Communications of the Association for Information Systems

Volume 22 Article 3

1-2008

Developing a Measurement Instrument for Subjective Aspects of Information Quality

Rosanne Price
Monash University

Dina Neiger
Monash University

Graeme Shanks
University of Melbourne, Graeme.Shanks@infotech.monash.edu.au

Communications of the Association for Information Systems

CAIS

Developing a Measurement Instrument for Subjective Aspects of Information Quality

Rosanne Price

Dina Neiger

Faculty of Information Technology, Monash University

Rosanne.Price@infotech.monash.edu.au

Graeme Shanks

Department of Information Systems, University of Melbourne

Abstract:

Having quality information is crucial for effective operations and decision making within organisations. The *InfoQual* framework provides a sound theoretical basis for defining information quality at three levels: syntactic (form), semantic (content), and pragmatic (usage). Objective measures can be defined for the syntactic and semantic levels. In this paper, we focus on the pragmatic level by developing and empirically testing an instrument that aims to measure subjective aspects of information quality based on the perceptions of information consumers. In combination, such a framework and instrument have the potential to aid organizations in identifying problems and planning improvement strategies for information quality.

Keywords: subjective information quality, instrument development, questionnaire

Volume 22, Article 3, pp. 49-74, January 2008

Communications of the Association for Information Systems

I. INTRODUCTION

There is increasing recognition of the importance of information quality (IQ) to organizations and the need for active management of IQ. As a starting point, organizations must be able to monitor the quality of the information they produce or use, including both stored data sets and the information retrieved from those data sets. This requires both a clear understanding of the IQ criteria that must be considered and a means of measuring quality based on these criteria. Essentially, the necessary foundation for IQ management is an effective means of defining and evaluating IQ.

IQ is commonly defined in terms of a set of quality criteria grouped into quality categories. Competing views of quality from the perspective of either stored data or received information focus on objective (i.e. relatively use-independent[1]) or subjective (i.e. use-dependent) quality criteria respectively. The former view is based on conformance to initial specifications (including specified integrity rules) or correspondence to represented real-world phenomena. The latter view is based on consumer judgments of perceived IQ in the context of data use, where perceptions are influenced by data delivery (e.g. interface quality) and consumer expectations.

A number of frameworks defining IQ have been proposed [see the survey in Eppler 2001] based on theoretical [Wand and Wang 1996], empirical [Kahn et al. 2002; Lee et al. 2002; Wang and Strong 1996], or intuitive [English 1999; Redman 1996] research approaches. A purely theoretical research approach to the definition of IQ categories and criteria is necessarily limited in scope to the objective view of IQ. Empirical or intuitive research approaches rely respectively on subjective information consumer judgments or personal experience rather than systematic theory. As a consequence, the resulting frameworks suffer from problems of consistency, particularly with respect to the definition of quality categories and classification of quality criteria into categories [see Eppler 2001 pp. 178; Lee et al. 2002 pp. 135; Price and Shanks 2004; Price and Shanks 2005b].

Price and Shanks [2004, 2005a, 2005b] have recently proposed that semiotic theory, the philosophical theory of signs, be used to address issues of both scope and consistency. In their IQ framework *InfoQual,* semiotics provides a theoretical and thus consistent basis for *defining* quality categories, *classifying* quality criteria, and *integrating* different IQ views and research approaches. Significantly, the fact that the last two steps follow implicitly and automatically from the first ensures their consistency and coherence respectively.

The objective view of IQ is represented by *InfoQual*'s *syntactic* and *semantic* quality categories; whereas the subjective view is represented by *InfoQual*'s *pragmatic* quality category. Theoretical techniques can be used to derive syntactic and semantic criteria. In contrast, empirical techniques are required to derive pragmatic criteria since their selection depends on understanding which specific information characteristics consumers consider important for assessing the suitability of available information for their use.

With respect to IQ evaluation, ad-hoc and problem-specific measures are the norm [Pipino et al. 2002]. Pipino et al. [2002] present three functional forms—simple ratio, min/max operators, and weighted average—that can be used for developing objective IQ metrics. English [1999, Ch 6 and 10] discusses automated tools for assessing design conformance and techniques for assessing correspondence between system and represented (e.g. real-world) values. For example, selective or random sampling can be used to compare system to represented real-world values either directly (e.g. using point of customer contact by front of house staff) or indirectly to a trusted source or surrogate database (e.g. using a telephone directory for customer addresses or phone numbers). Redman [1996, Ch 10] considers the use of statistical quality control to measure the rate at which defective (e.g. unsatisfactory) data or information is produced by business processes. Finally, questionnaires (called *measurement instruments*) have been developed to assess information consumer perceptions of IQ for specific business domains (e.g. in Barnes and Vidgen [2002]) or indirectly as one factor in a broader IS perspective (e.g. in the context of measuring IS satisfaction in Chin and Lee [2000]). Lee et al. [2002] developed an IQ instrument intended for general application; however, the underlying IQ framework used was empirically developed and thus is subject to issues of consistency as discussed earlier.

[1] See [Price and Shanks 2004] for a discussion of the relative degree of use-independence and objectivity.

Regardless of the objective IQ metrics used by an organization, subjective consumer-based perceptions represent a key indicator of IQ in practice since they represent actual use-based evaluation. Thus IQ assessment based on questionnaires reflects the information's *fitness for use.* While objective measurements of IQ may provide assistance in detecting problems of conformance (to system design) or correspondence (to the real-world), subjective measurements are required to account for problems experienced due to unsupported or changed consumer requirements or expectations. Therefore, the development of a questionnaire that can be used to assess consumer perceptions of IQ is critical to its evaluation and management in an organization.

Research reported in this paper describes *the development of such a measurement instrument for subjective IQ based on the pragmatic category of the* InfoQual *framework*. The resulting instrument, the *Subjective IQ Questionnaire (SIQQ)*, is intended as a generic instrument applicable to general business application domains and data types. The aim is to develop an instrument to assess information consumer perceptions of the quality for a given data set, where information consumers include information producers, managers, and end-users either internal or external to the organization. The goals of the paper are twofold: to report the results of the instrument development and to describe in detail the actual development process to serve as an aid to others considering instrument development.

The research reported in this paper was guided by standard instrument development methods (e.g. Moore and Benbasat [1991]; Ewing and Napoli [2005]) and the validation guidelines for IS positivist research proposed by Straub et al. [2004]. The mandatory validation guidelines given by Straub [2004, pp. 385, Table 1] encompass *content*, *construct,* and *reliability* aspects of validity. According to Straub et al. [2004, pp. 387], "*content validity* is established through literature reviews and expert judges or panels," *construct validity* is concerned with the operationalization of the measurement instrument [pp. 388], and *reliability* is "a statement about measurement accuracy" in terms of the stability and internal consistency of the measure [Straub 1989]. Based on these three aspects of validity the remainder of the paper is structured as follows.

Sections II and III of this paper address the issue of *content validity*. Section II describes the *InfoQual* framework and the development techniques used in the derivation of the framework's pragmatic criteria. Section III discusses the development and operationalization (including how and what data were collected) of the measurement instrument based on the framework's pragmatic category. Section IV is concerned with the techniques that are most suitable for assessment of *construct validity* and *reliability* in the current experimental context. Factor analysis techniques are used to assess *discriminant*, *convergent,* and *factorial* aspects of *construct validity* [Straub et al. 2004, pp. 410]. Cronbach's alpha and split sample analysis of the final factor solution are used to assess *reliability*. Section V contains a detailed presentation of the empirical results and Section VI discusses their implications for the final factor solution and measurement instrument. The paper is concluded in Section VII.

II. BACKGROUND: THE *INFOQUAL* FRAMEWORK

The IQ framework *InfoQual* provides the underlying theoretical basis for *SIQQ,* the subjective IQ instrument proposed in this paper. Previous publications describe in detail theoretical [Price and Shanks 2004], empirical, [Price and Shanks 2005a; Price and Shanks 2005b], and comparative aspects [Price and Shanks 2005b] of the *InfoQual* framework. The intention of this section is to provide sufficient detail to serve as context for the measurement instrument and associated field study described in the rest of the paper. Thus, we first give a general description of the framework's conceptual foundation (i.e. semiotics), structure, and criteria. Particular emphasis is given to describing the pragmatic criteria and their derivation, since the pragmatic category forms the basis for the subjective IQ instrument *SIQQ* described in Section III.

Classical philosophical semiotics forms the conceptual foundation of *InfoQual.* In particular, Peirce [1931-1935] and Morris [1938] describe communication via signs using three components and three levels. The components describe the representation, intended meaning, and use of a sign respectively. The *syntactic*, *semantic*, and *pragmatic* levels describe, respectively, relations between sign representations; between a sign representation and its meaning; and between a sign representation and its use.

