Antecedents of Information and System Quality: An Empirical Examination Within the Context of Data Warehousing

R. RYAN NELSON, PETER A. TODD, AND BARBARA H. WIXOM

R. RYAN NELSON is a Professor and the Director of the Center for the Management of Information Technology at the McIntire School of Commerce at the University of Virginia. His research interests are in the areas of technology adoption and innovation, information technology human capital, and project management. His research has been published in such journals as *Journal of Management Information Systems, MIS Quarterly,* and *Communications of the ACM.*

PETER A. TODD is the Chesapeake and Potomac Professor and Associate Dean for Graduate Programs at the McIntire School of Commerce at the University of Virginia. His research interests are in the areas of technology adoption and innovation, the role of IT in decision-making and human–computer interaction. He has published in a variety of journals including *MIS Quarterly* and *Information Systems Research.* He currently serves as a senior editor for *MIS Quarterly.*

BARBARA H. WIXOM is an Associate Professor of Commerce at the McIntire School of Commerce at the University of Virginia. She is a Fellow of the Data Warehousing Institute for her data warehousing research, Associate Editor for the *Journal of Business Intelligence,* and a two-time winner of the Society for Information Management Paper Competition. Professor Wixom teaches graduate and undergraduate classes in data warehousing, data management, and IT strategy, and her research focuses on data warehousing benefits and use. She has published in journals that include *Journal of Management Information Systems, MIS Quarterly,* and *Information Systems Research.*

ABSTRACT: Understanding the successful adoption of information technology is largely based upon understanding the linkages among quality, satisfaction, and usage. Although the satisfaction and usage constructs have been well studied in the information systems literature, there has been only limited attention to information and system quality over the past decade. To address this shortcoming, we developed a model consisting of nine fundamental determinants of quality in an information technology context, four under the rubric of information quality (the *output* of an information system) and five that describe system quality (the *information processing system* required to produce the output). We then empirically examined the aptness of our model using a sample of 465 data warehouse users from seven different organizations that employed report-based, query-based, and analytical business intelligence tools. The results suggest that our determinants are indeed predictive of overall information and system quality in data warehouse environments, and that our model strikes a balance between comprehensiveness and parsimony. We conclude with a discussion of the implications for both theory and the development and implementation of information technology applications in practice.

Journal of Management Information Systems / Spring 2005, Vol. 21, No. 4, pp. 199–235.

0742–1222 / 2005 $9.50 + 0.00.

KEY WORDS AND PHRASES: business intelligence software, data warehousing, information quality, information systems success, system quality.

QUALITY HAS EVOLVED INTO A CORE BUSINESS CONCEPT with multidisciplinary applications and dramatic implications for business value. When manufacturing firms were forced to come to terms with the quality challenge of the early 1980s, the total quality management (TQM) movement had a profound effect on product development [23, 25, 38]. Since that time, other disciplines, such as marketing and human resource management, have engaged in quality pursuits; the former citing evidence that the quality of customer service is often as important as the quality of the product [87], and the latter recognizing quality of work life as a key driver of employee retention [61]. In addition, overall measures of quality, such as those captured in the Baldridge Awards and Balanced Scorecard practices, have proliferated[1] [57]. Some researchers assert that quality of products and services is the single most important determinant of a business' long-term success [3, 15].

Despite increasing attention to the quality construct in the broader business literature, attention to information and system quality has become less central in recent years. Instead, in an effort to understand users' reactions to information technology (IT), researchers have focused on perceptions related to IT use, predominantly ease of use and usefulness, along with other related factors (e.g., [21, 77]). Although such perceptions have been important in explaining IT usage, they are relatively abstract and, as a result, provide limited guidance for system designers [31, 75]. Orlikowski and Iacono [59] have noted that such IT research, which employs a "proxy view" of technology, has lost its connection to the field's core subject matter—the IT artifact itself. We believe that identifying the dimensions of the IT artifact that shape quality can provide this connection.

Thus, the primary purpose of this research is to identify a set of antecedents that simultaneously define the nature of the IT artifact and drive information and system quality. Building on the findings reported in Wixom and Todd [84], we will empirically test the suitability of these determinants as aids in the prediction and understanding of quality within an IT context. A second research objective is to explore the area of data warehousing, specifically three popular business intelligence applications—predefined reports, ad hoc queries, and analytical tools. We hope to provide IT managers with a better understanding of these contemporary tools to help them create IT infrastructures that effectively support organizational decision-making.

The Quality Construct

THERE ARE MULTIPLE PERSPECTIVES on quality in the business literature. In a comprehensive review, Reeves and Bednar [65] identify four dominant views of quality:

quality as excellence, quality as value, quality as conformance with specifications, and quality as meeting expectations. The excellence view suggests that quality is assessed on some absolute standard. The value perspective refines that notion to suggest that the standards of excellence need to be assessed relative to the costs of achieving them. The conformance view further systematizes these ideas to suggest that quality be assessed in terms of a consistent and quantifiable delivery of value relative to a specific design ideal. Finally, the notion of quality as meeting expectations suggests that quality is defined by conformance to customer expectations that may relate to excellence, value, and other attributes that are salient to consumers in shaping their perceptions of quality.

Reeves and Bednar [65] note that quality assessments relative to expectations represent the most pervasive perspective on quality, with the critical exemplar being service quality. Zeithaml et al. [88] define service quality as the degree to which a service exceeds customer expectations. Further, they empirically identify a set of service attributes that collectively determine customer expectations about service quality; these service attributes include responsiveness, reliability, assurance, tangibility, and empathy. Consistent with the notion that salient beliefs about objects and behaviors shape broader attitudes [26], these five factors have been empirically tested across a variety of settings to establish their overall utility in shaping service quality.

In the information systems (IS) literature, quality has been a frequently referenced, but relatively ill-defined, construct (e.g., [4, 8, 41, 51, 71]). Furthermore, with the exception of IT service quality (e.g., [37, 40, 63]), the study of quality as a key dependent variable has been largely supplanted by usage in the IS literature. The following section provides a review of relevant literature while developing the theoretical context for the quality of IS.

The Theoretical Context for Quality

Some IT frameworks have been created to place quality into a broader theoretical context. Building on concepts from Shannon and Weaver [69] and Mason [52], DeLone and McLean [22] identify information and system quality as the key initial antecedents for IS success. Extending these notions, Seddon [67] developed a respecified model of IS success, which shows that information quality and system quality jointly influence perceptual measures of system benefit, represented by perceived usefulness and user satisfaction (which Seddon [67] defines as satisfaction with use). These, in turn, influence expectations about the benefits of future use, and subsequently, actual usage of IT, which can have a series of positive or negative organizational consequences [46, 50].

Related efforts have focused on empirically assessing the role of information and system quality as antecedents of satisfaction and usage in a variety of settings [64, 68, 70, 84, 85]. In general, such studies treat quality at a holistic level. However, it is clear that quality constructs are multidimensional [42, 64, 67]. More generally, Goodhue [31] notes that user evaluations of IS attributes can provide a basis for the determination of IS value.

One critical issue is determining what constitutes a "good" set of IS dimensions. We offer four key goals for the set of determinants that shape quality. Collectively, the dimensions should:

1. be complete (in the sense of explaining overall information and system quality);
2. be relatively parsimonious;
3. enhance understanding of the multifaceted nature of information and system quality; and
4. be actionable, in the sense that the dimensions can be influenced through system design or managerial intervention.

Using these four goals as a guide, we will turn to a derivation of the dimensions of information and system quality and to the integration of those into a model of quality.

Information Quality

Researchers have introduced a variety of definitions for information (or data) quality. In general, the definitions take either an intrinsic or a contextual view of information quality. The intrinsic view considers the properties of information largely in isolation from a specific user, task, or application. Thus, the intrinsic view reflects a measure of agreement between the data values presented by an IS and the actual values the data represents in the real world [47, 60], the degree to which data values are *not* inaccurate, outdated, and inconsistent [48], and the accuracy of information generated by an IS [31, 67, 82]. Although this is an important perspective, it is somewhat limited because it treats information as an object that can be assessed in isolation of the context to which it is applied. Thus, intrinsic quality is a necessary, but not sufficient, condition to determine information quality.

A context-based view extends the notion of information quality, suggesting that it needs to be defined relative to the user of the information, the task being completed, and the application being employed [47, 60]. From this perspective, information quality is assessed by the degree to which it is helpful in completing a particular task [27, 45, 62, 72, 74, 81, 82]. For example, this might be assessed abstractly in terms of the usefulness of the information in aiding decision-making. The context view expands the dimensions of information quality beyond accuracy to include dimensions such as relevance, completeness, and currency of the information that shape perceptions of quality in the context of use [82].

In addition to intrinsic and context-based dimensions of information quality, Wang and Strong [82] also suggest that there is a representational dimension. The role of format in information processing and decision-making has long been a topic of study in IS research (e.g., [11, 36, 76, 79, 83]). The representational dimension reflects the degree to which information presentation effectively facilitates interpretation and understanding; therefore, the format of the information is an important dimension of information quality [64].

Collectively, there are myriad dimensions that can be considered under the label of intrinsic, contextual, and representational information quality, and there is little consensus on what constitutes a complete and yet parsimonious set of information quality dimensions [80]. Building on the categorization of intrinsic, contextual, and representational dimensions provided by Wang and Strong [82], we have distilled a core set of information quality dimensions as follows: accuracy (reflecting intrinsic quality), completeness and currency (reflecting contextual quality), and format (reflecting representational quality). The dimensions, their derivation, and treatment in prior literature are shown in Appendix A.

Accuracy is most commonly defined as the correctness in the mapping of stored information to the appropriate state in the real world that the information represents [5, 27, 43]. Wand and Wang [80] further refine the notion of accuracy to include the idea that the information not only is correct, unambiguous, and objective, but also meaningful and believable. The key element of this refinement is the notion that there is an important perceptual component to accuracy. Information not only must be accurate but must also be perceived to be accurate [82]. A further extension to the notion of accuracy is consistency [5, 27, 34, 41], referring to the correctness of the relationship *between* or *among* multiple items of information and of information over time. In judging accuracy, we would assert that users assess perceptions of correctness of information extracted from systems over a protracted period of time. Their overall sense of accuracy may be shaped by the underlying correctness of the information, perceptions of the believability of the information, and the consistency of longitudinal experiences.

