguity seems to offset the negative

influence of ambiguity, although this effect is much stronger for high levels of ambiguity than for low levels.

Investigating H5, we tested for a moderating effect of task ambiguity on heuristic processing. Contrary to theory, there was no indication of a moderation effect, using both the strong interaction-term analysis and the split sample. It seems that the direct effect of heuristic processing on performance is not apparent in the presence of task ambiguity.

In summary, HSM predicts a lack of main effects by the moderators under investigation – expertise and ambiguity here. This was found to be the case for expertise, but task ambiguity had a significant negative effect on the performance. This contradicts dual-process theory but is consistent with other research emphasizing the negative outcomes that can result from high ambiguity task contexts. In testing the moderation effects of expertise (H1 and H2), results were consistent with theory for heuristic processing. For systematic processing, results were weak but in a direction consistent with theory. Task ambiguity (H4 and H5) did have a significant moderating effect on the association between systematic processing and performance outcomes, consistent with theory. However, task ambiguity was not found to moderate the association between heuristic processing and performance, contrary to theory. Table 3 below summarizes these findings. Table 2 presents detailed results of the moderation analyses.

We then examined whether expertise and/or task ambiguity *mediated* the association between processing mode and performance outcomes, to fully explore our theoretical model. Using the technique prescribed in [23], we performed a series of regressions for each case. To identify a possible mediation effect of expertise (E) on the association between systematic processing (SP) and performance outcome (P), we performed the following regressions: $P = C_1 + T (SP) + e_1$; $P = C_2 + T_1 (SP) + T_2 (E) + e_2$; $P = C_3 + \beta (E) + e_3$ and $E = C_4 + \alpha (SP) + e_4$. The non-standardized coefficients T, T_1, α, and β are shown in Fig. 4a. As evident, T and T_1 are equal, ($T_1 - T$) is 0 and the product ($\alpha * \beta$) is –0.000099, nearly 0. These statistics indicate that expertise does not mediate the association between systematic processing and performance outcomes. Using a similar procedure (see Fig. 4b for coefficients), analyses show that expertise does not mediate the association between heuristic processing and performance outcomes either. Expertise hence

Table 3
Summary of Findings.

	Independent	Moderator	Dep = Score
	Systematic processing		Adj. R^2 = .35, sig. F = .000
	Subjective Quality		Adj. R^2 = .10, sig. F = .016
H3	Expertise		Adj. R^2 = .02, sig. F = .15
H6	Task ambiguity		Adj. R^2 = .096, sig. F = .015
H1	Systematic processing	Expertise	Directionality Confirmed
H2	Heuristic processing	Expertise	Directionality & interaction effect confirmed
H4	Systematic processing	Task Ambiguity	Directionality & interaction effect confirmed
H5	Heuristic processing	Task Ambiguity	Non-significant

has no mediating effect on either association. We then undertook the same series of regressions to investigate any potential mediating effects of ambiguity, Fig. 5a and b present the values of the coefficients from these set of regressions. As illustrated in Fig. 5a, task ambiguity does not mediate the association between systematic processing and performance outcomes. However, it does appear to partially and weakly mediate the association between heuristic processing and performance, consistent with theory.