These components and levels can be used to describe an Information System (IS), since IS data can be regarded as "signs" that represent external[2] (e.g. "real-world") phenomena. Thus salary data for an employee has a stored representation (e.g. employee salary field), an intended meaning (e.g. employee's actual salary), and a use (e.g. payroll). Similarly, IS metadata (e.g. the integrity rule *emp.sal≥0*) can be regarded as signs for external definitions, rules, or documentation relevant to an application or data model (e.g. employee salary must be non-negative). In the IS context, the three semiotic levels can then be used to describe relations between IS data and metadata (both sign

[2] *External* here refers to something in the domain being modeled (represented) by the database and IS, thus "external" to the database and IS.

representations); between IS data and represented real-world phenomena (a sign representation and its intended meaning); and between data and use (a sign representation and its use).

Quality categories are defined based on the desirable characteristics at each of these levels, i.e. conformance (of data to metadata), correspondence (of data to real-world phenomena), and suitability (of data for use). In the context of employee salary data, *syntactic*, *semantic*, and *pragmatic* quality aspects relate to whether such salary data conforms to relevant integrity rules (e.g. *emp.sal≥0*), whether it matches actual employee salaries, and whether it is useful for a given purpose (e.g. payroll).

A summary of the semiotic IQ framework *InfoQual* is shown below in Table 1. Individual categories, their criteria, and the derivation process used for those criteria are discussed further in the two subsections following the table.

Table 1. Quality Criteria by Category

Syntactic Criteria (based on rule conformance)
Conforming to data integrity rules. Data follows specified database integrity rules.
Semantic Criteria (based on external correspondence)
Mapped completely. Every external phenomenon is represented.
Mapped consistently. Each external phenomenon is either represented by at most one identifiable data unit or by multiple but consistent identifiable units or by multiple identifiable units whose inconsistencies are resolved within an acceptable time frame.
Mapped unambiguously. Each identifiable data unit (e.g. relational tuple) represents at most one specific external phenomenon.
Mapped meaningfully. Each identifiable data unit represents at least one specific real-world phenomenon.
Phenomena mapped correctly. Each identifiable data unit maps to the correct external phenomenon.
Properties mapped correctly. Non-identifying (i.e. non-key) attribute values in an identifiable data unit match the property values for the represented external phenomenon.
Pragmatic Criteria (use-based consumer perspective)
Accessible. Data is easy and quick to retrieve.
Suitably presented. Data is presented in a manner appropriate for its use, with respect to format, precision, units, and the types of data displayed.
Flexibly presented. Data can be easily manipulated and the presentation customized as needed, with respect to aggregating data and changing the data format, precision, units, or types of data displayed.
Understandable. Data is presented in an intelligible (i.e. comprehensible) manner.
Timely. The currency (age) of the data is appropriate to its use.
Secure. Data is appropriately protected from damage or abuse (including unauthorized access, use, or distribution).
Allowing access to relevant metadata. Appropriate metadata is available to define, constrain, and document data.
Perceived to be conforming to data integrity rules. Data follows specified database integrity rules.
Perceived to be complete. There are no data missing, i.e. every external phenomenon is represented in the data.
Perceived to be reliable. The data is dependable, i.e. there is a correct one-to-one mapping (i.e. correspondence) of external phenomena to data.

The Objective IQ View: Syntactic and Semantic Categories and Criteria

The *syntactic quality category* describes the degree to which stored data conform to stored meta-data. A single syntactic criterion of *conforming to metadata* can be derived directly from the definition of the syntactic quality category, where *metadata* includes database definitions, documentation, and integrity rules, i.e. the data schema. This definition is then operationalized as *conforming to specified data integrity rules* in order to serve as a practical basis for syntactic quality assessment. Essentially, this assumes that important requirements for conformance to definitions and documentation have been specified in terms of integrity rules.

The *semantic quality category* describes the degree to which stored data corresponds to represented external phenomena. The derivation of semantic quality criteria is based on Wand and Wang's [1996] work. They use ontological theory to formally define quality criteria describing real-world to IS transformations that are free of data deficiencies. As described in Price and Shanks [2004, 2005b], the list of criteria is amended for inclusion in the *InfoQual* framework to account for differences in goals and in the unit of analysis. The result describes a correct transformation as one where every external phenomenon that is relevant to the organization is represented consistently, unambiguously, and correctly without any meaningless (spurious) data (i.e. data that does not map to any external phenomenon of interest).

The Subjective IQ View: The Pragmatic Category and Criteria

The *pragmatic quality category* describes the degree to which stored data is suitable and worthwhile for a given use. Derivation of pragmatic criteria requires the use of empirical techniques to solicit consumer input on the appropriateness of the pragmatic criteria since by definition they relate to the subjective consumer perspective. Both extant literature and empirical methods were used to derive pragmatic criteria, as described in Price and Shanks [2004, 2005a, 2005b].

First, an initial set of criteria were derived based on an analytic review of literature guided by clearly delineated set of goals and requirements. For example, one requirement was that selected criteria must be general, i.e. applicable across application domains and data types. The resulting list was then refined using empirical techniques. In this context, focus groups were considered the preferred empirical technique because of their highly interactive nature, allowing for a full exploration of relevant (and possibly contentious) issues based on a direct exchange of views between participants. Three focus groups were conducted to solicit feedback from IT practitioners, IT academics, and end-users respectively, where participants of the first two groups had direct responsibility for or research interest in IQ. Participants were asked to evaluate the list of criteria and their definitions for clarity, validity (i.e. importance), completeness, and independence. The end-user group also served to clarify the vocabulary understood by end-users in the lead-up to instrument development. As a preliminary step to composing instrument items (i.e. questions), we observed end-user responses to the wording of criteria definitions and the vocabulary they used to describe quality concerns.

The resulting list of pragmatic criteria is shown in the last (third) section of Table 1 and in Figure 1 (with sub-criteria shown). Note that in Figure 1, Level 2 represents an additional level of detail (i.e. sub-criteria) that explicates specific aspects of Level 1 criteria. Thus *easy* and *quick* (i.e. easily and quickly accessible) comprehensively describe *accessible*. However, Level 1 criteria are not necessarily subsumed by Level 2 and thus may include aspects not explicitly described in Level 2. For example, the colour scheme and illumination level of information presentation may also influence presentation quality but did not figure significantly in the literature or focus group discussions and so do not warrant separate criteria. Thus *suitably presented* and *flexibly presented* are not subsumed by their subcriteria.

The first seven pragmatic criteria (listed in the relevant section of Table 1 and in the leftmost group of Level 1 criteria in Figure 1) pertain either to the delivery and/or the usability of the retrieved data. The remaining three pragmatic criteria relate to consumer *perceptions* of the syntactic and semantic criteria (so-labeled in Figure 1) described earlier. These are included because an information consumer's subjective and use-based judgment may differ considerably from objective and relatively use-independent measurement of the same quality criterion. For example, consumers may consider a data set to be incomplete (.e.g. based on their use of a new application requiring data not previously considered relevant) even though the same data set is judged to be complete using objective methods (e.g. comparison to a trusted but less recently compiled source).

With respect to perceptions of semantic criteria, the more general term *reliable* was used in place of the original more specific semantic mapping criteria described in the previous subsection on objective IQ. As evident from the individual opinion form and focus group feedback, it was quite difficult for respondents (especially end-users) to distinguish between the different semantic criteria and their corresponding mapping cardinalities as originally defined. Therefore, the term *reliable* was used instead. Respondents felt that this term was more intuitively understandable and could be used to summarize the group of criteria.

On the basis of the literature review and focus group feedback, Price and Shanks [2005b] concluded that the criteria shown in Figure 1 represent "essential and distinct" but inter-dependent aspects of the pragmatic category of information. Interdependencies between the values of pragmatic criteria relating to data delivery (such as *suitably* and *flexibly presented, understandable, timely, accessible*, and *secure*) are acknowledged by Price and Shanks [2005b]. For example, the presentation of data affects its perceived comprehensibility, accessibility, and security.