Beyond accuracy, the quality of information also can be shaped by *completeness.* Completeness refers to the degree to which all possible states relevant to the user population are represented in the stored information [5, 27, 34, 80]. It is important to recognize that the assessment of completeness only can be made relative to the contextual demands of the user and that the system may be complete as far as one user is concerned, but incomplete in the eyes of another. While completeness is a design objective, its assessment is based on the collective experience and perceptions of the system users.

In addition to completeness, *currency* has been identified as an important factor in contextual information quality [4, 8, 16, 35, 54]. Currency refers to the degree to which information is up to date, or the degree to which the information precisely reflects the *current* state of the world that it represents.[2] Currency is a contextual attribute of system quality to the extent that its assessment is dependent on task and user perceptions [6]. Users may have different demands for currency and, as a consequence, information that is viewed as current for one task may be viewed as too dated for another. Again, user perceptions of currency relative to the task demands over time will be an important determinant of information quality.

The final dimension of information quality captured in Table 1 is *format.* Format is tied to the notion of representational quality [4, 47, 53, 82]. Format refers to the degree to which information is presented in a manner that is understandable and interpretable to the user, and thus aids in the completion of a task. There is significant research on information presentation, and the one consistent conclusion from this line of research is that the suitability of a particular presentation is highly contingent

Table 1. Information Quality Dimensions

Dimension	Definition	Information quality category
Accuracy	The degree to which information is correct, unambiguous, meaningful, believable, and consistent.	Intrinsic
Completeness	The degree to which all possible states relevant to the user population are represented in the stored information.	Extrinsic; contextual
Currency	The degree to which information is up-to-date, or the degree to which the information precisely reflects the current state of the world that it represents.	
Format	The degree to which information is presented in a manner that is understandable and interpretable to the user and thus aids in the completion of a task.	Extrinsic; representational

Note: The dimensions, their derivation, and treatment in prior literature are shown in Appendix A.

on the manner in which the presentation matches the demands of the task and the mental model employed by the user [78]. Thus, the assessment of format will be shaped by the perceptions of the user completing different tasks with the system over time.

To summarize, a number of factors have been identified and labeled as dimensions of information (data) quality as illustrated in Appendix A; however, an integration of the literature based on Wang and Strong's [82] organizing framework suggests that these factors can be reduced to a relatively concise set of determinants of information quality (see Table 1). The dimensions we identify are accuracy, completeness, currency, and format. Collectively, these four dimensions appear to capture the key elements of information quality by taking into account the intrinsic properties of information quality related to correctness, the contextual factors driving task performance, as well as the representational characteristics of information quality. For each dimension, it is important to recognize that quality is not assessed in an absolute and objective sense, but rather that the assessment of quality is tied to the perceptions of information consumers who are working on specific tasks within specific contexts.

Applying our four-model assessment criteria, we conclude that these dimensions enhance understanding of the multifaceted nature of information quality by capturing dimensions related to the intrinsic, extrinsic, and representational views of information quality and by emphasizing the importance of context and perception in the overall quality assessment. Further, we propose that these dimensions are actionable in that they can guide a designer to refine specific facets of a system in an effort to enhance quality, and the dimensions are relatively parsimonious, distilling over 30 dimensions represented in the literature into four key constructs related to quality.

What remains to be determined is the extent to which this explanation is complete in the sense of explaining the variation in a global assessment of information quality. Prior to describing the empirical study that will provide this assessment, we turn our attention to an assessment of the dimensions of system quality.

System Quality

Overall, system quality has received less formal treatment than information quality in the IS literature. In addition, elements of system quality often are intermingled with dimensions that are closely related to service quality and ease of use. For example, Bailey and Pearson [4] include a variety of system dimensions that relate to IS services in their studies of user satisfaction. Rai et al. [64] equate system quality with operational measures of ease of use. Although these constructs clearly are related, they are not the same. A system that is perceived to be easy to use may also be perceived to be high quality; therefore, ease of use may be a consequence of system quality. Similarly, systems bundled with a high level of IT service may be viewed to be of higher quality, making service quality a covariate to the quality of the system.

Such interrelationships are the cornerstone of system success models (e.g., [67]). These interrelationships make it all the more important to ensure conceptual clarity in the specification and distinction of constructs. In this regard, we would suggest that there are unique dimensions that act as antecedents to system quality that are distinct from either ease of use or service factors.

To the extent that information quality is related most closely to the *output* of an IS, system quality reflects the *information processing system* required to produce that output. Thus, the dimensions of system quality represent user perceptions of interaction with the system over time. In this sense, higher-quality systems should be perceived as easier to use and, ultimately, have higher levels of usefulness and use [21].

System interaction typically occurs within an organizational context with the goal of completing a particular task; therefore, it is useful to consider dimensions of system quality using a spectrum that ranges from system to task. System dimensions are those characteristics of a system that are largely invariant across different uses and can be assessed independent of task, context, or application. Task dimensions are those for which an assessment will depend on the task and setting. Drawing on over 20 studies that define dimensions of system quality, our assessment of the literature (summarized in Table 2 and detailed in Appendix B) suggests that there are five key dimensions to system quality: accessibility, reliability, flexibility, response time, and integration. Accessibility and reliability are, to a large extent, system dimensions. They represent defined properties that are largely independent of usage. Response time, flexibility, and integration are characteristics that are perhaps best evaluated in the context of specific tasks and should be considered task-related. Below, we explore each in turn.

Accessibility represents the degree to which a system and the information it contains can be accessed with relatively low effort [4, 53, 54, 82]. Access to information can be viewed as a necessary condition for system quality. It is a system property to

Table 2. System Quality Dimensions

Dimension	Definition	System quality category
Accessibility	The degree to which a system and the information it contains can be accessed with relatively low effort.	System-related
Reliability	The degree to which a system is dependable (e.g., technically available) over time.	
Response time	The degree to which a system offers quick (or timely) responses to requests for information or action.	Task-related
Flexibility	The degree to which a system can adapt to a variety of user needs and to changing conditions.	
Integration	The degree to which a system facilitates the combination of information from various sources to support business decisions.	

Note: The dimensions, their derivation, and treatment in prior literature are shown in Appendix B.

the extent that the system itself is either accessible to a user or not accessible, regardless of the task that the user is trying to accomplish.

Reliability refers to the dependability of a system over time [10, 71, 73]. It can be defined objectively as the technical availability of the system and can be concretely measured by metrics such as uptime, downtime, or mean time between failures. Despite the fact that reliability can be measured objectively, it also is true that individuals may have perceptions of reliability that are independent of measured reliability. Consider a user who only works with a system once a week for a short period of time. A moment of downtime during that time may have a significant detrimental effect on reliability. Thus, user perceptions of reliability are key to determining system quality.

Response time refers to the degree to which a system offers quick (or timely) responses to requests for information or action [4, 20, 24, 35]. Different kinds of systems (e.g., transaction processing, decision support) often are designed or optimized to provide certain response times, and users may perceive the response time of a system based on the kind of task that they are performing. For example, users may be very tolerant of long response times for an Internet application, but they would be much less tolerant of a similar response time in a desktop application. To the extent that this is the case, we argue that response time is a task-related property of a system, and one in which user perceptions may vary from objective measures. While the two should be related in most cases, ultimately it is the perceptions, not the objective measures, that will guide perceptions of quality and usage behavior.

Flexibility relates to the degree to which a system can adapt to a variety of user needs and to changing conditions [4, 33, 54, 82]. The definition of flexibility suggests the need to adapt to changing conditions and different user needs, making it a task property of system quality. To the extent that a system will be used over time and must provide information as input to a wide variety of decision tasks, flexibility can be expected to be a key determinant of quality. The relative importance of flexibility in determining quality may depend on the degree to which task demands change over time. In a data warehouse context, for example, we might expect that flexibility is less important in the context of predefined reports (which provide information for static tasks) and more important for querying and analysis, which are less structured and more likely to change over time.

Finally, *integration* refers to the degree to which a system facilitates the combination of information from various sources to support business decisions [4, 53, 82]. The need for integration will vary across tasks and contexts, and thus, integration represents a task-related property. Tasks that are more interdependent will require systems that facilitate integration to a greater degree than systems that support largely independent tasks [32].

Determinants of Information and System Quality

When considering information and system quality together, it is useful to think of information as the product of a system and the system as the information processing system that produces the information [22]. As noted above, the key dimensions of information quality are *accuracy, completeness, currency,* and *format.* The key determinants of system quality are *accessibility, reliability, response time, flexibility,* and *integration.* Collectively, these determinants should explain information and system quality, and they indirectly should influence user perceptions about satisfaction with the information and system (see Figure 1).

As explained earlier, the literature suggests that system factors may influence a user's perception of, or satisfaction with, the information provided by the system [12]. Moreover, past confusion in differentiating system quality from information quality factors (see Appendices A and B) suggests that crossover or interaction effects may exist between the two constructs. Therefore, the research model includes crossover relationships from quality (information and system) to satisfaction (system and information) as well as an interaction effect of information and system quality on information satisfaction and system satisfaction (see Figure 1). These relationships explore the possibility that more complex quality/satisfaction relationships may exist.