4. Discussion

Assessment of information quality *in use* is problematic because information that is valuable and informative to one person may not be valuable and informative to the next, even when the information is objectively accurate and consistent. While other researchers have acknowledged the importance of understanding the subjective aspects of information quality assessment [36], none have proposed a theoretically-grounded approach for doing so. This exploratory research provides a theoretical model for understanding how characteristics of the user and the task – in this case expertise and ambiguity – affect how users process new information (quality metadata) in a quality assessment context. And because user expertise and task ambiguity are aspects of this context rather than invariant properties of information quality, this framework offers a means for understanding how the context affects the quality assessment process. For designing high quality information delivery systems, fitness-for-
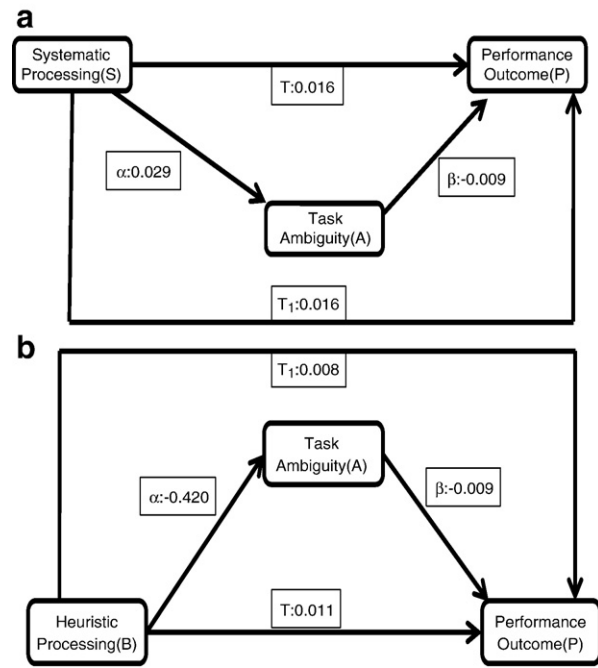
Fig. 5. a and b: Examining the mediation effect of Expertise.

use begs research, and the dual-process framework applied here offers a theoretically robust means of doing so. We are not suggesting that we should abandon our quest for high levels of information quality along traditional and objective quality dimensions. On the contrary, this research highlights the importance of integrating contextual and objective quality assessment processes. The contribution of the theoretical framework presented here is that it enables just such integration, simultaneously examining both contextual and objective quality attributes, reflecting how decision-makers actually process received information. The dual-process theories offer a clear mechanism for understanding when and why one type of processing is more likely to predominate in particular contexts. By using this referent theory, we can account for variations in quality assessments of identical information across decision makers and decision contexts.

For example, in this study, substantially more of the objective performance scores were explained by systematic processing for those with above median expertise than below median expertise. It seems that experts are more likely than novices to approach this task systematically, as predicted by theory. Further, since this task did not require prior expertise, and expertise did not directly increase performance (H3), findings suggest that experts are more motivated to process systematically in this context. It follows then that we may be able to improve the performance of novices by motivating them to systematically process quality metadata. An implication for managers is that they may be able to improve the performance of new hires by motivating them to process their decision tasks systematically. This presents an interesting avenue for future research, since this could be accomplished through training programs. Also as predicted by theory, participants with low levels of expertise have stronger associations between subjective quality assessments and objective performance than did those with high levels of expertise. Again, expertise does not directly affect performance, but the type of information processing that is undertaken does, and this in turn seems to vary with level of expertise. Interestingly, heuristic quality perceptions accounted for 15% of the variance in the objective outcome scores of these participants, regardless of expertise level. Clearly heuristic assessments of quality play a role in performance for this task, underscoring the need to better understand contextual aspects of information quality assessment.
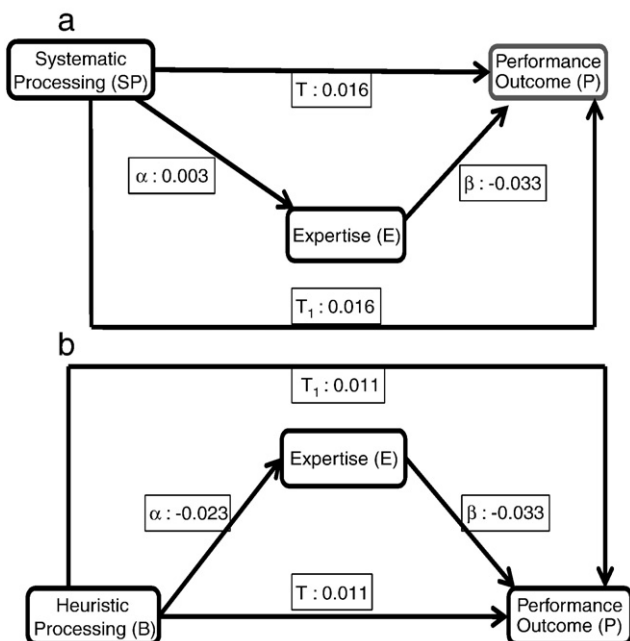
Fig. 4. a and b: Examining the mediation effect of Expertise.