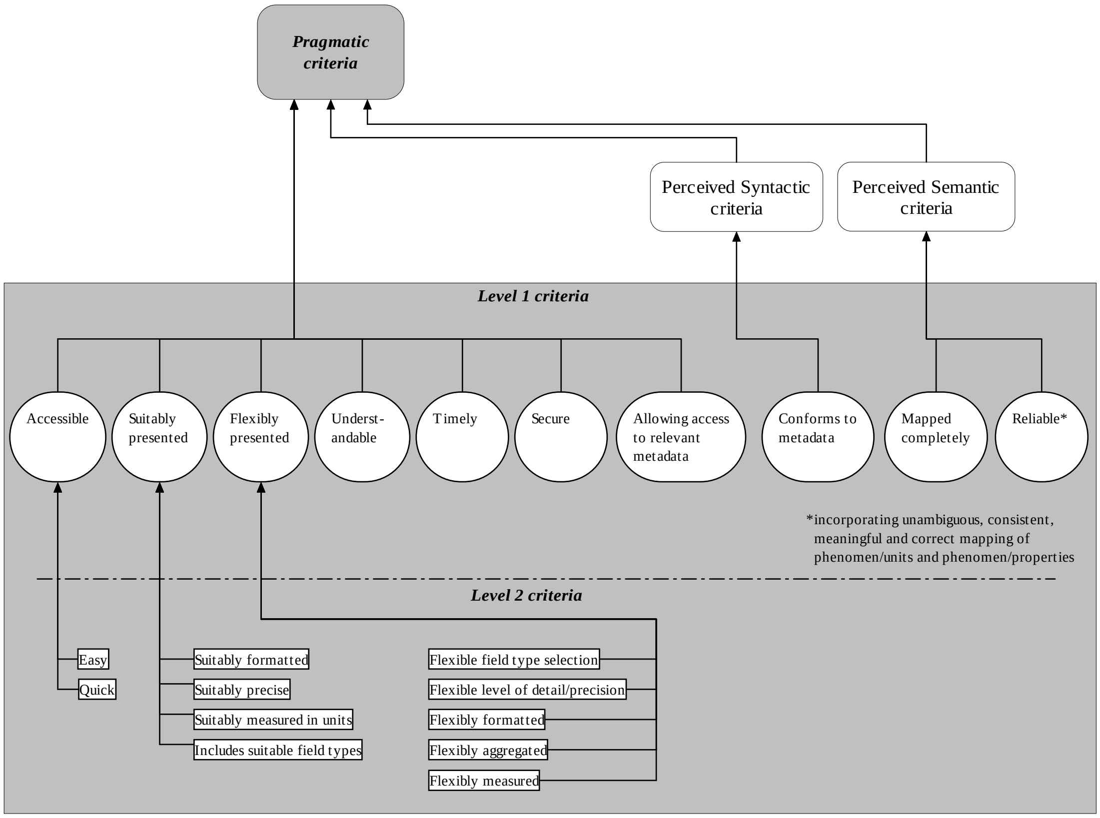

Figure 1. Pragmatic Criteria of the SIQ Framework

In the next section, we discuss the development of an initial measurement instrument based on the pragmatic criteria. The instrument is intended to allow assessment of subjective IQ for a given data set using the pragmatic criteria.

III. METHODOLOGY

Design of the Measurement Instrument for InfoQual's Pragmatic Category

As is often the case in the context of empirical research, the quality criteria themselves are not directly observable and are therefore referred to as *latent variables* or *latent constructs* or, in the context of factor analysis, *factors* or *dimensions*. Accordingly, a questionnaire (also referred to as the *measurement instrument* or simply the *instrument*) is developed with the aim of representing these *latent variables* in the form of measurable *observed variables* (also called *observed constructs, indicators,* instrument *items* or *questions*) for the purpose of measurement. In the remainder of this section, issues associated with the development of such instrument are discussed. We will use the terms *latent variables* or *factors* and *observed variables, items* or *questions* in the remainder of the paper.

Segars [1997] and others (e.g. [Hair et al. 1998, pp. 98]) suggest that "measurement of latent variables be accomplished through use of multi-item scales" with factor analysis then used to establish *construct validity* [Straub et al., 2004]. Accordingly, a five-item Likert scale that includes *strongly disagree*, *disagree*, *neutral*, *agree* and *strongly agree* items was used to measure user (i.e. information consumer) satisfaction with the quality of information in the context of the semiotic framework.

Having defined the measurement scale, the next issue that needs to be considered is the number of questions that are "intended to be alternative indicators of the same underlying construct [i.e. latent variable]" [Segars, 1997, pp. 2] to be included in the instrument. Hair et al. [1998, pp. 98] suggest at least five questions per criterion (i.e. proposed factor), while other authors (e.g. [Garson 2005; Segars 1997]) suggest that three questions are sufficient. Due to the large number of criteria, it was decided to include at least three rather than five questions for each of the second-level criteria (i.e. subcriteria) and for each of those first-level criteria that were not subsumed by their subcriteria (see the second subsection in Section II for further explanation).

The wording of questions in the measurement instrument was guided by previous research in assessment of user perception of and satisfaction with IQ (e.g. [Barnes and Vidgen 2002; Chin and Lee 2000; Wang and Strong 1996]), empirical testing of the semiotic framework that included focus groups involving in-depth discussions with IQ experts and with end-users[Price and Shanks 2005b], as well as general principles of good questionnaire design (e.g. [Cavanna et al. 2001; Groves 1989]). The goals in selecting questions were to ensure that they were:

- consistent with the definitions of the framework criteria;
- consistent with feedback from focus groups;
- short (less than 20 words) and clear;
- used consistent terminology;
- not double-barreled, recall-dependent, leading, or loaded; and
- adopted from other validated instruments whenever possible (i.e. without compromising the previous goals listed).

As recommended by Cavanna et al. [2001], questions were randomly ordered in the questionnaire to avoid bias; however, negative questions were interspersed with positive questions to avoid automatic response patterns at one end of the scale. The resulting 66-item questionnaire is included in Appendix 1, with questions that were taken directly from previously validated instruments footnoted accordingly.

The measurement instrument proposed in this section was developed to ensure content validity, in other words, the measures were chosen to "capture the essence" [Straub et al. 2004, pp. 386] of the quality criteria derived on the basis of literature review and focus groups, erring on the side of inclusion while acknowledging that some criteria may not meet *construct validity* requirements and will subsequently be excluded. Thus, second-level criteria are treated as separate latent variables in designing the original measurement instrument; although it is likely that they may be combined in the final factor solution. Similarly, user perceptions of semantic and syntactic quality aspects may not be retained in the final factor solution. While these criteria are treated as latent variables during initial instrument design; we acknowledge that users (particularly end-users) may not have access to, be concerned with, or understand the technical aspects of semantic and syntactic concepts.

Users are likely to view syntactic and semantic concepts through "pragmatic lenses," i.e., based on their understanding of the application domain and their experience of using the information. For example, an end-user does not know the actual integrity rules implemented but rather has a personal view of the application rules appropriate to their use of the data. Furthermore, syntactic criteria relating to the form of the information may be viewed through "semantic lenses." An example would be data that violates specified syntactic formatting rules but is still able to be matched correctly to the represented real-world phenomena by the end-user and thus is not viewed as being illegal. In fact, users can only judge the legality (i.e. syntax) or validity (i.e. semantic) of stored data in a database indirectly as reflected in the presentation of retrieved data. So if a user query retrieves an incomplete set of information they are likely to conclude that the information in the database is incomplete even when it is present and could be retrieved by a different query. Figure 2 illustrates the so-called "onion model" of a user's perception of information.

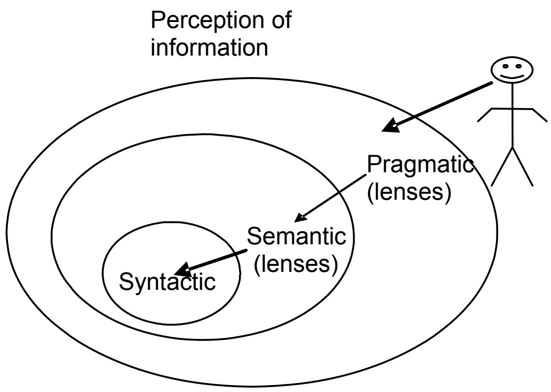

Figure 2. A User's View of Pragmatic Criteria (An Onion Model)

Price and Shanks [2005b] discuss the consequent possibility that objective measurements of syntactic and semantic quality may not match information consumer perceptions (i.e. subjective measures of the same criteria) and the potential value of being able to measure such discrepancies for identifying and solving IQ problems. A further implication is that the set of criteria used to describe objective IQ (i.e. in syntactic and semantic categories) may not be suitable for describing subjective perceptions of those criteria by information consumers (i.e. in the pragmatic category). In other words, if the layers in the onion model illustrated in Figure 3 have low level of transparency, it is likely that there will be a lack of *discriminant validity* between criteria measuring perceptions of syntactic and semantic categories and other pragmatic criteria.

Survey Design and Administration

The measurement instrument discussed in the previous section was used as the basis for an IQ survey aimed at collecting data to assess the *construct validity* and *reliability* of the proposed instrument. In addition to the questions, the survey included explanatory notes and provided respondents with the space to comment.

The number of questions included in the instrument dictates the minimum sample size necessary for multivariate data analysis. Hair et al. [1998, pp. 98] recommend that the sample size should be at least 100 and should have "at least five times as many observations as there are [observed] variables". This suggests that the sample size should exceed 300 valid responses to enable analysis of a survey based on a 66-question instrument. Garson [2005] notes that there are many different rules of thumb quoted in the literature for determining sample size but none have any sound theoretical justification. The rule adopted is said to be one of the most commonly used.

To facilitate a good response to the survey, the information consumer sample and associated data set were selected to ensure that survey respondents have a strong interest in the quality of the data they evaluated. Thus, the survey was designed to solicit university student feedback on the quality of enrolment and class allocation information regularly accessed by the students. Specifically, the Web Enrolment System (WES) and Class Allocation System (Allocate+) were used in this study. These systems are both Web-based systems and are used to store and manage personal and enrolment information about students.