Empirical Study

A CROSS-SECTIONAL SURVEY WAS CONDUCTED to test the model in Figure 1. The context of the survey was user experiences with a data warehouse. Specifically, survey participants were asked to report on their experiences with three types of business intelligence tools most commonly employed to access and analyze data warehouse information: (1) predefined reporting software, (2) query tools, and (3) analysis tools. Predefined reporting is software that is set up by the data warehouse project team and is run by users on a regular basis to provide predetermined information. Query tools allow users to extract information for themselves to satisfy unplanned, nonroutine information needs. Analytical tools allow the manipulation and modeling of information

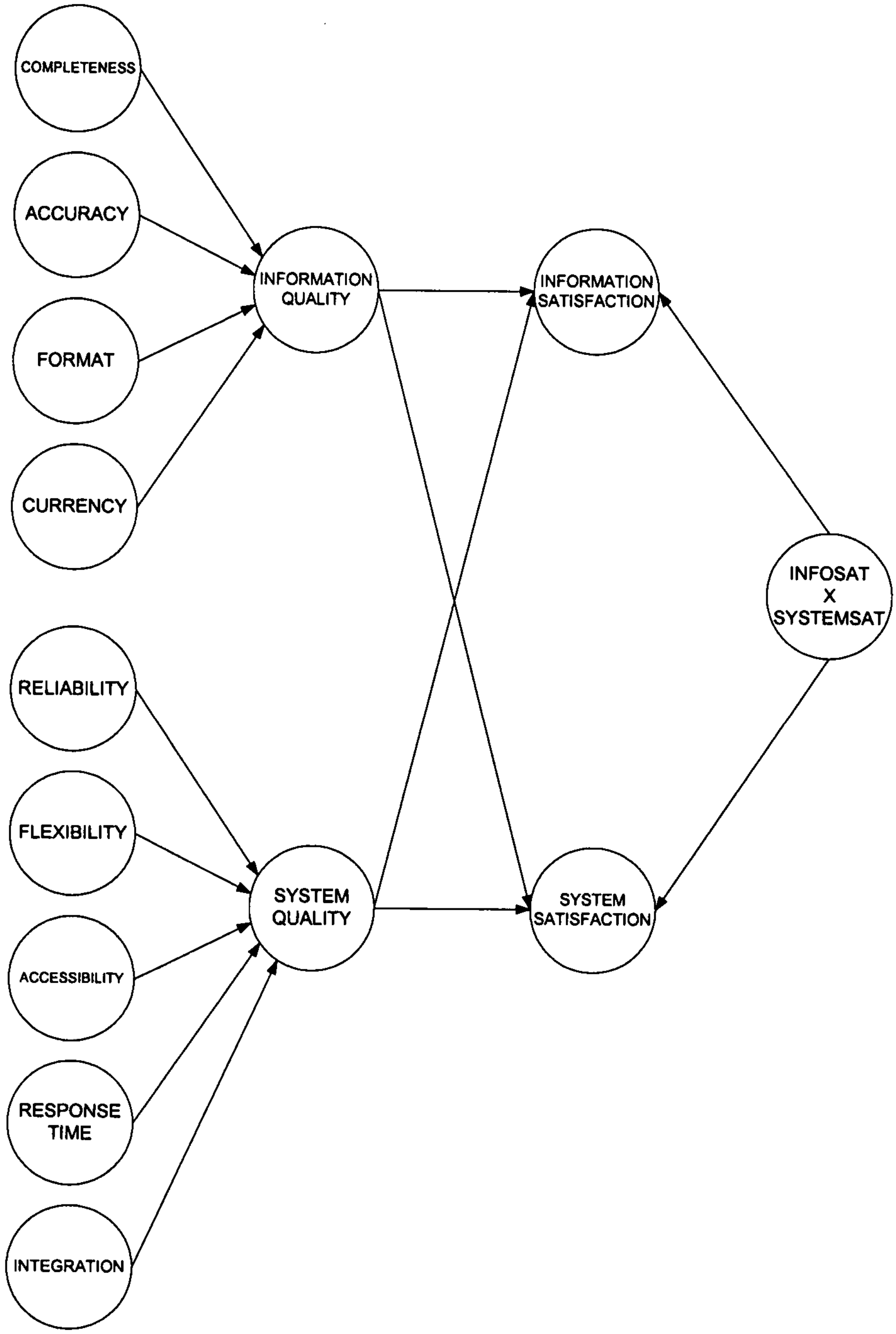

Figure 1. Determinants of Information and System Quality

extracted from a data warehouse. The following sections describe the study, including the instrument development process and the sample that was used.

Instrument Development

Development of the survey instrument followed the process proposed by Moore and Benbasat [55]. A literature review was conducted to locate past operational measures of the constructs under investigation, groups of questions were compiled from validated instruments to represent each construct, and wording was modified to fit the data warehouse context to be studied. Next, ten professors and graduate students sorted the items into separate categories, identifying ambiguous or poorly worded items. Items were removed and minor wording changes were made prior to a second round of sorting, which did not uncover further problems. The three items that were categorized most accurately were selected for each quality dimension and included on the survey instrument in a random order.[3] Each question was measured on a seven-point Likert-type scale, ranging from (1) strongly disagree to (7) strongly agree.

Before implementing the survey, the instrument was further reviewed by academics and practitioners with knowledge of survey design, IS quality, and data warehousing. Minor changes were made based on their suggestions. The resulting survey was then pilot-tested using respondents from a large public university to identify problems with the instrument's wording, content, format, and procedures. For this pilot test, surveys were distributed to 250 active users of a large public university's data warehouse; 73 responded, resulting in a 29 percent response rate. Pilot participants completed the instruments and provided written comments about length, wording, and instructions. Two of the participants were interviewed to gain a richer understanding of the feedback. The data were analyzed regarding the internal consistency reliability of the constructs using Cronbach's alpha, and each exceeded the accepted 0.7 level of reliability [58].

Based on the results of the pilot sample, minor modifications were made to the survey design. The final survey included items measuring the constructs from Figure 1 as well as a series of demographic and self-reported usage items. The specified items and descriptive statistics, organized by construct, are shown in Table 3.

Sample

Study participants were solicited via an e-mail announcement sent to members of the Data Warehousing Institute offering a free study to assess the success of their organization's data warehouse software. Seven organizations from a variety of industries (e.g., health care, consumer goods, financial services, and government) agreed to participate. Each organization was asked to distribute paper-based surveys via interoffice mail to all of the active users of its data warehouse. All surveys were confidential; no identifying personal information was collected. Completed surveys were collected by a contact person at each organization and returned to the researchers. Response rates varied across organizations (see Table 4), with an overall study response rate of 21 percent, yielding 465 completed surveys.

Table 3. Constructs and Descriptive Statistics

	Predefined reports		Query		Analysis	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Information quality						
Completeness	$\alpha = 0.90$ Fornell = 0.94		$\alpha = 0.88$ Fornell = 0.93		$\alpha = .89$ Fornell = 0.93	
____ provides me with a complete set of information.	4.58	1.77	4.78	1.70	4.65	1.61
____ produces comprehensive information.	4.88	1.70	5.08	1.62	4.96	1.57
____ provides me with all the information I need.	4.15	1.87	4.39	1.87	4.32	1.72
Accuracy	$\alpha = 0.87$ Fornell = 0.90		$\alpha = 0.85$ Fornell = 0.89		$\alpha = 0.83$ Fornell = 0.88	
____ produces correct information.	5.14	1.60	5.16	1.51	5.10	1.42
There are few errors in the information I obtain from ____.	4.75	1.78	4.79	1.67	4.71	1.55
The information provided by ____ is accurate.	5.04	1.66	5.09	1.54	5.02	1.45
Format	$\alpha = 0.89$ Fornell = 0.92		$\alpha = 0.88$ Fornell = 0.91		$\alpha = 0.87$ Fornell = 0.90	
The information provided by ____ is well formatted.	4.93	1.68	4.61	1.68	4.63	1.61
The information provided by ____ is well laid out.	5.10	1.57	4.86	1.63	4.81	1.50
The information provided by ____ is clearly presented on the screen.	5.23	1.55	5.08	1.53	5.04	1.44
Currency	$\alpha = 0.93$ Fornell = 0.94		$\alpha = 0.93$ Fornell = 0.94		$\alpha = 0.93$ Fornell = 0.94	
____ provides me with the most recent information.	5.05	1.87	5.11	1.90	4.98	1.76
____ produces the most current information.	4.96	1.79	4.95	1.78	4.93	1.63
The information from ____ is always up to date.	4.71	1.77	4.70	1.78	4.64	1.66

Information quality	α = .94 Fornell = 0.94		α = 0.91 Fornell = 0.93		α = 0.90 Fornell = 0.92	
Overall, I would give the information from ____ high marks.	5.09	1.68	5.09	1.68	5.00	1.59
Overall, I would give the information provided by ____ a high rating in terms of quality.	5.10	1.63	5.13	1.56	4.97	1.52
In general, ____ provides me with high-quality information.	5.11	1.61	5.15	1.56	5.02	1.51
Information satisfaction	α = 0.93 Fornell = 0.96		α = 0.91 Fornell = 0.96		α = 0.90 Fornell = 0.95	
Overall, the information I get from ____ is very satisfying.	4.89	1.80	5.02	1.71	4.87	1.68
I am very satisfied with the information I receive from ____.	4.84	1.78	5.02	1.75	4.92	1.69
System quality						
Reliability	α = 0.90 Fornell = 0.93		α = 0.89 Fornell = 0.92		α = 0.90 Fornell = 0.92	
____ operates reliably.	5.10	1.73	4.91	1.65	4.94	1.60
____ performs reliably.	5.15	1.66	5.02	1.64	4.97	1.58
The operation of ____ is dependable.	5.10	1.56	4.96	1.60	4.87	1.59
Flexibility	α = 0.86 Fornell = 0.90		α = 0.86 Fornell = 0.90		α = 0.85 Fornell = 0.89	
____ can be adapted to meet a variety of needs.	4.28	1.99	5.19	1.71	5.07	1.66
____ can flexibly adjust to new demands or conditions.	3.73	1.86	4.81	1.75	4.63	1.70
____ is versatile in addressing needs as they arise.	4.00	1.83	4.91	1.67	4.80	1.57
Integration	α = 0.89 Fornell = 0.91		α = 0.88 Fornell = 0.92		α = 0.90 Fornell = 0.92	
____ effectively integrates data from different areas of the company.	4.78	1.89	4.75	1.88	4.62	1.84
____ pulls together information that used to come from different places in the company.	5.14	1.77	5.09	1.75	4.83	1.75
____ effectively combines data from different areas of the company.	4.93	1.77	4.94	1.73	4.76	1.72

(continues)

Table 3. Continued

	Predefined reports		Query		Analysis	
	Mean	Std. dev.	Mean	Std. dev.	Mean	Std. dev.
Accessibility	$\alpha = 0.90$ Fornell = 0.92		$\alpha = 0.89$ Fornell = 0.92		$\alpha = 0.89$ Fornell = 0.92	
____ allows information to be readily accessible to me.	5.27	1.70	5.18	1.69	5.03	1.65
____ makes information very accessible.	5.16	1.69	5.12	1.70	4.97	1.65
____ makes information easy to access.	5.14	1.70	4.79	1.78	4.73	1.68
Response time	$\alpha = 0.80$ Fornell = 0.87		$\alpha = 0.79$ Fornell = 0.87		$\alpha = 0.78$ Fornell = 0.86	
It takes too long for ____ to respond to my requests. (RC)	4.26	1.90	4.12	1.86	4.14	1.74
____ provides information in a timely fashion.	5.07	1.67	4.93	1.73	4.80	1.63
____ returns answers to my requests quickly.	4.90	1.72	4.73	1.72	4.66	1.63
System quality	$\alpha = 0.91$ Fornell = 0.94		$\alpha = 0.90$ Fornell = 0.93		$\alpha = 0.91$ Fornell = 0.93	
In terms of system quality, I would rate ____ highly.	4.91	1.69	4.88	1.66	4.84	1.60
Overall, ____ is of high quality.	5.12	1.55	5.08	1.53	5.04	1.47
Overall, I would give the quality of ____ a high rating.	4.97	1.62	5.04	1.59	4.92	1.57
System satisfaction	$\alpha = 0.92$ Fornell = 0.95		$\alpha = 0.91$ Fornell = 0.95		$\alpha = 0.92$ Fornell = 0.95	
All things considered, I am very satisfied with ____.	4.61	1.94	4.49	1.92	4.44	1.89
Overall, my interaction with ____ is very satisfying.	4.65	1.82	4.52	1.84	4.47	1.80

RC = reverse coded.