In this study we found that it is not only the expertise of the decision-maker that affects information processing in this experimental task, but also how ambiguous he or she perceives the task to be. Importantly, we did not manipulate the task – all participants received the same instructions and task process. Therefore the effect of task ambiguity here is a perceptual phenomenon, yet findings indicate a strong interaction effect with systematic processing. Those that perceived the task as highly ambiguous had a much stronger association between systematic processing and outcome scores than did those that perceived the task to be less ambiguous. This is consistent with the dual-process theories, since ambiguity tends to motivate systematic processing. Since variations in task ambiguity are perceptual here, these findings demonstrate how perceptions of the task can affect information quality assessments even when task characteristics are not objectively different. For practitioners this suggests that there are important performance implications of task ambiguity perceptions; perceptions that managerial interventions may be able to alter. Counter-intuitively but consistent with theory, those that perceive the task to be highly ambiguous approached the problem more systematically, and had better performance, than those that viewed the task as less ambiguous. In exploratory mode, we investigated the possibility that expertise and ambiguity might mediate the associations between processing mode and performance outcomes. Were this to be the case, expertise and ambiguity would be serving to absorb the effects of processing mode, effects not accounted for by the dual-process theories. Analyses confirmed that the effects of expertise and ambiguity are primarily moderation and not mediation ones. Finally, and contrary to predictions based on the dual-process theories, we found a significant negative association between task ambiguity and performance. This conforms to standard theories of the negative effects of task ambiguity on performance. However, consistent with the dual-process theories, this negative impact seems to be overcome by systematic processing. The strength and direction of these effects is consistent with dual-process theory, but further research is needed to understand just how systematic processing works to offset the negative effects of perceived task ambiguity.

It is important to note that this first attempt at model validation was exploratory, and while it is generally consistent with theory, we acknowledge some limitations of this study. The use of a laboratory experiment allowed us to control for noise but reduced the realism of the task and the generalizability of the findings. Participants were all graduate students who were similar in age and were pursuing the same graduate degree in management. The decision making process investigated here focuses on the specific domain of marketing (although these students were Information Systems Masters students and so were not domain experts). Clearly these results can only be generalized to similar decision scenarios that can be mapped to this context. The test utilized a relatively small sample size and hence we were limited in our ability to apply advanced statistical analyses such as structural equation modeling. In this study we examined a limited model and more research is needed to understand all the potential interaction effects. For these reasons, we emphasize that the main contribution of this research is identification and application of appropriate theory for understanding subjective aspects of information quality assessment. Further validation of the framework in this context and other contexts is necessary to confirm these exploratory findings.

5. Conclusions

Overall, these findings suggest that characteristics of the decision-maker and the context play a significant role in the assessment of information quality. While research has acknowledged this [12,43], there has been a gap in theory that prevented adequate modeling of this, and hence appropriate integration with the standard data quality literature. By offering a theoretically-grounded model of this process, this research takes a first step towards closing this gap. We hope that integration of this theory with the standard information quality research will one day lead us to a deeper understanding of how contextual dimensions of quality interact with objective quality dimensions to affect information usage in decision tasks and decision outcomes.

Decision making is significantly affected by the quality of the information provided. Today's decision environments are characterized by high volumes of complex information, presenting a challenge to data managers charged with maintaining quality levels. Considerable advances have been made in improving objective aspects of quality, such as its accuracy and consistency, and also in supporting quality assessment with certain forms of metadata. However, users assess quality on the basis of how well the information suits their own needs, in the context of a particular task. This issue of the fitness of the information *for use in context* has not been addressed theoretically or empirically in the information quality literature. This research contributes by building a theoretical framework based on the dual-process theories of human cognition, specifically HSM here. The model enables investigation of the full process of information quality assessment, a process in which both systematic and heuristic information processing operate and affect each other. The research model suggests that the presence of contextual moderators – in this case expertise and task ambiguity – affect the balance of systematic and heuristic processing, which in turn affects performance outcomes. The model also suggests that users invoke subjective judgments in their quality assessment by applying heuristics about fitness-in-use. Using a cognitive approach, this research offers an explanation for ways that the decision context affects users' trade-offs between systematic and heuristic processing – tradeoffs that affect which information quality factors users attend to and utilize in their decision making. Further exploration of the model should help us understand ways to design decision support environments such that they take a variety of individual and contextual factors into account.