The only information that students are able to access in WES is information that is personal (e.g. their own examination results or academic record), relevant to their enrolment (e.g. exam timetable for subjects they are enrolled in), or of a generic nature (e.g. payment options for fees). Similarly, in Allocate+ students can only view information about activities for those subjects in which they are enrolled. Both systems allow students to change some information about themselves (e.g. update their contact details or change allocation to a different tutorial). Other information is updated by university staff (e.g. examination results or lecture timetable). The status of some information is time-dependent. For example, after a certain cut-off point, students are unable to update their enrolment or allocate themselves to a tutorial.

The frequency of use of the two systems varies within the academic year. For example, Allocate+ is used intensively during the first three weeks of the semester (while students are settling their timetable) and after that it is used only occasionally by some students to look up their lecture or tutorial details. Similarly, WES is used intensively by students during the enrolment period, at the end of the semester to check examination details, and at the end of the examination period to look up results. The IQ survey was conducted at the end of the first semester. In order to minimize any non-sampling error associated with the ability to remember the information stored in the system (e.g. [Groves 1989]), screen snapshots were used to remind students about the types of information included in the two systems.

A survey of students during their class time was considered to be the most likely way of ensuring the number of responses required in the time available for this research. Seven lecturers within the faculty of Information Technology of Monash University allowed researchers to conduct the survey during their class time. Due to the relatively homogeneous nature of the student body surveyed (especially with respect to age and end-user role), the large number of questions in the original measurement instrument, and limitations with respect to the time allocated for survey administration, additional questions relating to demographic and summary information of respondents were not included in the survey.

Selection of student cohorts for participation in the survey was opportunistic and subject to the availability of lecture and tutorial time. The survey was distributed either in lecture or all of the tutorials of a first year subject, two second-year subjects, one third-year subject, and three postgraduate subjects of the information systems faculty undergraduate degrees. Questionnaires were distributed to students at the beginning of lectures or tutorials by one of the researchers and collected by the researcher either at the end of the session or on completion of the questionnaire (subject to the each lecturer's preference). Responses were anonymous: no identification information was included on the questionnaires. The researcher, tutor, and/or lecturer were present in class while students completed the surveys. This arrangement facilitated a high response rate (at least 85 percent of the student attendees responded) with a total of 402 students completing the surveys. Out of 402 completed surveys (150 WES and 252 Allocate+), 324 (81 percent) surveys (126 WES and 198 Allocate+) were found to have complete and valid responses that were included in the data analysis. Only 23 percent of valid responses included comments, with nearly two-thirds of these comments noting the repetitive nature of the questionnaire.

Due to the high rate of valid responses, the simultaneous collection of surveys, and the decision not to include demographic or other summary information about respondents in the survey, nonresponse bias analysis was not undertaken on the collected data. It is recommended that in further instrument testing using a cross-sectional survey (discussed as possible future work in Section VII), where a much lower response rate and less simultaneous survey collection is expected (and given the reduced size of the instrument after factor analysis), some demographic data about respondents (such as respondents' age and sex, experience with the organization, and information usage level) is also collected to facilitate such analysis (e.g. [Nelson et al., 2005]).

After the data were collected, negatively worded questions were converted to positive questions by reversing the order of the responses (e.g. a *strongly disagree* answer to a negatively worded question such as Question 37 was treated as a *strongly agree* answer to a positively converted Question 37) in order to simplify interpretation of data analysis.

IV. DATA ANALYSIS

The objective of the IQ survey described in the previous section is to provide data that can be used to identify and validate relationships between the pragmatic criteria of the *InfoQual* framework (i.e. proposed *latent variables*) and the measurement instrument expressed in the form of survey questions (i.e. *observed variables* or *items*). Accordingly, the purpose of data analysis is twofold. Firstly, to identify the underlying dimensions (i.e. *factors*) in the data that reflect the commonality between survey questions, and, secondly, to derive a measurement instrument that enables valid measurement of the pragmatic category within the *InfoQual* framework.

The majority of questions within the proposed instrument were not previously validated in the context of IQ theory or were validated in the context of substantially different theoretical frameworks. Accordingly, the methodology chosen for the analysis of the IQ survey data must be suitable for validation of an instrument in early stages of its development where the relationship between most observed variables and latent variables is either uncertain or completely unknown.

Other important considerations for the choice of the methodology are:

1. the presence of correlations between observed variables as a result of common underlying dimensions (i.e. *factors*);
2. the interdependencies between criteria as discussed in Section II;
3. the five-item Likert measurement scale that precludes the data from being normally distributed; and
4. the large number of observed variables (66) that are being considered and the relatively large sample size available for the analysis (324 cases).

Factor analysis techniques have been extensively used within the social and information sciences (e.g. [Cramer 2003; Straub et al. 2004]) in the context of validating measurement instruments. Of the two types of factor analysis techniques discussed in the literature, exploratory rather than confirmatory factor analysis is more suitable for use with an instrument in the early stages of its development [Byrne 2001, pp. 5, 99]. It has the further advantage of not requiring that data conform to statistical assumptions such as normality (see point number 3 above). A choice must then be made between two variants of exploratory factor analysis, principle components analysis and common factor analysis. Hair recommends that common factor analysis be used when the objective is to understand the "underlying factors or dimensions that reflect what [observed] variables share in common" [Hair et al. 1998, pp. 100], as is the case in this study. Common factor analysis is based on the variance that is shared by each observed variable with all other observed variables in the analysis (often referred to as *common variance*).

Factor analysis techniques require decisions to be made with respect to the number of factors to be extracted from the data and the rotation procedure used to assist with the interpretation of the factors. Latent root and scree test criteria, providing measurements of the amount of variance accounted for by each factor, usually dictate the initial choice of the number of factors to be extracted. The *latent root* or *eigenvalue*[3] *greater than 1* criterion is based upon the requirement "that any individual factor should account for the variance of at least a single variable if it is to be retained for interpretation" [Hair et al. 1998, pp. 103]. The *scree test* is "derived by plotting the latent roots [also

[3] These terms refer to the amount of variance in the data that is accounted for by each factor.

referred to as *eigenvalues*] against the number of factors in their order of extraction [in a scree plot] . . . the point at which the curve first begins to straighten out is considered to indicate the maximum number of factors to extract" [Hair et al. 1998, pp. 104]. Orthogonal or oblique rotation methods are used to reduce ambiguities in the factor structure. The oblique rotation method allows correlated factors and thus is more consistent with the underlying theoretical framework (see points number 1 and 2 above).

Interpretation of the rotated factor structure is guided by the size of factor loadings (indicating the degree of correspondence between observed variables and factors) in the rotated factor pattern matrix (Hair et al. 1998, pp. 106). The higher the absolute value of the loading of a observed variable on a factor, the more representative that variable is of the factor. While Straub et al. [2004] suggest that items with loadings of less than ±0.40 can be dropped from the instrument, Hair et al. [1998, pp. 111] state that "loadings greater than ±0.30 are considered to meet the minimal level." Hair et al. [1998, pp. 111-112] note that as the number of items and/or sample size increases, the acceptable level decreases. Given the large number of items and the large sample size in this study (see point 4 above), loadings greater than ±0.30 are considered significant and acceptable. To ensure discriminate validity of the factor solution, Hair et al. [1998] and Straub et al. [2004] recommend that items which load significantly (especially with similar size loadings) on multiple factors be eliminated from the instrument as they fail to discriminate between factors.

Further statistical criteria for determining the final number of factors are based on the composition of the factors. To facilitate interpretability of the solution, factors on which less than three items load significantly may be dropped (e.g. [Benamati and Lederer 1998]). Straub et al. [2004] also suggest that factors must be internally consistent (i.e. *reliable*) and propose that Cronbach's α coefficient of at least 0.6 can be used as a cut-off for the acceptable level of internal consistency within each factor.

Hair et al. [1998, pp. 128] acknowledge that while the optimal validation of factor analysis involves confirmatory factor analysis on an entirely new sample this approach is rarely feasible. Accordingly, they suggest that splitting the sample into two equal samples and re-estimating the factor model on the two samples provides a way of assessing stability of results. Hair et al. [1998, pp. 128] refer to this technique as *split sample analysis*.

As can be seen from the earlier discussion, the heuristics that are used to finalize the factor structure are somewhat arbitrary. Because of this, the most important criterion for finalizing the factor structure and retaining items is considered to be the conceptual soundness of the solution that can only be established in the context of the appropriate theoretical framework (e.g. [Hair et al. 1998, pp. 110, 114]).

To summarize, in the context of the methodological considerations discussed in the previous section, the following data analysis and heuristics have been selected for the analysis of the IQ survey in the context of the pragmatic category within the *InfoQual* framework.

Analysis

- Exploratory factor analysis using common factor analysis procedures with the initial number of factors to be extracted based on the latent root criterion, with scree test and interpretability criteria to be used to finalize the factor structure.
- Split sample analysis of the final factor structure.