Table 4. Participating Firms

Company	Surveys sent*	Surveys returned	Response rate (percent)
A. Health care	129	40	31
B. Packaged goods	300	92	31
C. Financial services	179	23	13
D. Health care	108	42	39
E. Public sector	1,200	172	14
F. Public sector	231	61	26
G. Public sector	66	35	53
Overall	2,213	465	21

* Number of surveys sent to each company. We cannot be certain that all surveys sent were distributed to data warehouse users. Thus, our effective response rate is likely somewhat higher than reported here.

The average age of the respondents was 42 years, and 40 percent were male. The respondents had an average of 12 years tenure with their organization and 18 years average total work experience. Their positions in the organizations varied from clerical to senior management, with 58 percent being analysts, and they represented different functional areas across the organization. The demographic profile of the sample is shown in Table 5.

All respondents used at least one of the three kinds of business intelligence tools, and they represent all levels of both relative and absolute use of the software. The histograms in Figure 2 show the number of respondents who reported their relative and absolute use of each tool across the seven possible levels of usage (i.e., 1 = low use; 7 = high use). Relative use followed an approximately normal distribution, although there were many infrequent users of the analysis and predefined reporting tools. Absolute use followed more of a uniform distribution and had a large number of respondents reporting low levels of usage for all three tools. Collectively, the varied usage levels suggest a reasonable variance in user interactions and experiences with the business intelligence technology.

Results

THE RESEARCH MODEL WAS TESTED using partial least squares (PLS), a structural modeling technique that is well suited for assessing complex predictive models [7, 18, 86]. PLS concurrently tests the psychometric properties of the scales used to measure the variables in the model (i.e., the measurement model) and analyzes the strength and direction of the relationships among the variables (i.e., the structural model) [49]. PLS Graph version 2.91 [19] was used for the analysis, and the bootstrap resampling method (100 resamples) used to determine the significance of the paths within the structural model.

Table 5. Study Participants

	Number	Percent
Organizational level		
Senior management	13	3
Middle management	95	22
First-level supervisor	48	11
Analyst	257	58
Clerical	27	6
Functional area		
Accounting	22	5
Finance	79	17
Human resources	22	5
Information systems	37	8
Marketing and sales	82	18
R&D	96	21
Other	116	25
Gender		
Male	180	40
Female	270	60
Average age: 42 years		
Average years at company: 12 years		
Average years in workforce: 18 years		

Measurement Model

The test of the measurement model included the estimation of internal consistency and the convergent and discriminant validity of the instrument items for each of the three technologies (predefined reports, query tools, and analysis tools). Table 3 lists the survey scales and their internal consistency reliabilities for responses across each of the three technologies. All reliability measures were well above the recommended level of 0.70, indicating adequate internal consistency [58].

These items also demonstrated satisfactory convergent and discriminant validity. Convergent validity is adequate when constructs have an average variance extracted (AVE) of at least 0.5 [28]. For satisfactory discriminant validity, the AVE from the construct should be greater than the variance shared between the construct and other constructs in the model [18]. This was true for each of our constructs (see Appendix C). Convergent validity is also demonstrated when items load above 0.50 on their associated factors; all of the measures have significant loadings above the suggested threshold (see Appendix D).

Structural Model

The path coefficients and explained variance for the structural models are shown in Figure 3. Overall, the four determinants of information quality explained over 75

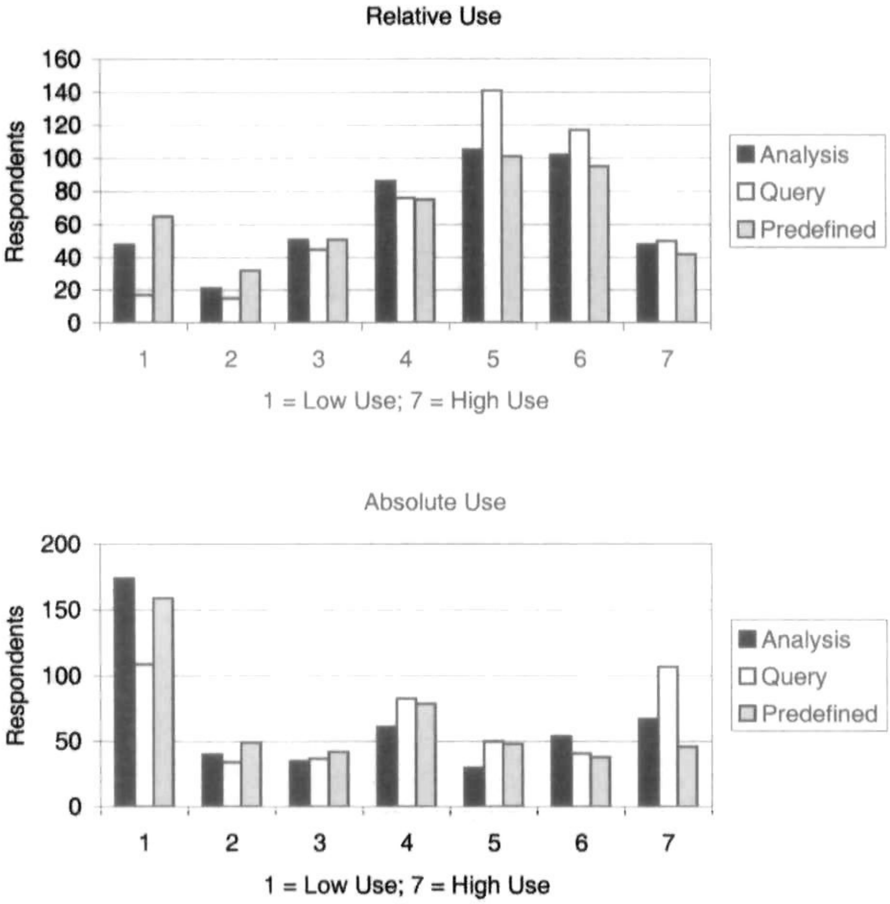

Figure 2. Sample Business Intelligence Tool Usage

percent of the variance in information quality (predefined = 0.779; query = 0.782; analysis = 0.761). Completeness, accuracy, and format are significant drivers of information quality for predefined reports, query tools, and analysis tools. Currency was not significant in any case.

Accuracy appears to have the strongest influence on information quality for all of the business intelligence tools. The dimension is most influential for predefined reports and somewhat less so for query and analysis tools. Completeness is the second most influential determinant, and it is more important to analysis tools than it is for predefined reports. Format is slightly less influential, and it has a slightly stronger effect within the context of predefined reporting tools.

Overall, the five determinants of system quality explained approximately 75 percent of the variance in system quality (predefined = 0.731; query = 0.743; analysis = 0.759). Reliability, flexibility, accessibility, and integration are significant drivers of

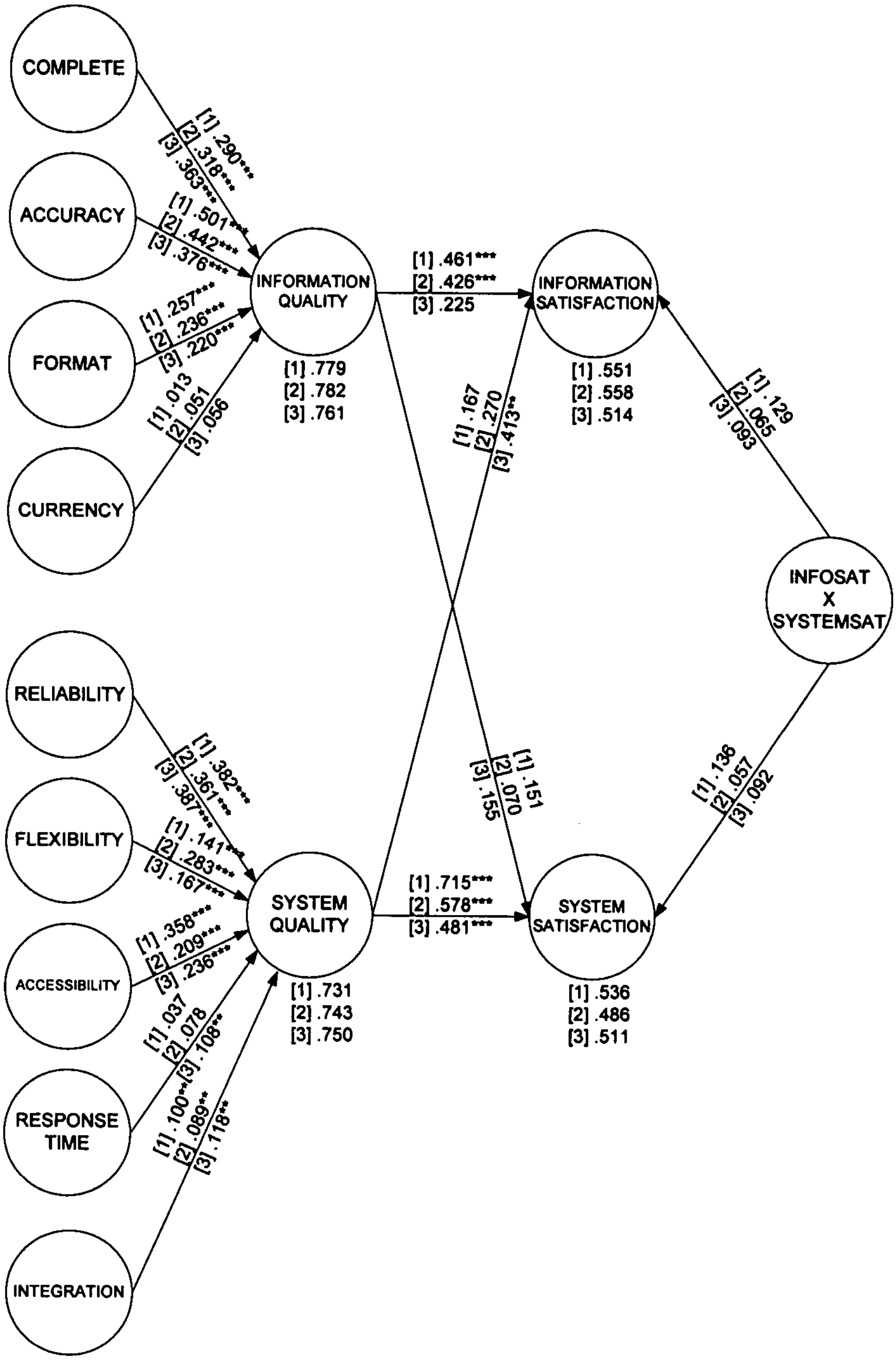

Figure 3. Research Model Results.