Appendix A. Perceptual measures

Hueristic Processing (from Bailey and Pearson's measure of information satisfaction [3])

To measure the use of heuristic data quality assessment, participants were asked to rate attributes of data quality as they were reflected the dataset. It included the following data quality attributes, measured as a separate 7-point Likert scale question:

Worthless – valuable
Uninformative – informative
Harmful – helpful
Useless – useful

These items were selected based on the pilot test indicating that these reflected the participants' use of these attributes to heuristically assess data quality rather than perform a systematic analysis of attributes of the data set.

Systematic Processing

Like heuristic processing, systematic processing is difficult to measure since it is cognitive, and most of the previous research in this area has increased levels of systematic processing by inducing it. We created the following items for reporting levels of systematic processing. They seek to assess how much effort participants had put into finding the optimal solution to the assigned task. We reasoned that the harder participants felt that they had worked, and the better they believed their solution to be, the more likely it was that they engaged in systematic processing. By contrast, use of heuristic processing tends to result in low confidence that an optimal solution was arrived at [24]. The pilot test indicated adequate convergent validity for the measure.

To what extent did you work hard to achieve an optimal solution to this exercise?
  (Very little.......to a great extent)
My solution to this exercise was an extremely effective one
  (Agree.......disagree)
The process I used to solve this exercise was a very efficient one
  (Agree.......disagree)

Expertise (from Stamm and Dube's measure of expertise [34]).

How knowledgeable are you about advertising planning?
  (not knowledgeable.......very knowledgeable)
To what extent are you an expert on the topic of allocating advertising resources?
  (agree.......disagree)
I am an expert in marketing
  (agree.......disagree)
I know very little about marketing
  (agree.......disagree)

Ambiguity (adapted from Daft and Macintosh's measure of task equivocality [8].

The task instructions were ambiguous
  (agree.......disagree)
When I received this task I was not sure what to do
  (agree.......disagree)
I was totally clear about what to do when I received this task
  (agree.......disagree)
I was very unsure about what I was supposed to do
  (agree.......disagree)