Heuristic

- Items to be retained in the instrument must have a factor loading in the final structure that is at least ±0.30 and must be able to discriminate between factors;
- Each factor retained in the final structure must have at least three items loading significantly on it and must be internally consistent with a Cronbach's α coefficient of at least 0.60;
- The factor solution must be conceptually sound; and
- Factor solutions on the split samples must be comparable.

The SPSS for Windows package version 12.0.1 [SPSS Inc. 2005] is one of the statistical packages that is commonly used for factor analysis (e.g. [Hair et al. 1998]) and is able to produce the required heuristics. Results of data analysis using SPSS are summarized in the next section.

V. RESULTS

While acknowledging that "critical assumptions underlying factor analysis [are] more conceptual than statistical" [Hair et al. 1998, pp. 99], a number of authors (e.g. [Field 2000; Hair et al. 1998; Segars 1997]) provide guidelines for numerical assessment of survey data against factor analysis assumptions. These guidelines and corresponding assessment of the survey data are summarized in Table 2. As is evident from Table 2, data collected in the IQ survey satisfies the requirements of exploratory factor analysis methodology.

In addition to the minimum requirements described in Table 2, Hair et al. [1998] and others (e.g. [Schwab 2003]) recommend identification of potential multivariate outliers using the Mahalanobis distance measure. It is important to note that none of the authors recommend automatic exclusion of cases that are identified as potential outliers. Analysis of the 324 valid responses identified 40 potential outliers. A data entry error was found and corrected as a result of this analysis. Potential outliers were found to have a large number of "neutral" responses to the questions. These are considered legitimate satisfaction responses and therefore none of the potential outliers were excluded from the analysis. We now describe the results of the factor analysis in detail.

Table 2. Assessment of Factor Analysis Requirements for the IQ Survey

Requirement description (based on Hair et al. 1998, pp. 99)	**Empirical results**
Correlations present greater than 0.30	24.4% of correlations between survey questions were greater than 0.3
Small anti-image or partial correlations	Over 90% of partial correlations between survey items are less than 0.1, all are less than 0.2
Barlett test of sphericity	Barlett test is significant at 0.000 level confirming presence of non-zero correlations
Measure of sampling adequacy (MSA) for each survey question	MSA is greater than 0.8 for all survey questions, this suggests that all items can be included in the factor analysis
Overall Measure of sampling adequacy (MSA)	Overall MSA is 0.9, this suggests that at the overall level factor analysis is also "above meritorious"

Factor analysis procedures were applied iteratively to the data until the factor solution that satisfied all of the heuristics described in Section IV was achieved. The results of each iteration are summarized in Table 3.

As can be seen from Table 3, common factor analysis of the complete instrument with latent root criterion and oblique rotation (i.e. the first iteration) produced an initial 16-factor solution. In this solution, 13 questions (Q39, Q61, Q43, Q56, Q42, Q54, Q14, Q28, Q16, Q52, Q5, Q7, Q10) did not have any loadings greater than ±0.30 and were therefore dropped from the instrument and later analysis.

Factor analysis of the remaining 53 questions produced a 12-factor solution with five further questions (Q47, Q53, Q20, Q1, Q26) not loading significantly on any of the factors. Accordingly, these five items were also eliminated from the instrument. Factor analysis on the remaining 48 questions produced an 11-factor solution using the latent root criterion. However, examination of the scree plot for the point at which the curve levels off (indicating that little additional variance would be explained by adding additional factors as explained in Section IV) suggested that a six-to-eight factor solution may be more appropriate.

Table 3. Factor Analysis Iterations

Iteration	No. of factors decision	Factor solution (%variance explained)	Reason for exclusion	Questions excluded (underlined)
1 (66 questions)	MinEigen>1	16 (51.5)	Items with all loadings in [-0.3,0.3]	Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11Q12 Q13 Q14 Q15 Q16 Q17Q18 Q19 Q20 Q21 Q22 Q23Q24 Q25 Q26 Q27 Q28 Q29Q30 Q31 Q32 Q33 Q34 Q35Q36 Q37 Q38 Q39 Q40 Q41Q42 Q43 Q44 Q45 Q46 Q47Q48 Q49 Q50 Q51 Q52 Q53Q54 Q55 Q56 Q57 Q58 Q59Q60 Q61 Q62 Q63 Q64 Q65Q66
2 (53 items)	MinEigen>1	12 (50.241)	Items with all loadings in [-0.3,0.3]	Q1 Q2 Q3 Q4 Q6 Q8 Q9 Q11 Q12 Q13 Q15 Q17 Q18 Q19 Q20 Q21 Q22 Q23Q24 Q25 Q26 Q27 Q29 Q30 Q31 Q32 Q33 Q34 Q35Q36 Q37 Q38 Q40 Q41 Q44 Q45 Q46 Q47 Q48 Q49 Q50 Q51 Q53 Q55 Q57 Q58 Q59 Q60 Q62 Q63 Q64 Q65 Q66
3 (48 items)	MinEigen>1	11 (50.581)	1. Items with all loadings in [-0.3,0.3] (Q23) 2. Less than three items loaded on a factor (Q2, Q48, Q63) 3. Items fail to discriminate between factors (Q30)	Q2 Q3 Q4 Q6 Q8 Q9 Q11 Q12 Q13 Q15 Q17 Q18 Q19 Q21 Q22 Q23 Q24 Q25 Q27 Q29 Q30 Q31 Q32 Q33 Q34 Q35Q36 Q37 Q38 Q40 Q41 Q44 Q45 Q46 Q48 Q49 Q50 Q51 Q55 Q57 Q58 Q59 Q60 Q62 Q63 Q64 Q65 Q66
4 (43 items)	A priori 8 factors	8 (48.382)	Not applicable	None
5 (43 items)	A priori 7 factors	7 (46.7)	Items fail to discriminate between factors	Q3 Q4 Q6 Q8 Q9 Q11 Q12 Q13 Q15 Q17 Q18 Q19 Q21 Q22 Q24 Q25 Q27 Q29 Q31 Q32 Q33 Q34 Q35Q36 Q37 Q38 Q40 Q41 Q44 Q45 Q46 Q49 Q50 Q51

Table 3. Factor Analysis Iterations

Iteration	No. of factors decision	Factor solution (%variance explained)	Reason for exclusion	Questions excluded (underlined)
				Q55 Q57 Q58 Q59 Q60 Q62 Q64 Q65 Q66
6 (40 items)	A priori 7 factors	7 (46.707)	Items are loaded on factors that are not conceptually sound	Q3 Q4 Q6 Q8 Q9 Q11 Q12 Q13 Q15 Q17 Q18 Q19 Q21 Q24 Q25 Q27 Q29 Q31 Q32 Q33 Q34 Q35Q36 Q37 Q38 Q40 Q44 Q45 Q46 Q50 Q51 Q55 Q57 Q58 Q59 Q60 Q62 Q64 Q65 Q66
7 (36 items)	A priori 6 factors	6 (45.866)	Items fail to discriminate between factors	Q3 Q6 Q8 Q9 Q11 Q12 Q13 Q15 Q17 Q18 Q19 Q21 Q24 Q25 Q27 Q29 Q31 Q32 Q33 Q34 Q35Q36 Q37 Q38 Q40 Q44 Q45 Q46 Q50 Q51 Q55 Q57 Q59 Q60 Q62 Q65
8 (35 items)	A priori 6 factors	6 (45.958)	Items with all loadings in [-0.3,0.3]	Q3 Q6 Q8 Q9 Q11 Q12 Q13 Q15 Q17 Q18 Q19 Q21 Q24 Q25 Q27 Q29 Q31 Q32 Q33 Q34 Q36 Q37 Q38 Q40 Q44 Q45 Q46 Q50 Q51 Q55 Q57 Q59 Q60 Q62 Q65
9 (34 items)	A priori 6 factors	6 (46.537)	No further items excluded as all of the requirements for the factor analysis specified in the heuristics from Section IV are satisfied.	Q3 Q6 Q8 Q9 Q11 Q12 Q13 Q17 Q18 Q19 Q21 Q24 Q25 Q27 Q29 Q31 Q32 Q33 Q34 Q36 Q37 Q38 Q40 Q44 Q45 Q46 Q50 Q51 Q55 Q57 Q59 Q60 Q62 Q65

In the 11-factor solution, Q23 did not load significantly on any of the factors and was therefore excluded from the instrument. As most of the questions in this iteration did load significantly on at least one factor, the ability of the items to discriminate between factors and composition of the factors were also examined. As a result of this examination, it was found that three questions loaded significantly on multiple factors and three factors had less than three questions that loaded significantly on them. Of the three questions that had multiple significant loadings, only

Q30 had similar size loadings on multiple factors. Accordingly, Q30 was excluded from the instrument, while the other two questions were retained. Factors 6, 10, and 11 had less than three questions loaded on them, thus supporting a choice of an eight-factor solution consistent with the scree test. Items which only had significant loadings on these factors 6, 10, and 11 (Q63 on Factor 6 and Q2 and Q48 on Factor 10) were therefore excluded from the instrument.