Notes: 1 = Predefined reporting software; 2 = query tools; 3 = analysis tools; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

system quality for the three kinds of business intelligence tools, whereas response time is only significant in the case of analysis tools.

Reliability appears to have the strongest influence on system quality across the three business intelligence tools. On average, accessibility and then flexibility are the next most influential factors, although the two positions are reversed for query tools. Thus, while reliability appears to have a universally high level of effect on the assessment of system quality, and integration appears to have a consistently weak effect, , the relative effect of accessibility and flexibility appears to be more tool-dependent.

Within this specific context of data warehousing, we did not find an interaction between system and information satisfaction. And, the crossover effects from information and system quality to information and system satisfaction were not significant within the contexts of predefined reporting and query tools. However, it should be noted that within the context of analysis tools, the path leading from system quality to information satisfaction was significant (0.413), and the path from information quality to information satisfaction was not significant. This will be explored further next.

Discussion

THIS STUDY DEVELOPS AND TESTS A MODEL that explains how various system and information attributes influence information and system quality and, ultimately, user satisfaction. Based on empirical results, it appears that our set of determinants is indeed predictive of overall information and system quality in data warehouse environments and that our model strikes a balance between being comprehensive and parsimonious. At the same time, it is clear from the empirical results that the determinants of quality are not all equivalent in their predictive power within the context of data warehousing.

For information quality, accuracy is the dominant determinant across all three data warehouse technologies. As the core intrinsic attribute of information quality, this is not surprising, especially in a data warehousing context. At the same time, this suggests that there is significant perceived variance in accuracy that reflects overall perceptions of information quality. Further, it reinforces the importance of managing information accuracy as a key determinant of quality. It should not be assumed that users would universally recognize data warehouse initiatives as providing high-quality data.

Completeness and format are the next most influential determinants across the three data warehousing technologies. The importance of completeness may be accentuated in the data warehouse environment, because the integration of disparate information sources is a key data warehouse implementation objective. Format also has a consistent effect across the three technologies, which is again consistent with a warehouse's information and decision-making orientation. Our results suggest that these three dimensions collectively account for three-quarters of the variance in information quality. Whereas subsequent research would be required to test the consistency and robustness of this finding across other contexts, it provides an initial indication that a small number of factors can be identified and managed to influence information quality.

The results suggest that a data warehouse project should emphasize accuracy, completeness, and format as the primary drivers of information quality. Ultimately, experimental tests would need to be conducted to assess the degree of causality.

These results also suggest that more attention needs to be given to the differences across varying technologies. Although the three information quality dimensions are significant across the three business intelligence tool contexts, there do seem to be some differences in the strength of the relationships. For example, completeness seems to increase in importance moving from predefined reports to query tools to analysis tools. The reverse is true for accuracy and format. If you consider that predefined reports are much more restrictive tools, these patterns make sense. A user may expect a predefined reporting tool to provide less complete information (i.e., access to data is restricted to what has been predefined), and that same user would expect more accurate information (because the report has been well planned and defined) in a better format.

The absence of a significant effect for currency likely stems from the nature of the items that measured it. The currency questions were asked in absolute terms; for example, "The information from the tool is always up-to-date." Given that standard data warehouses, by definition, contain historical data, it is not surprising that absolute currency did not influence a user's satisfaction with data warehouse information. It would be interesting to see if currency matters more within technology contexts such as online trading or real-time warehousing where absolute currency is critical. A specific example of the latter case is the use of data mining in homeland security applications. The National Science Foundation's Computer and Information Sciences and Engineering directorate has recently begun sponsoring research in this area.[4]

Turning to system quality, reliability is consistently the most influential determinant. As a key system dimension, reliability appears to have an influence similar to accuracy. Together, the two results suggest the primacy of intrinsic and system properties in shaping quality. From a design perspective, this suggests that there are stable attributes that can be managed to influence quality. These attributes should be primary concerns in system design.

Accessibility is similar in magnitude to reliability for predefined reports and next in magnitude overall for the three business intelligence tools. Such an ordering makes intuitive sense. Reliability relates to a perception of system dependability. Will the system be there when it is needed? Assuming that it is available, does it provide access to the right information? Finally, flexibility has a significant influence and appears to be particularly important for query tools. Because query tools, by definition, support ad hoc, unplanned requests, the relative importance of flexibility seems plausible.

Integration had a consistent effect across the three tool contexts. Considering that most data warehousing tools are expected to provide integrated data (this is often the main point of data warehousing), it seems logical that integration should be a relevant attribute in the context of data warehouse technologies.

Response time was not significant for the three tools. It may be that response time for business intelligence tools is relatively less critical to the perception of overall system quality. Because a data warehouse does not typically support an ongoing, real-time information process, it does seem reasonable that little weight would be put on response time unless problems were extreme or tactical in nature. This is consistent with our argument for currency as it relates to information quality. Once again, online trading systems and homeland security applications are the types of systems that would seem to necessitate fast response time.

The crossover effects found in the research model also shed light on our understanding of quality. They highlight a continuum of interactivity that makes system characteristics more prominent with analysis tools than with the other two tool contexts. Data analysis tools have environments in which the interface and the user's interaction with the interface have obvious effects on the data that the analysis tool provides, and it may be hard to differentiate an access tool from the output it provides. Given this, it makes sense that system quality affects the user's information satisfaction. We were surprised that this influence was so strong and that information quality did not influence information satisfaction within the analysis tool context. It would be interesting for future researchers to explore exactly what characteristic(s) of analysis tools lead to the ambiguity between the system and its output.

To summarize our results, it appears that users of business intelligence tools want to have a reliable, easy-to-access, integrated, and flexible system that provides accurate, complete, and well-formatted information, and these users have different levels of expectations for different kinds of business intelligence tools. The restrictiveness of a tool may be an important characteristic that varies one's expectations. Users appear less concerned with whether information is current, in an absolute sense, or with a particular response time. These results have implications for both research and the development and implementation of IT applications, but only after considering the limitations inherent to this study.

Limitations

The following three limitations need to be considered before discussing implications for research and practice. First, this research was conducted within the specific domain of data warehousing. As a result, it is uncertain whether or not the findings can be applied more broadly or to other specific forms of technology. Second, this research was based on a cross-sectional survey, and the study contains the typical limitations associated with this kind of research methodology. Third, this research included only specific factors based on prior research, but did not test the universal set of antecedents for information and system quality.

Directions for Research

Although previous studies have identified attributes and conceptual linkages related to quality, a comprehensive focus on the dimensions of information and system quality has been surprisingly absent from the IT literature. One of the challenges in assessing quality is to find definitions and measures that enhance understanding, can be practically implemented, and have the potential to guide management action [65]. In this regard, one of the key contributions of our work is the identification of a comprehensive set of determinants that predicts quality, is relatively parsimonious, and most important, enhances understanding. We feel that this model can serve as a powerful lens both in interpreting the results of prior investigations and in shaping rigorous research models for future inquiry that attempt to tie together system quality, satisfaction, and use.

The antecedents of IS usage and value now can be evaluated using the lens of quality in the design of an IT artifact. More specifically, information quality (completeness, accuracy, format, and currency) and system quality (reliability, flexibility, integration, accessibility, and response time) now can be connected to user satisfaction (information and system satisfaction, respectively), as depicted in Figure 1. Our results suggest a way to integrate elements of the technology artifact with the user perceptions and use that helps to avoid the potential pitfalls associated with taking a "proxy view" of IT [59].

Further testing of the model should examine its robustness and stability across IT environments. We believe that evolving business or technological factors may alter the nature and relative strength of the relationships in a conceptual model, such as the one developed here. Our results suggest that some factors are more important than others in the data warehousing context that we examined. It is not clear if these results will be stable across technologies or applications. In fact, one would suppose that theorizing about the relative effect of the quality attributes for different forms of technology would be an important avenue of research to pursue. Contemporary examples include Web-based applications [42] and mobile Internet services [17]. In addition to online trading and homeland security, contemporary applications that should present a different test of the model would be reservation systems and some real-time inventory systems (requiring currency, integration, and response time).

There also is the opportunity to look at changes in the model across different kinds of technologies within the same context (similar to this study) with the purpose of investigating exactly which technology characteristics influence the strength of the quality antecedents. For example, this study seems to suggest that the restrictiveness of a tool makes a difference in shaping quality perceptions. Further, other nontechnical characteristics, such as task type or user demographics, may play important roles in understanding quality. Within data warehousing, users with tactical versus strategic tasks or ones with enterprise-level access to data versus departmental access to data may cause different variations of the research model even when the same business intelligence tool is used. This needs to be studied further.

In addition to examining the stability of results across technologies and applications, the model should be tested with alternative perspectives on measurement. We have taken the approach of asking for absolute assessments of quality attributes that are interpreted relative to individual need. Alternative approaches to assessing quality based on expectation gaps, as is typical in the service quality literature [37], might also be considered.

In addition, future research could explore the relationship of different development methodologies and techniques to information and system quality. For example, do object-oriented practices lead to more flexible applications? Does extreme programming increase application reliability? Thus, the model provides an avenue to conceptualize how various design methods and techniques might influence system and information quality through the individual quality attributes. This provides an approach for linking design methodologies to their ultimate effects on system usefulness, usability, and use.