References

[1] T. Anderson, The Penalties of Poor Data, 2005 Whitepaper published by GoImmedia. com and the Data Warehousing Institute dw-institute.com, www.golmmedia.com/whitepapers/poordata.pdf.
[2] D. Arnott, G. Pervan, Eight Key Issues for the Decision Support Systems Discipline, Decision Support Systems 44 (3) (2007).
[3] J.E. Bailey, S.W. Pearson, Development of a Tool for Measuring and Analyzing Computer User Satisfaction, Management Science 29 (5) (1983).
[4] D. Ballou, R.Y. Wang, H. Pazer, G.K. Tayi, Modeling Information Manufacturing Systems to Determine Information Product Quality, Management Science 44 (4) (1998).
[5] M.D. Byrne, ACT-R/PM and Menu Selection: Applying a Cognitive Architecture to HCI, International Journal of Human Computer Studies 55 (1) (2001).
[6] S. Chaiken, A. Lieberman, A.H. Eagly, Heuristic and Systematic Information Processing Within and Beyond the Persuasion Context, in: J.S. Uleman, J.A. Bargh (Eds.), Unintended Thought, Guilford Press, New York, 1989.
[7] I. Chengalur-Smith, D.P. Ballou, H.L. Pazer, The Impact of Data Quality Information on Decision Making: An Exploratory Study, IEEE Transactions on Knowledge and Data Engineering 11 (6) (1999).
[8] R.L. Daft, N.B. Macintosh, A Tentative Exploration into the Amount and Equivocality of Information Processing in Organizational Work Units, Administrative Science Quarterly 26 (1981).
[9] J.J. Dijkstra, User Agreement with Incorrect Expert System Advice, Behavior & Information Technology 18 (6) (1999).
[10] M.J. Dutta-Bergman, The Impact of Completeness and Web Use Motivation on the Credibility of e-health Information, Journal of Communication 54 (2) (2004).
[11] A.H. Eagly, S. Chaiken, The Psychology of Attitudes, Hartcourt Brace College Publishers, New York, 1993.
[12] C.W. Fisher, I. Chengalur-Smith, D.P. Ballou, The Impact of Experience and Time on the Use of Data Quality Information in Decision Making, Information Systems Research 14 (2) (2003).
[13] B.J. Fogg, C. Soohoo, D.R. Danielson, L. Marable, J. Stanford, E.R. Tauber, How Do Users Evaluate Credibility of Web Sites? – A Study with over 2500 Participants, Conference on Designing for User Experiences, San Francisco, CA, 2003.
[14] M.F. Goodchild, R. Jeansoulin, Editorial in GeoInformatica 2 (3) (1998).
[15] T. Hong, Contributing Factors to the Use of Health-related Websites, Journal of Health Communication 11 (2) (2006).
[16] J.J. Illies, R. Reiter-Palmon, The Effects of Type and Level of Personal Involvement on Information Search and Problem Solving, Journal of Applied Social Psychology 34 (8) (2004).
[17] A. Janzone, J. Borzovs, An Approach to Geographical Data Quality Evaluation, Proceedings of the International Baltic Conference on Databases and Information Systems, July 2006.
[18] B.K. Kahn, D.M. Strong, R.Y. Wang, Data Quality Benchmarks: Product and Service Performance, Communications of the ACM 45 (4) (2002).
[19] D. Kahneman, Attention and Effort, Prentice Hall, Englewood Cliffs, N.J., 1973.
[20] Y.S. Kang, Y.J. Kim, Do Visitors' Interest Level and Perceived Quantity of Web Page Content Matter in Shaping the Attitude Toward a Web Site? Decision Support Systems 42 (2) (2006).
[21] Y.W. Lee, L.L. Pipino, J.D. Funk, R.Y. Wang, Journey to Data Quality, MIT Press, Cambridge, MA, 2006.
[22] S.H. Li, B.S. Lin, Assessing Information Sharing and Information Quality in Supply Chain Management, Decision Support Systems 42 (3) (2006).
[23] D.P. MacKinnon, G. Warsi, J.H. Dwyer, A Simulation Study of Mediated Effect Measures, Multivariate Behavioral Research 30 (1995).
[24] A. Parssian, S. Sarkar, V.S. Jacob, Assessing Data Quality for Information Products – Impact of Selection, Projection, and Cartesian Product, Management Science 50 (7) (2004).
[25] R.E. Petty, J.T. Cacioppo, Communication and Persuasion: Central and Peripheral Routes to Attitude Change, Springer-Verlag, New York, 1986.
[26] L.L. Pipino, Y.W. Lee, R.Y. Wang, Data Quality Assessment, Communications of the ACM 45 (4) (2002).
[27] S. Raghunathan, Impact of Information Quality and Decision-maker Quality on Decision Quality: A Theoretical Model and Simulation Analysis, Decision Support Systems 26 (4) (1999).
[28] T.C. Redman (Ed.), Data Quality for the Information Age, Artech House, Boston, MA, 1996.
[29] L.P. Robert, A.R. Dennis, Paradox of Richness: A Cognitive Model of Media Choice, IEEE Transactions on Professional Communication 48 (1) (2005).
[30] G. Shankaranarayanan, Y. Cai, Supporting Data Quality Management in Decision-Making, Decision Support Systems 42 (1) (2006).
[31] G. Shankaranarayanan, A. Even, The Metadata Enigma, Communications of the ACM 49 (2) (2006).
[32] G. Shankaranarayanan, S. Watts, A. Even, The Role of Process Metadata and Data Quality Perceptions in Decision Making: and Empirical Framework and Investigation, Journal of Information Technology Management 17 (1) (2006).
[33] G. Shankaranarayanan, M. Ziad, R.Y. Wang, Managing Data Quality in Dynamic Decision Environments: An Information Product Approach, Journal of Database Management 14 (4) (2003).
[34] K. Stamm, R. Dube, The Relationship of Attitudinal Components to Trust in Media, Communication Research 21 (1) (1994).
[35] F. Strack, R. Deutsch, The Two Sides of Social Behavior: Modern Classics and Overlooked Gems in the Interplay of Automatic and Controlled Processes, Psychological Inquiry 14 (2003).
[36] D.M. Strong, Y.W. Lee, R.Y. Wang, Data quality in context, Communications of the ACM 40 (5) (1997) 103–110.
[37] S.A. Sussman, W. Siegal, Informational Influence in Organizations: An Integrated Approach to Knowledge Adoption, Information Systems Research 14 (1) (2003).
[38] G.K. Tayi, D.P. Ballou, Examining Data Quality, Communications of the ACM 41 (2) (1998).
[39] J.B. Walther, Z.M. Wang, T. Loh, The Effect of Top-level Domains and Advertisements on Health Website Credibility, Journal of Medical Internet Research 6 (3) (2004).
[40] Y. Wand, R.Y. Wang, Anchoring Data Quality Dimensions in Ontological Foundations, Communications of the ACM 39 (11) (1996).
[41] R.Y. Wang, A Product Perspective on Data Quality Management, Communications of the ACM 41 (2) (1998).
[42] R.Y. Wang, V.C. Storey, C.P. Firth, A Framework for Analysis of Data Quality Research, IEEE Transactions on Data and Knowledge Engineering 7 (4) (1995).
[43] R.Y. Wang, D.M. Strong, Beyond Accuracy: What Data Quality Means to Data Users, Journal of Management Information Systems 12 (4) (1996).
[44] S. Watts, W. Zhang, Knowledge Adoption in Online Communities of Practice, Systemes d'Information et Management 1 (9) (2004).