Factor analysis with an eight-factor a priori statistical criterion (i.e. choice of the number of factors was determined in advance of the factor analysis run) was then conducted on the remaining 43 questions. Analysis of the resulting scree-plot and the factor loading pattern matrix suggested that a seven-factor solution would be more appropriate. Results of factor analysis with a seven-factor a priori statistical criterion suggested that three more questions should be excluded (Q22, Q41, Q49) as they failed to discriminate between factors.

The factor analysis with a priori seven-factor statistical criterion was then re-run on the remaining 40 questions. The results of this analysis were examined in the context of the *InfoQual* framework to ensure that factors were conceptually sound. As a result of this examination it was found that one of the factors (Factor 7) did not add to the interpretability of the solution. Questions that loaded significantly on Factor 7 were negatively worded versions of the positive questions that loaded significantly on other factors. For example, the most representative question of Factor 7 (loading 0.592)—question 58 ("It is not possible to customize information format")—is also a negative equivalent of the most representative question of Factor 4 (loading 0.755) —question 9 ("The format of retrieved information can easily be changed as needed"). Given that negatively worded questions can be confusing to respondents (as indicated by the higher variance shown in the questions' responses according to Parasuraman et al [1991]) and the meaning of Factor 7 is fully overlapped with other factors in the solution, it was decided to exclude all questions (Q4, Q58, Q64, Q66) that loaded significantly on factor 7 and limit the factor solution to six factors with 36 questions remaining.

Two more iterations of factor analysis with a priori six-factor statistical criterion (refer to Table 3) resulted in exclusion of question 35 (on the basis that it did not discriminate between factors) and question 15 (on the basis that none of its loadings were significant). The resulting final six-factor solution included 34 questions from the original instrument. All factors within this solution have been found to be internally consistent (i.e. *reliable*) while satisfying mandatory *discriminate*, *convergent* and *factorial validity* criteria as defined by Straub et al. [2004].

The final factor solution is summarized in Appendix 2 and Figure 3.

In Figure 3, abbreviated names of pragmatic criteria are listed for each of the questions included in the factor solution (for example, Q40 has the label of *flexible (measure)* that abbreviates the pragmatic criterion "flexibly presented, flexible level of detail/precision"). Appendix 2 includes expanded descriptions of the *InfoQual* criteria. Negatively worded questions that were converted to positive prior to the analysis are indicated with a *(cp)* notation both in Figure 3 and Appendix 2. The names of individual factors in the final factor solution were selected to be congruent with the questions (i.e. *items*) loading on each factor while relating as much as possible to the original framework. Thus, the term *complete* was used for Factor 3 since the majority of the questions loading on this factor related directly to this concept. Similarly, the first factor has been labeled *useful* as it incorporates various aspects of information suitability and usefulness to the respondent (e.g. timeliness (Q34), understandablity (Q11), and suitable format (Q33)).

In the final factor solution, all questions relating to information accessibility (Factor 6 in Figure 3) loaded negatively on Factor 6, which implies a negative interpretation of Factor 6 in survey responses (i.e. information is not accessible). For ease of presentation and without altering the meaning of the factor, this factor has been treated as a positive factor (i.e. information is accessible) by using absolute values of loadings in Figure 3 and Appendix 2. Correlations between factors were calculated using this positive interpretation.

As was expected, a number of factors were interdependent. Interdependency was indicated by the correlations between factors available in the SPSS common factor analysis output (refer to Table 4). Correlations between factors of greater than ±0.3 are shown in Figure 3.

Figure 3. Final Factor Solution

Table 4. Correlations between Final Factors (Correlations between factor 6 and other factors were changed to reflect the positive nature of this factor)

Factor	1. Useful	2. Flex. content	3. Complete	4. Flex. layout	5. Secure	6. Accessible
1. Useful	1	0.219	0.216	0.326	0.312	0.574
2. Flex content	0.219	1	-0.122	0.433	-0.005	0.123
3. Complete	0.216	-0.122	1	0.029	0.177	0.302
4. Flex layout	0.326	0.433	0.029	1	0.187	0.279
5. Secure	0.312	-0.005	0.177	0.187	1	0.467
6. Accessible	0.574	0.123	0.302	0.279	0.467	1

The final factor structure was re-estimated on two samples (162 respondents each) derived by randomly splitting the original sample in half. Differences of note between these samples were:

- Changes to factor loadings for Q36 and Q37 (sample 1) that caused these questions to load on factors 6 (accessible) and 3 (complete) respectively instead of their original loading on factor 1 (useful); however, both questions retained significant loadings on factor 1;
- Reductions to the size of factor loadings for Q29 (sample 1) from 0.357 to 0.299 on Factor 1 (useful) and Q57 (sample 2) from 0.398 to 0.270 on Factor 3 (complete);
- Reversal of signs on some factors that did not affect the interpretation of the relationship between questions or factors.

Since these differences did not affect the interpretation of the underlying factor structure described in Figure 3 and Appendix 2, it was concluded that this structure is stable within the sample. The next section provides a detailed discussion of the results and their implications for the final factor solution and *SIQQ* measurement instrument, presented at the end of the section.

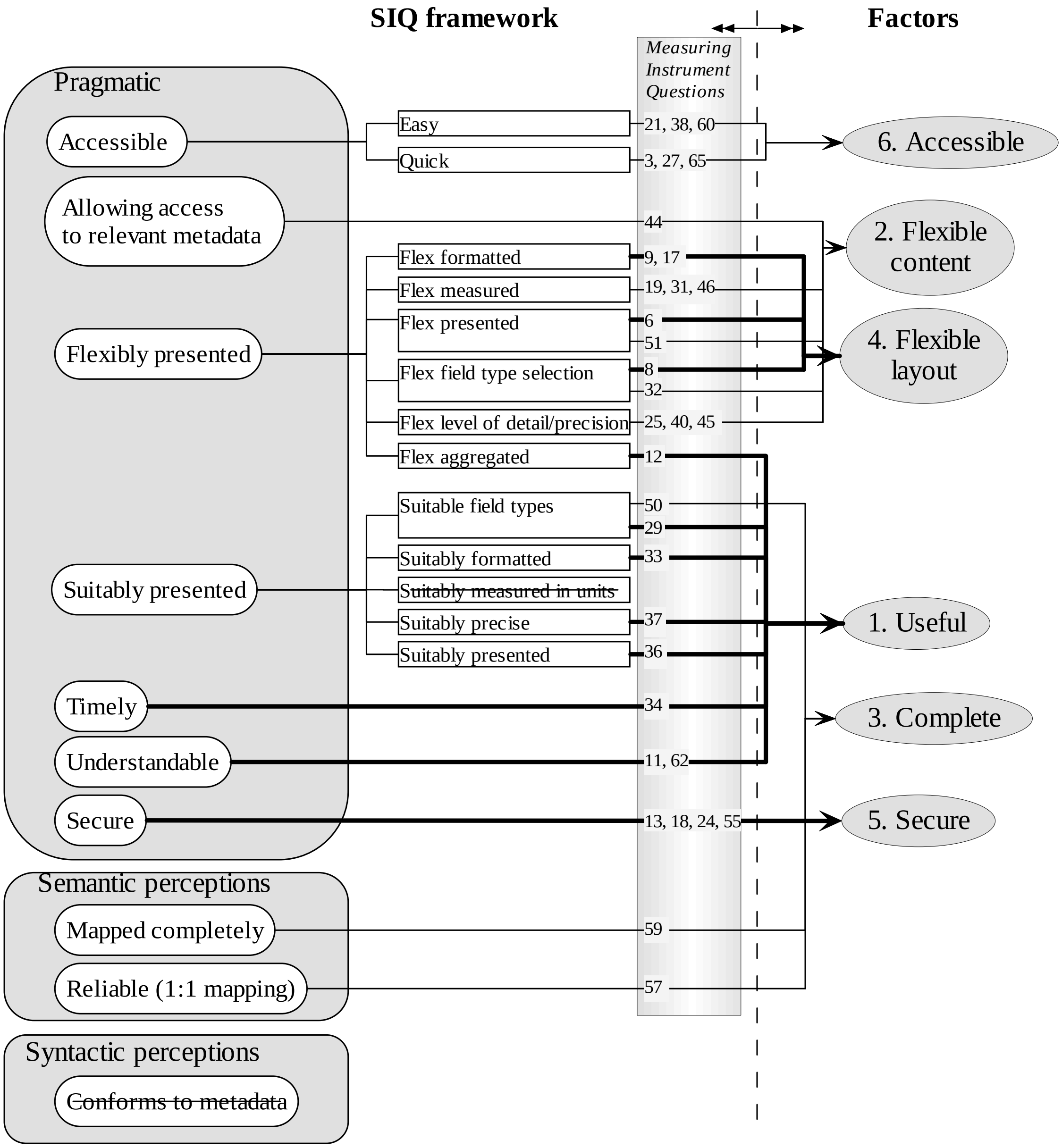

Figure 4. Correspondence between SIQ Framework and Final Factor Solution

VI. DISCUSSION

The comparison between the originally proposed factors for subjective IQ (i.e. based on the pragmatic category of *InfoQual*) and the final factor solution is shown in Figure 4. The pragmatic criteria and sub-criteria of *InfoQual* (i.e. proposed, theoretically-derived factors or latent variables, reproduced from Figure 1 in Section II) are displayed in the two left sections of Figure 4 respectively. Moving right, the next section of Figure 4 shows the numbers of those measurement instrument questions remaining after factor analysis (i.e. the observed variables used to operationalize the latent variables, refer to Appendices 1 and 2 for the question text associated with each question number). Finally, the rightmost section of Figure 4 displays the final factor solution (i.e. empirically-derived latent variables, reproduced from Figure 3). The correspondence of the final measurement instrument questions with the proposed and final factors they represent is shown by connecting lines.