Implications for Practice

This study contributes to practice in three important ways. First, the development of a model linking specific quality antecedents to quality and then to satisfaction should ultimately allow system designers to better understand and influence system use. For example, in the case of the three business intelligence tools studied here, our findings suggest that designers should be focusing on producing accurate, complete, and clearly formatted information in their quest to enhance information quality. With respect to system quality, designers will have the greatest effect on quality if they focus on creating a reliable, accessible, integrated, and flexible system. The crossover relationships suggest that when designing analysis tools, designers may need to know that the quality of the system strongly affects the users' ultimate satisfaction with its output and focus on improving system quality accordingly. As data is collected across technologies and in different environments, it should become clearer to what extent there are dominant quality characteristics, secondary factors, and special-purpose quality factors that will guide design. Thus, for example, if accuracy were to emerge as the dominant determinant of information quality, it would suggest the need to pay particular attention to the way validation techniques are embedded in applications.

Second, the set of nine determinants can play a useful role in the comparison of competing vendor solutions for a commercial off-the-shelf application. For example, it would be useful to be able to compare competing business intelligence solutions with respect to their relative weighting on both information quality (completeness, accuracy, format, and currency) and system quality (reliability, flexibility, integration, accessibility, and response time) as appropriate. As we begin to understand the relative importance of the various quality factors, it should be possible to use measures diagnostically in the evaluation of software options.

Finally, understanding the quality factors should serve as an aid to implementation management. Managers can shape communication strategies, testing and trial phases, and technology rollouts so that they better lead to positive user reactions. For example, usability testing approaches can help to shape perceptions of information clarity. Descriptions of information quality assurance processes may be needed to ensure accuracy. Attention to the various quality dimensions should provide both managers and designers additional tools for managing implementation processes.

Conclusion

BASED ON THE THEORETICAL LINKAGE between quality and usage, the primary objective of this study was to increase our understanding of the key dimensions of information and system quality. More specifically, we sought to identify a comprehensive, yet parsimonious, set of determinants that help predict the quality of an IT artifact. Based on the literature, a total of nine fundamental determinants of IT-related quality were identified, four under the rubric of information quality (the *output* of an IS) and five that describe system quality (the *information processing system* required to produce the output).

The nine determinants collectively explained a substantial portion of the variance, over 75 percent overall for information quality and approximately 74 percent overall for system quality. We believe that these results support the aptness of our model within a data warehousing context, and that the model strikes a balance between comprehensiveness and parsimony, while providing additional understanding of how quality perceptions are shaped. In conclusion, we hope that this research will serve as a catalyst for action, encouraging both researchers and practitioners to focus on quality as a core concept within the IT discipline.

Acknowledgments: The authors thank Izak Benbasat, Dale Goodhue, Stefano Grazioli, Michael Morris, Peter Seddon, and anonymous reviewers for their helpful comments on earlier drafts of this paper.

NOTES

1. It is estimated that 50 percent of large companies in North America and Europe are using the balanced scorecard for measuring the quality of financial, customer, internal, and human resources performance [57].

2. Currency is also commonly referred to as timeliness (see Appendix A) and has been further refined by Ballou et al. [6] to include both currency (age of data) and volatility (how long an item remains valid).

3. Two questions were used to form each of the two scales to measure information satisfaction and system satisfaction.

4. www.nsf.gov/od/lpa/news/03/fact030124.htm (accessed July 22, 2004).

REFERENCES

1. Agmon, N., and Ahituv, N. Assessing data reliability in an IS. *Journal of Management Information Systems, 4,* 2 (Fall 1987), 34–44.

2. Ahituv, N. Assessing data reliability in an information system. *MIS Quarterly, 4,* 4 (1980), 61–75.

3. Anderson, C.R., and Zeithaml, C.P. Stage of the product life cycle, business strategy, and business performance. *Academy of Management Journal, 27,* 1 (1984), 5–24.

4. Bailey, J.E., and Pearson, S.W. Development of a tool for measuring and analyzing computer user satisfaction. *Management Science, 29,* 5 (1983), 530–545.

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

6. Ballou, D.P.; Wang, R.; Pazer, H.; and Tayi, G.K. Modeling information manufacturing systems to determine information product quality. *Management Science, 44,* 4 (1998), 462–484.

7. Barclay, D.; Higgins, C.; and Thompson, R. The partial least squares approach to causal modeling, personal computing adoption and use as an illustration. *Technology Studies, 2,* 2 (1995), 285–309.

8. Barki, H., and Huff, S.L. Change, attitude to change, and decision support system success. *Information & Management, 9,* 5 (1985), 261–269.

9. Baroudi, J., and Orlikowski, W. A short-form measure of user information satisfaction: A psychometric evaluation and notes on use. *Journal of Management Information Systems, 4,* 4 (1988), 44–59.

10. Belardo, S.; Karwan, K.R.; and Wallace, W.A. Managing the response to disasters using microcomputers. *Interfaces, 14,* 2 (1984), 29–40.

11. Benbasat, I., and Dexter, A.S. An experimental evaluation of graphical and color-enhanced information presentation. *Management Science, 31,* 11 (1985), 1348–1365.

12. Bennett, J.L. Analysis and design of the user interface for decision support systems. In J.L. Bennett (ed.), *Building Decision Support Systems.* Reading, MA: Addison-Wesley, 1983, pp. 41–64.

13. Blaylock, B.K., and Rees, L.P. Cognitive style and the usefulness of information. *Decision Sciences, 15,* 1 (1984), 74–90.

14. Bovee, M.; Srivastava, R.P.; and Mak, B. A conceptual framework and belief-function approach to assessing overall information quality. *International Journal of Intelligent Systems, 18,* 1 (January 2003), 51–74.

15. Buzzell, R., and Gale, B. *The PIMS Principles: Linking Strategy to Performance.* New York: Free Press, 1987.

16. Cappiello, C.; Francalanci, C.; and Pernici, B. Time-related factors of data quality in multi-channel information systems. *Journal of Management Information Systems, 20,* 3 (Winter 2003–4), 71–91.

17. Chae, M., and Kim, J. Information quality for mobile Internet services: A theoretical model with empirical validation. In V. Storey, S. Sarkar, and J. DeGross (eds.), *Proceedings of the Twenty-Second International Conference on Information Systems.* Atlanta: Association for Information Systems, 2001, pp. 43–54.

18. Chin, W.W. Issues and opinions on structural equation modeling. *MIS Quarterly, 22,* 1 (1998), vii–xvi.

19. Chin, W.W., and Frye, T.A. PLS graph, 2.91.03.04. Department of Decision and Information Systems, University of Houston, 1996.

20. Conklin, J.H.; Gotterer, M.H.; and Rickman, J. On-line terminal response time: The effects of background activity. *Information & Management, 5,* 3 (1982), 169–173.

21. Davis, F.D. Perceived usefulness, perceived ease of use, and user acceptance. *MIS Quarterly, 13,* 3 (1989), 319–341.

22. DeLone, W.H., and McLean, E.R. Information systems success: The quest for the dependent variable. *Information Systems Research, 3,* 1 (1992), 60–95.

23. Deming, W.E. *Quality, Productivity, and Competitive Position.* New York: Free Press, 1986.

24. Doll, W.J., and Torkzadeh, G. The measure of end-user computing satisfaction. *MIS Quarterly, 12,* 2 (1988), 259–274.

25. Feigenbaum, A.V. *Total Quality Control,* 4th ed. New York: McGraw-Hill, 1991.

26. Fishbein, M., and Ajzen, I. *Belief, Intention, Attitude, and Behavior.* Reading, MA: Addison-Wesley, 1975.

27. Fisher, C.W., and Kingma, B.R. Criticality of data quality as exemplified in two disasters. *Information & Management, 39,* 2 (2001), 109–116.

28. Fornell, C., and Larcker, D.F. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18,* 3 (1981), 39–50.

29. Franz, C.R., and Robey, D. Organizational context, user involvement, and the usefulness of information systems. *Decision Sciences, 17,* 3 (1986), 329–355.

30. Gallagher, C.A. Perceptions of the value of a management information system. *Academy of Management Journal, 17,* 1 (1974), 46–55.

31. Goodhue, D.L. Understanding user evaluations of information systems. *Management Science, 41,* 12 (1995), 1827–1845.

32. Goodhue, D.L., and Thompson, R.L. Task-technology fit and individual performance. *MIS Quarterly, 19,* 2 (1995), 213–236.

33. Halloran, D.; Manchester, S.; Moriarty, J.; Riley, R.; Rohrman, J.; and Skramstad, T. Systems development quality control. *MIS Quarterly, 2,* 4 (1978), 1–13.

34. Huh, Y.U.; Keller, F.R.; Redman, T.; and Watkins, A.R. Data quality. *Information and Software Technology, 32,* 8 (1990), 559–565.

35. Ives, B.; Olson, M.H.; and Baroudi, J.J. The measurement of user information satisfaction. *Communications of the ACM, 26,* 10 (1983), 785–793.

36. Jarvenpaa, S.L. The effect of task demands and graphical format on information processing strategies. *Management Science, 35,* 3 (1989), 285–304.

37. Jiang, J.J.; Klein, G.; and Carr, C.L. Measuring information system service quality: SERVQUAL from the other side. *MIS Quarterly, 26,* 2 (2002), 145–161.

38. Juran, J.M, and Godfrey, A.B. *Juran's Quality Handbook,* 5th ed. New York: McGraw-Hill, 1999.

39. Kahn, B.K.; Strong, D.M.; and Wang, R.Y. Information quality benchmarks: Product and service performance. *Communications of the ACM, 45,* 4 (2002), 184–192.

40. Kettinger, W.J.; Lee, C.C.; and Lee, S. Global measures of information service quality: A cross-national study. *Decision Sciences, 26,* 5 (1995), 569–589.

41. King, W.R., and Epstein, B.J. Assessing information system value: An experimental study. *Decision Sciences, 14,* 1 (1983), 34–45.

42. Klein, B.D. Assessment of information quality: Does the medium matter? In D. Galletta and J. Ross (eds.), *Proceedings of the Ninth Americas Conference on Information Systems.* Atlanta: Association for Information Systems, 2003, pp. 2348–2352.

43. Klein, B.D.; Goodhue, D.L.; and Davis, G.B. Can humans detect errors in data? Impact of base rates, incentives, and goals. *MIS Quarterly, 21,* 2 (1997), 169–194.

44. Larcker, D.F., and Lessig, V.P. Perceived usefulness of information: A psychometric examination. *Decision Sciences, 11,* 1 (1980), 121–133.

45. Lee. Y. Crafting rules: Context-reflective data quality problem solving. *Journal of Management Information Systems, 20,* 3 (Winter 2003–4), 93–116.

46. Lee, Y., and Strong, D. Knowing-why about data processes and data quality. *Journal of Management Information Systems, 20,* 3 (Winter 2003–4), 13–39.