Stephanie Watts is an Assistant Professor of Information Systems at the Boston University School of Management. She was previously on the faculty of the Weatherhead School at Case Western Reserve University. Her research focuses on the role that information technology plays in organizational knowledge transfer, with a focus on mediated knowledge sharing and cognition. She has published academic papers in such journals as Management Science, Information Systems Research, Organization Science, Journal of Strategic Information Systems, Information and Management, and Journal of Computer Mediated Communication. She has consulted to numerous organizations throughout North America.

G. Shankaranarayanan (Shankar) is a faculty member in the TOIM Division of Babson College. He obtained his Ph.D. in Management Information Systems from The University of Arizona, Eller School of Management. His research interests cover three primary areas: (1) modeling and managing data and metadata in information systems, (2) managing data quality for decision support, and (3) economic perspectives in data management. His research has appeared in journals including Journal of Database Management, Decision Support Systems, Communications of the ACM, Communications of the AIS, Journal of Information Technology Management, Journal of Computer Information Systems, International Journal of Information Quality, IEEE Transactions on Knowledge and Data Engineering, and the ACM Journal of Data and Information Quality. His research has won the best paper awards at the International Conference on Information Quality (ICIQ) and at the Workshop of Information Technology and Systems (WITS). He serves as an Area Editor of the International Journal of Information Quality and as an Associate Editor of the ACM Journal for Data and Information Quality.

Adir Even, a faculty member of the Department of Industrial Engineering and Management at Ben-Gurion University of the Negev, Israel, received his doctoral degree from Boston University School of Management in 2008. Adir explores the contribution of data resources to value-gain and profitability from both theoretical and practical perspectives, and studies implications for system design, data warehousing, business intelligence, and data quality management. His research has been published in journals such as IEEE/TDKE, CACM, CAIS, and DATABASE.