The most noteworthy result of the empirical study is the reduction of the 21 *InfoQual* criteria (i.e. proposed *latent variables* or *factors*) that were used to define the measurement instrument within the pragmatic category of the *InfoQual* framework (10 first-level and 11 second-level) to six factors in the factor structure as illustrated in Figure 4. Given that the stated aim in constructing the measurement instrument was to err on the side of inclusiveness (see Section III), some reduction is to be expected. In particular, the reduction to first-level criteria and the combination of perceived syntactic and semantic quality aspects with other pragmatic aspects was anticipated (see Section III). In general, the reduction in the number of criteria (i.e. proposed factors) is supported by the IQ survey respondents' persistent comments regarding the repetitive nature of the questionnaire. Given that, at most, three questions were designed to measure the same concept, these comments suggest that users are unable to differentiate between the nuances of question wording that were used to separate between related concepts on the basis of the focus group and expert feedback. This is not surprising given the differences in the mode of data collection used to refine the pragmatic category criteria (i.e. focus groups) and the survey.

Within a survey, respondents read and answer questions quickly, often without taking the time to understand or think about the questions. For example, a respondent may not be able to distinguish between question 28, "The measurement units used for information are appropriate for your needs" and question 31, "The units of measurement used for retrieved information can be easily changed as needed" even though these questions reflect different criteria (suitability (Q28) and flexibility (Q31)) within the framework.

The combination of criteria (shown in Figure 4) in the final factor solution is discussed in the first sub-section. The representation of syntactic and semantic perceptions and the representation of pragmatic criteria in the final measurement instrument are discussed in the next two subsections respectively. The last subsection presents the final factor structure and measurement instrument.

Combination of Criteria in the Factor Solution

The anticipated outcome (as discussed in Section III) of the reduced structure is the combination of the second-level criteria within a single factor. For example, the results of the empirical study suggest that all of the second-level criteria for *flexibly presented* and *suitably presented* be treated as constructs within broader, higher-level criteria, as shown in Figure 4. This finding is consistent with other empirical studies, such as for example, an empirical examination by Nelson et al. [2005]. Similarly, the combination of second-level criteria — *easily accessible* and *quickly accessible* — into the single broader first-level criterion *accessible* is consistent with the framework definition.

Other anticipated combinations relate to the "onion model" of transparency discussed in Section III. The combination of the two first-level criteria relating to semantic perceptions, *mapped completely* and *reliable* (i.e. mapped reliably), with the second-level criterion *includes suitable field types* is a reflection of the extent to which users view semantic concepts through "pragmatic lenses"—as anticipated in Section III. This is discussed further in the second (following) subsection, in the context of the representation of syntactic and semantic perceptions in the final measurement instrument.

A further change from the factor structure initially proposed was that the first-level criterion *allowing access to relevant metadata* was combined in the factor structure with criteria relating to aspects of flexible presentation (and thus subsumed) in the final factor *flexible content*. Since the concept of metadata relates to the syntactic quality category, this result may be another manifestation of users viewing syntactic concepts through "pragmatic lenses."

We note that the initial version of *InfoQual* [Price and Shanks 2004a] did not include the criterion *allowing access to relevant metadata*. The subsequent suggestion to add this criterion came from an academic focus group participant during the process of empirically refining the initial literature-based list of pragmatic criteria. The surveyed end-users clearly do not share this view. This result may be a consequence of the particular databases and participant cohorts surveyed. Since the fields in the databases selected (i.e. WES and Allocate+) were well-understood by the survey participants (i.e. students), participants may not have fully understood or related to the questions based on this criterion. For example, students may not have distinguished either between the *availability* and *presentation* of metadata or between the presentation of information *content* and *metadata.* Thus the survey questions relating to metadata access may have been interpreted as the ability to change the presentation to include metadata such as documentation of the abbreviations or codes used for a particular data field.

As in the case of the criterion *access to metadata,* the demotion of *timely* (and its combination with presentation aspects of quality) observed in the final factor solution is congruent with the initial version of *InfoQual* [Price and Shanks 2004]. There, *timely* is included only as a subcriterion of the criterion *suitably presented*. Its consequent elevation to a separate criterion was based on suggestions by IT professionals in the business and academic focus groups. The contrasting view evidenced by surveyed end-users is consistent with an empirical study reported recently by Nelson et al. [2005]. They found that *currency* (the dimension incorporating aspects of timeliness) was not a significant effect within the data warehouse information context. (e.g. [Nelson et al. 2005])

The observed combination of presentation aspects of quality with the *understandable* criterion is consistent with the findings of recent empirical studies (e.g. [Nelson et al. 2005]) and can be explained by the interdependencies acknowledged in Section II. For example, the definition of the *format* dimension within the Nelson, Todd, and Vixom (NTV) framework [Nelson et al. 2005] incorporates aspects of both the *understandable* and *suitably presented* criteria that have been combined within the *useful* factor in our study.

Representation of Syntactic and Semantic Perceptions in the Instrument

The final instrument does not include any of the items (i.e. questions) relating to syntactic perceptions and only two out of eight items relating to semantic perceptions. (Thus, the first-level *InfoQual* criterion relating to syntactic perceptions is not represented as a final factor and the two first-level *InfoQual* criteria relating to semantic perceptions have a combined representation in the final factor solution.) These results clarify the degree of transparency in the "onion model" of the IQ discussed in Section III. The absence of items relating to perceptions of syntactic quality (i.e. survey questions derived from the *InfoQual* syntactic category criterion) suggests that users do not "see" these items in their own right but perceive them only through aspects of pragmatic and semantic quality. For example, it is likely that an end-user would not normally be able to distinguish *incorrectly* formatted information (i.e. violating syntactic integrity rules relating to format) separately from *inappropriately* formatted information (i.e. information not suitably presented). Thus only the latter aspect (i.e. *Information is presented in an appropriate format*) is retained in the final instrument.

Similarly, the user's view of the semantic layer (i.e. semantic quality perceptions) is influenced by the pragmatic layer (i.e. use and delivery quality aspects). Thus users do not distinguish between missing information field *types* and missing information *values* or between *presentation* and *existence* of information. Hence the item *Some types of information that you need are missing* (derived from the second-level criteria *suitable field types*) loads with the item *The information is missing some required values* (derived from the first-level criterion *mapped completely*). This suggests that subjective measures of IQ relating to the perceived semantic criteria within the pragmatic category in the *InfoQual* framework are either influenced by or cannot be differentiated from the other pragmatic criteria in the framework and relate primarily to the degree of user satisfaction with information. Therefore, when the primary purpose of an assessment is evaluation of the syntactic or semantic quality of information; objective measures (e.g. integrity checking or matching of stored data with external phenomena) rather than subjective measures (e.g. a survey of information users' perceptions) measures should be used.

Representation of Pragmatic Criteria in the Instrument

Importantly, the final measurement instrument incorporates items from every first- and second-level criterion in the pragmatic category with a single exception at each level. The first-level syntactic criterion *conforms to metadata* is not represented. This was expected, as discussed in the preceding subsection and in Section III. Although no items are retained for the second-level criterion *suitably measured in units*, the second-level criteria *flexibly measured* (in units) and *flexible level of detail/precision* is represented in the final instrument (in the *flexible content* factor). Essentially, if units can be presented flexibly they can be presented suitably; thus the latter may be seen as subsumed in the former. Furthermore, the overall suitability of the information is part of the *useful* factor. Therefore, it is not surprising that the suitability of measurement units was found to be an insignificant construct within the framework. These results are consistent with the collapsing of second-level criteria into the broader concepts of IQ, as discussed in the first subsection discussing the combination of criteria in the factor solution.