47. Lee, Y.; Strong, D.; Kahn, B.; and Wang, R. AIMQ: A methodology for information quality assessment. *Information & Management, 40,* 2 (December 2002), 133–146.

48. Levitin, A.V., and Redman, T.C. Data as a resource: Properties, implications, and prescriptions. *Sloan Management Review, 40,* 1 (1998), 89–101.

49. Lohmoller, J.B. *Latent Variable Path Modeling with Partial Least Squares in Predictive vs. Structural Modeling: PLS vs. ML.* Heidelberg: Physica-Verlag, 1989.

50. Madnick, S.E.; Wang, R.Y.; and Xian, X. The design and implementation of a corporate householding knowledge processor to improve data quality. *Journal of Management Information Systems, 20,* 3 (Winter 2003–4), 41–69.

51. Mahmood, M.A. System development methods—A comparative investigation. *MIS Quarterly, 11,* 3 (1987), 293–311.

52. Mason, R.O. Measuring information systems output: A communication systems approach. *Information & Management, 1,* 5 (1978), 219–234.

53. Miller, H. The multiple dimensions of information quality. *Information Systems Management, 13,* 2 (1996), 79–82.

54. Miller, J., and Doyle, B.A. Measuring the effectiveness of computer-based information systems in the financial services sector. *MIS Quarterly, 11,* 1 (1987), 107–124.

55. Moore, G., and Benbasat, I. Development of an instrument to measure the perceptions of adopting and information technology innovation. *Information Systems Research, 2,* 3 (1991), 192–222.

56. Morey, R.C. Estimating and improving the quality of information in an MIS. *Communications of the ACM, 25,* 5 (1982), 337–342.

57. Norton, D.P. *Balanced Scorecard Report.* Boston: Harvard Business School Press, 2002.

58. Nunnally, J.C. *Psychometric Theory.* New York: McGraw-Hill, 1978.

59. Orlikowski, W., and Iacono, C.S. Research commentary: Desperately seeking the "IT" in IT research—A call to theorizing the IT artifact. *Information Systems Research, 12,* 2 (2001), 121–134.

60. Orr, K. Data quality and systems theory. *Communications of the ACM, 41,* 2 (1998), 66–71.

61. Perlow, L.A. *Finding Time: How Corporations, Individuals, and Families Can Benefit from New Work Practices.* Ithaca: Cornell University Press, 1997.

62. Pipino, L.L.; Lee, Y.W.; and Wang, R.Y. Data quality assessment. *Communications of the ACM, 45,* 4 (2002), 211–218.

63. Pitt, L.; Watson, R.T.; and Kavan, C.B. Service quality: A measure of information systems effectiveness. *MIS Quarterly, 19,* 2 (1995), 173–187.

64. Rai, A.; Lang, S.S.; and Welker, R.B. Assessing the validity of IS success models: An empirical test and theoretical analysis. *Information Systems Research, 13,* 1 (2002), 50–69.

65. Reeves, C., and Bednar, D.A. Defining quality: Alternatives and implications. *Academy of Management Review, 19,* 3 (1994), 419–445.

66. Robey, D. User attitudes and management information system use. *Academy of Management Journal, 22,* 3 (1979), 527–538.

67. Seddon, P. A respecification and extension of the DeLone and McLean model of IS success. *Information Systems Research, 8,* 3 (1997), 240–253.

68. Seddon, P., and Kiew, M.-Y. A partial test and development of the DeLone and McLean model of IS success. In J.I. DeGross, S.L. Huff, and M.C. Munro (eds.), *Proceedings of the International Conference on Information Systems.* Atlanta: Association for Information Systems, 1994, pp. 99–110.

69. Shannon, C.E., and Weaver, W. *The Mathematical Theory of Communication.* Urbana: University of Illinois Press, 1949.

70. Shaw, N.G. Capturing the technological dimensions of IT infrastructure change: A model and empirical evidence. *Journal of the AIS, 2,* 8 (2002), 1–33.

71. Srinivasan, A. Alternative measures of system effectiveness: Associations and implications. *MIS Quarterly, 9,* 3 (1985), 243–253.

72. Strong, D.M.; Lee, Y.W.; and Wang, R.Y. Data quality in context. *Communications of the ACM, 40,* 5 (1997), 103–110.

73. Swanson, E. B. Management information systems—Appreciation and involvement. *Management Science, 2,* 2 (1974), 178.

74. Tayi, G.K., and Ballou, D.P. Examining data quality. *Communications of the ACM, 41,* 2 (1998), 54–57.

75. Taylor, S., and Todd, P.A. Understanding information technology usage: A test of competing models. *Information Systems Research, 6,* 2 (1995), 144–176.

76. Tractinsky, N., and Meyer, J. Chartjunk or goldgraph? Effects of presentation objectives and content desirability on information processing. *MIS Quarterly, 23,* 3 (1999), 397–421.

77. Venkatesh, V., and Davis, F.D. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46,* 2 (2000), 186–204.

78. Vessey, I. Cognitive fit: A theory-based analysis of the graphs versus tables literature. *Decision Sciences, 22,* 2 (1991), 219–240.

79. Vessey, I., and Galletta, D. Cognitive fit: An empirical study of information acquisition. *Information Systems Research, 2,* 12 (1991), 63–84.

80. Wand, Y., and Wang, R.Y. Anchoring data quality in ontological dimensions. *Communications of the ACM, 39,* 11 (1996), 86–96.

81. Wang, R.Y. A product perspective on total data quality management. *Communications of the ACM, 41,* 2 (1998), 58–65.

82. Wang, R.Y., and Strong, D.M. Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12,* 4 (Spring 1996), 5–34.

83. Washburne, J.N. An experimental study of various graphic, tabular and textual methods of presenting quantitative material. *Journal of Educational Research, 18,* 6 (1927), 465–476.

84. Wixom, B.H., and Todd, P.A. User acceptance and user satisfaction: Towards a theoretical integration. *Information Systems Research,* forthcoming.

85. Wixom, B.H., and Watson, H.J. An empirical investigation of the factors affecting data warehousing success. *MIS Quarterly, 25,* 1 (2001), 17–41.

86. Wold, H., and Joreskog, K. *Systems Under Indirect Observation: Causality, Structure, Prediction.* Amsterdam: North-Holland, 1982.

87. Zeithaml, V.A. Service quality, profitability, and the economic worth of customers: What we know and what we need to learn. *Academy of Marketing Science Journal, 28,* 1 (2000), 67–86.

88. Zeithaml, V.A.; Parasuraman, A.; and Berry, L.L. *Delivering Quality Service.* New York: Free Press, 1990.

89. Zmud, R.W. An empirical investigation of the dimensionality of the concept of information. *Decision Sciences, 9,* 2 (1978), 187–195.

Appendix A. Information Quality Dimensions

Construct name	Definition	References
	Accuracy	
Absolute reliability	When the database mirrors reality.	[1]
Accuracy	Agreement with a real-world entity, a value stored in another database, or the results of an arithmetic computation.	[43]
Accuracy	Data are certified error-free, accurate, correct, flawless, reliable, and precise. Errors can be easily identified, and the data has integrity.	[82]
Accuracy	It does not contain a substantial amount of bias.	[41]
Accuracy	Measure of agreement with an identified source.	[34]
Accuracy	Reflects the underlying reality.	[53]
Accuracy	The correctness of the output information.	[4, 8, 9, 35, 54]
Accuracy	The frequency of errors in the data.	[33]
Accuracy	The level of accuracy expected from the system.	[66]
Accuracy	The recorded value conforms to the real-world fact or value.	[27]
Accuracy	The recorded value is in conformity with the actual value.	[5]
Accuracy	The extent that data from the system requires correction.	[29]
Accurate	Accurate, believable.	[89]
Believable	Complete, consistent, credible source, and accurate.	[39, 82]
Coherence	How well the information hangs together and is consistent.	[53]
Consistency	It is not noticeably impaired as to comparability.	[41]
Consistency	The recorded value is the same in all cases.	[27]
Consistency	The representation of the data value is the same in all cases.	[5]
Consistency	Two sets of data do not conflict with each other.	[34]
Correct	No multiple states mapped to the same IS states.	[80]
Error rate	The amount of errors found in the data.	[56]
Free of error	The extent to which information is correct and reliable.	[39]
Factual	Factual, true.	[89]

(continues)

Appendix A. Continued

Construct name	Definition	References
	Accuracy (continued)	
Integrity	Accurate, complete, consistent, and existent.	[14]
Meaningful	No map to a wrong state.	[80]
Objectivity	Unbiased, unprejudiced, and impartial.	[39, 82]
Precision	The variability of the output information from that which it purports to measure.	[4,8, 9, 35]
Reasonable	Logical, sensible.	[89]
Reliability	The consistency and dependability of the output information.	[4, 8, 9, 35]
Reliability	Accurate, reliable, valid, true.	[30]
Reputation	The extent to which information is highly regarded in terms of its source or content.	[39]
Similarity	Accuracy of the reported data.	[2]
Unambiguous	No map to a meaningless IS state.	[80]
Validity	It can be verified as being true and satisfying appropriate standards related to other dimensions.	[53]
	Completeness	
Completeness	No missing IS states.	[80]
Completeness	Breadth, depth, and scope of information contained in the data.	[82]
Completeness	All values for a certain variable are recorded.	[5]
Completeness	The degree to which values are present in a data collection.	[27]
Completeness	The set contains all relevant data.	[34]
Completeness	Complete.	[53]
Completeness	The comprehensiveness of the output information.	[4, 8, 9, 35, 54]
Completeness	It is not sufficient alone.	[41]
Completeness	Complete, enough.	[30]
Importance and usableness	Sufficient, the portion of information essential for completing a specific decision.	[13, 44]
Quantity	Complete, effective, material, sufficient.	[89]
Appropriate amount of information	The extent to which the volume of information is appropriate for the task at hand.	[39]

Currency		
Currency	Data are up to date.	[34]
Currency	The age of the output information.	[4, 8, 35, 54]
Timeliness	The age of the data.	[82]
Timeliness	The recorded value is not out of date.	[5, 27]
Timeliness	Currency. Refers to the age of the primitive data units used to produce the information products.	[6]
Timeliness	Current.	[16, 30, 53, 89]
Timeliness	Up-to-date information.	[24]
Volatility	How long an item remains valid.	[6]
	Format	
Arrangement	Orderly, precise.	[89]
Concise	Well-presented, concise, compactly represented, well-organized, aesthetically pleasing, form of presentation, well-formatted, format of the data.	[39, 82]
Consistent representation	Data are consistently presented in the same format, consistently represented, consistently formatted, and data are compatible with previous data.	[39, 82]
Ease of understanding	Easily understood, clear, readable.	[82]
Format	How the information is presented to the customer.	[53]
Format	The material design of the layout and display of the output contents.	[4]
Format	Information is presented in useful format and is clear.	[24]
Format	The medium, ordering, and graphic design of a report.	[2]
Format	Readable, orderly, logical, clear, simple.	[30]
Format	The amount of complex recalculations or adjustments necessary in order to use the information presented to complete a specific decision.	[13, 44]
Interpretability	The extent to which information is in appropriate languages, symbols, and units, and the definitions are clear.	[14, 39, 82]
Quantitativeness	It is primarily quantitative.	[41]
Readable	Clear, convenient, readable, simple.	[89]
Relevancy	The extent to which information is applicable and helpful for the task at hand.	[14, 39]
Understandability	It is understandable.	[41]
Value added	The extent to which information is beneficial and provides advantages from its use.	[39]

Note: The studies included in this table are those that *explicitly* defined the constructs that were used.

Appendix B. System Quality Dimensions

Construct Name	Definition	References
Reliability		
Reliable	Valid.	[89]
Reliable	Low percentage of hardware and software downtime.	[54]
Reliability	Recoverability, the ability of the system to restore bad data and/or malfunctioning equipment to an operational state.	[33]
Technical quality	Technical problems, such as bugs, crashes, and so on.	[70]
Response time		
Response time	The difference between the time the information is received and the time it was requested.	[2]
Response time	The time interval from the time the operator depressed the transmit key until the first response character appeared on the screen.	[20]
Response/turnaround time	The elapsed time between a user-initiated request for service or action and a reply to that request.	[4]
Timeliness	The availability of the output information at a time suitable for its use.	[4, 8, 35]
Timeliness	The extent to which the information is sufficiently up-to-date for the task at hand.	[39]
Timeliness	Getting the information you need in time.	[24]
Timeliness	Timeliness of report delivery to users.	[54]
Timeliness	Data is generated frequently.	[41]
Timeliness	Data is available without delay.	[41]
Timeliness	The time it takes a system to respond to user needs.	[33]

Accessibility		
Accessibility	Accessible, retrievable, speed of access, and available.	[14, 39, 82]
Accessibility	Information can be obtained when needed.	[53]
Accessibility	Available.	[82]
Convenience of access	Ease or difficulty with which the user may act to utilize the capability of the computer system.	[4, 54]
Ease of operation	Easily joined, changed, updated, downloaded or uploaded. Data can be used for multiple purposes, manipulable, aggregated, reproduced, data can be integrated, and customized.	[82]
Security	Appropriate restrictions to maintain security.	[39]
Flexibility		
Flexibility	Adaptable, flexible, extendable, and expandable.	[82]
Flexibility	The capacity of the IS to change or to adjust in response to new conditions, demands, or circumstances.	[4, 54]
Flexibility	System responsiveness to changing user needs.	[54]
Flexibility	The extent to which system features and options lend themselves to accommodating change without modification to programs.	[33]
Ease of manipulation	The extent to which information is easy to manipulate and apply to different tasks.	[39]
Integration		
Compatibility	Information can be combined with other information and delivered to the customer.	[53]
Compatibility	The compatibility of the new software package with other software on your computer, old software that was replaced, other software packages in the company, other software in the marketplace, and the hardware in your computer.	[70]
Integration	The ability of systems to communicate/transmit data between systems servicing different functional areas.	[4]
Variety of data and data sources	Having a variety of data and data sources.	[82]

Note: The studies included in this table are those that *explicitly* defined the constructs that were used.

Appendix C. Correlations of Latent Variables

	COMP	ACCU	FORM	CURR	RELI	FLEX	INTE	TIME	ACCE	INFQ	SYSQ	SYSS	INFS
Predefined reports													
COMP	0.91												
ACCU	0.551	0.87											
FORM	0.664	0.525	0.89										
CURR	0.556	0.618	0.477	0.92									
RELI	0.618	0.688	0.585	0.566	0.90								
FLEX	0.686	0.355	0.474	0.382	0.427	0.86							
INTE	0.678	0.453	0.496	0.495	0.477	0.513	0.88						
TIME	0.558	0.560	0.573	0.561	0.745	0.487	0.478	0.83					
ACCE	0.689	0.536	0.637	0.515	0.699	0.554	0.559	0.687	0.89				
INFQ	0.734	0.776	0.648	0.639	0.724	0.562	0.617	0.626	0.703	0.91			
SYSQ	0.772	0.728	0.713	0.544	0.768	0.572	0.573	0.684	0.785	0.853	0.91		
SYSS	0.647	0.397	0.604	0.360	0.576	0.582	0.482	0.567	0.711	0.648	0.730	0.95	
INFS	0.659	0.516	0.559	0.492	0.595	0.586	0.516	0.543	0.624	0.754	0.716	0.779	0.96
Analysis													
COMP	0.91												
ACCU	0.619	0.84											
FORM	0.655	0.564	0.87										
CURR	0.577	0.569	0.489	0.91									
RELI	0.684	0.677	0.613	0.559	0.89								
FLEX	0.736	0.545	0.555	0.497	0.630	0.85							
INTE	0.672	0.501	0.531	0.527	0.578	0.639	0.89						
TIME	0.631	0.550	0.583	0.566	0.650	0.555	0.501	0.82					
ACCE	0.776	0.574	0.716	0.528	0.680	0.706	0.550	0.671	0.89				
INFQ	0.769	0.756	0.682	0.657	0.747	0.699	0.685	0.667	0.734	0.89			
SYSQ	0.781	0.712	0.687	0.552	0.791	0.713	0.632	0.670	0.754	0.857	0.90		
SYSS	0.647	0.397	0.604	0.360	0.576	0.582	0.482	0.567	0.711	0.648	0.730	0.95	
INFS	0.659	0.516	0.559	0.492	0.595	0.586	0.516	0.543	0.624	0.754	0.716	0.779	0.95

Queries

COMP	0.90												
ACCU	0.603	0.86											
FORM	0.624	0.535	0.88										
CURR	0.582	0.559	0.406	0.92									
RELI	0.628	0.645	0.613	0.496	0.89								
FLEX	0.765	0.537	0.541	0.564	0.557	0.86							
INTE	0.625	0.490	0.489	0.463	0.543	0.568	0.89						
TIME	0.614	0.534	0.558	0.542	0.716	0.584	0.426	0.83					
ACCE	0.740	0.518	0.630	0.487	0.710	0.674	0.499	0.662	0.89				
INFQ	0.751	0.790	0.645	0.651	0.713	0.694	0.652	0.633	0.672	0.90			
SYSQ	0.776	0.729	0.680	0.548	0.770	0.720	0.582	0.678	0.751	0.863	0.90		
SYSS	0.648	0.450	0.540	0.425	0.611	0.647	0.427	0.572	0.756	0.620	0.696	0.95	
INFS	0.661	0.570	0.563	0.542	0.620	0.658	0.499	0.573	0.679	0.733	0.728	0.789	0.96

Notes: Diagonal elements are the square root of AVE. These values should exceed the interconstruct correlations for adequate discriminant validity. COMP = completeness; ACCU = accuracy; FORM = format; CURR = currency; RELI = reliability; FLEX = flexibility; INTE = integration; TIME = response time; ACCE = accessibility; INFQ = information quality; SYSQ = system quality; INTN = intention; ATTI = attitude; EASE = ease of use; USEF = usefulness; SYSS = system satisfaction; INFS = information satisfaction.

Appendix D. Factor Loadings

Factor	Factor loadings		
	Predefined reports	Query tools	Analysis tools
Completeness			
____ provides me with a complete set of information.	0.91	0.91	0.91
____ produces comprehensive information.	0.91	0.90	0.91
____ provides me with all the information I need.	0.91	0.89	0.90
Accuracy			
____ produces correct information.	0.91	0.89	0.89
There are few errors in the information I obtain from ____.	0.78	0.77	0.73
The information provided by ____ is accurate.	0.91	0.91	0.89
Format			
The information provided by ____ is well formatted.	0.89	0.89	0.86
The information provided by ____ is well laid out.	0.92	0.91	0.91
The information provided by ____ is clearly presented on the screen.	0.85	0.83	0.84
Currency			
____ provides me with the most recent information.	0.92	0.92	0.92
____ produces the most current information.	0.93	0.94	0.93
The information from ____ is always up to date.	0.91	0.90	0.89
Information quality			
Overall, I would give the information from ____ high marks.	0.90	0.89	0.88
Overall, I would give the information provided by ____ a high rating in terms of quality.	0.92	0.91	0.89
In general, ____ provides me with high-quality information.	0.92	0.91	0.90
Information satisfaction			
Overall, the information I get from ____ is very satisfying.	0.96	0.96	0.95
I am very satisfied with the information I receive from ____.	0.96	0.96	0.95

Reliability			
____ operates reliably.	0.90	0.90	0.91
____ performs reliably.	0.92	0.91	0.91
The operation of ____ is dependable.	0.87	0.86	0.86
Flexibility			
____ can be adapted to meet a variety of needs.	0.83	0.82	0.81
____ can flexibly adjust to new demands or conditions.	0.87	0.88	0.86
____ is versatile in addressing needs as they arise.	0.88	0.89	0.88
Integration			
____ effectively integrates data from different areas of the company.	0.87	0.88	0.88
____ pulls together information that used to come from different places in the company.	0.88	0.88	0.87
____ effectively combines data from different areas of the company.	0.90	0.91	0.92
Accessibility			
____ allows information to be readily accessible to me.	0.88	0.89	0.90
____ makes information very accessible.	0.91	0.91	0.90
____ makes information easy to access.	0.88	0.88	0.88
Response time			
It takes too long for ____ to respond to my requests. (RC)	0.78	0.80	0.80
____ provides information in a timely fashion.	0.83	0.80	0.77
____ returns answers to my requests quickly.	0.89	0.88	0.88
System quality			
In terms of system quality, I would rate ____ highly.	0.90	0.88	0.89
Overall, ____ is of high quality.	0.91	0.91	0.91
Overall, I would give the quality of ____ a high rating.	0.93	0.92	0.91
System satisfaction			
All things considered, I am very satisfied with ____.	0.94	0.94	0.94
Overall, my interaction with ____ is very satisfying.	0.96	0.96	0.96

RC = reverse coded.