The Final Factor Structure and Measurement Instrument

Figure 5 shows the final factor solution (reproduced from Figure 3). As discussed in the first subsection, the *Useful* factor represents a combination of the first-level pragmatic criteria *timely, understandable,* and *suitable.* Similarly, the *Complete* factor represents a combination of the first-level pragmatic criteria *reliable* and *complete*. Ideally, the combined factor should have a name different from but representative of the set of contributing criteria (as is true for the *Useful* factor). However, it is difficult to find a word that clearly and unambiguously represents the combination of *Complete* and *Reliable.* Words commonly used in the literature (e.g. trustworthy, believable, reputable) are so broad in meaning as to be subject to misinterpretation. Therefore, we have retained the factor name *Complete* (see Section V) since the majority of items loading on the factor were from the pragmatic criterion *complete.* Finally, we note that in further testing of the instrument using a cross-sectional survey (discussed in Section III); it would be worth considering both the given factor solution and an alternate solution that combines the factors *flexible content* and *flexible layout* (into a single *flexible presentation* factor). This recommendation is based on the conceptual soundness requirement (referred to as the most important criterion for determining the factor structure in Section IV), the high degree of conceptual overlap shown in Figure 4, the high correlation between the two factors (see Figure 3), and the improved congruence with the original theoretical framework.

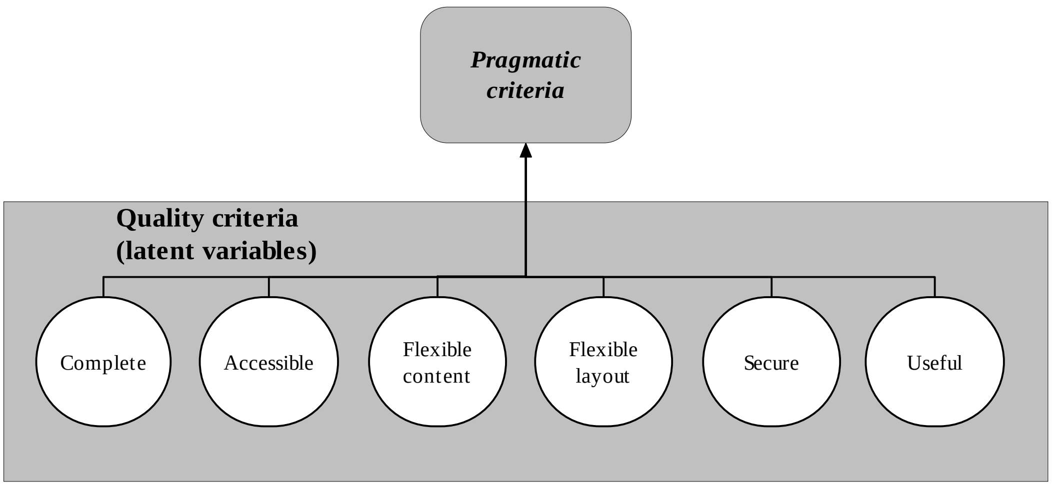

Figure 5. Final Factor Solution

The resulting measurement instrument, the *Subjective IQ Questionnaire* (*SIQQ*), is shown in Appendix 3, with the total number of questions reduced from 66 in the originally proposed instrument to 34 after factor analysis.

VII. CONCLUSIONS AND FUTURE WORK

This paper reports the development of an instrument intended to measure subjective aspects of IQ based on information consumer perceptions. The development consisted of two stages, reported in detail in this paper as an aid to others planning to use factor analysis to develop an instrument. Initial instrument design was based on the theoretically-grounded semiotic IQ framework *InfoQual.* This was followed by a survey using exploratory factor analysis for instrument refinement and validation.

The resulting factor solution was consistent with our expectations that second-level criteria would be combined into a single factor. It was further consistent with our theorized "onion" model of user perceptions—that users view conformance and correspondence concepts (i.e. represented by syntactic and semantic criteria) indirectly in terms of presentation concepts (i.e. represented by pragmatic criteria). Essentially, information consumers view quality in terms of data "use" rather than "form" or "content". Furthermore, with a single exception at each level (explained by the onion model and the second-level criteria combinations respectively), all of the first- and second-level pragmatic criteria from the *InfoQual* framework are represented directly in the final instrument. The final validated instrument, the *Subjective IQ Questionnaire* (*SIQQ*), consists of 34 questions.

In future work, a cross-sectional survey (such as conducted in [Nelson et al. 2005; Ewing and Napoli 2005]) should be conducted for further validation to address limitations in the current study with respect to diversity of information types and users.

SIQQ provides a measurement tool supporting assessment of the pragmatic category of the *InfoQual* framework, and in conjunction with objective measures such as integrity or correspondence checks, allows organizations to comprehensively assess the quality of their information. The advantage of such a bipartite approach to assessment is that subjective and objective views of quality can be compared for the same data set. This has the potential to facilitate both problem identification and problem source analysis (e.g. through highlighting discrepancies between objective assessments and information consumer perceptions). *SIQQ* can be used to assess and compare the views of different types of stakeholders in an organization. The detailed record of the instrument's development process, including the rationale for decisions made in terms of relevant design considerations, can potentially aid others intending to develop an instrument. In combination, the *InfoQual* framework and associated instrument *SIQQ* support organizations in understanding, monitoring, and identifying problems in IQ.

ACKNOWLEDGEMENTS

An Australian Research Council Discovery Grant funded this research.

LIST OF REFERENCES

Editor's Note: The following reference list contains hyperlinks to World Wide Web pages. Readers who have the ability to access the Web directly from their word processor or are reading the paper on the Web, can gain direct access to these linked references. Readers are warned, however, that:
1. These links existed as of the date of publication but are not guaranteed to be working thereafter.
2. The contents of Web pages may change over time. Where version information is provided in the References, different versions may not contain the information or the conclusions referenced.
3. The author(s) of the Web pages, not AIS, is (are) responsible for the accuracy of their content.
4. The author(s) of this article, not AIS, is (are) responsible for the accuracy of the URL and version information.

Barnes, S. J. and R. T. Vidgen. (2002). "An Integrative Approach to the Assessment of E-Commerce Quality," *Journal of Electronic Commerce Research* (3) 3, pp. 114-127.

Benamati, S. and A. L. Lederer. (1998). "Coping with Rapid Change in Information Technology," *Proceedings of the 1998 ACM SIGCPR Conference on Computer Personnel Research*, ACM Press, pp. 37-44.

Byrne, B. M. (2001). *Structural Equation Modelling with AMOS: Basic Concepts, Applications and Programming*, New Jersey: Lawrence Erlbaum Associates.

Cavanna, R. Y., B. L. Delahaye, and U. Sekaran. (2001). *Applied Business Research: Qualitative and Quantitative Methods*, Queensland, Australia: John Wiley & Sons.

Chin, W. W., and M. K. O. Lee. (2000). "A Proposed Model and Measurement Instrument for the Formation of IS Satisfaction: The Case of End-User Computing Satisfaction," *Proceedings of the International Conference on Information Systems 2000*, Brisbane, Queensland, Australia., pp. 553-563.

Cramer, D. (2003). *Advanced Quantitative Data Analysis*, Berkshire, England: McGraw-Hill Education.

English, L. P. (1999*). Improving Data Warehouse and Business Information Quality: Methods for Reducing Costs and Increasing Profits*, New York: Wiley Computer Publishing.

Eppler, M. J. (2001). "The Concept of Information Quality: An Interdisciplinary Evaluation of Recent Information Quality Frameworks," *Studies in Communication Sciences* 1, pp. 167-182.

Ewing, M. T. and J. Napoli. (2005). "Developing and Validating a Multidimensional Nonprofit Brand Orientation Scale," *Journal of Business Research* 58, pp. 841-853.

Field, A. (2000). *Discovering Statistics Using SPSS for Windows*, London: SAGE Publications.

Garson, D. (2005). "Factor Analysis," http://www2.chass.ncsu.edu/garson/pa765/factor.htm (Last accessed 3/8/2005)

Groves, R. M. (1989). *Survey Errors and Survey Costs*, New York: John Wiley & Sons Inc.

Hair, J. F. Jr., R. E. Anderson, R. L. Tatham, and W. C. Black. (1998). *Multivariate Data Analysis, 5th ed*., New Jersey: Prentice-Hall International Inc.

Kahn, B. K., D. M. Strong, and R. Y. Wang. (2002). "Information Quality Benchmarks: Product and Service Performance," *Communications of the ACM* (45) 4, pp. 184-192.

Lee, Y. W., D. M. Strong, B. K. Kahn, and R. Y. Wang. (2002). "AIMQ: A Methodology for Information Quality Assessment," *Information and Management* 40, pp. 133-146.

Morris, C. (1938). "Foundations of the Theory of Signs," in *International Encyclopedia of Unified Science*, vol. 1, London: University of Chicago Press.

Nelson, R. R., P. A. Todd, and B. H. Wixom. (2005). "Antecedents of Information And System Quality: An Empirical Examination within the Context of Data Warehousing," *Journal of Management Information Systems* (21) 4, pp. 199-235.

Parasuraman, A., L. L. Berry, and V. A. Zeithaml. (1991). "Refinement and Reassessment of the SERVQUAL Scale," *Journal of Retailing* (67) 4, pp. 420-450.

Peirce, C. S. (1931-1935). *Collected Papers*, Cambridge: Harvard University Press.

Price, R. and G. Shanks. (2004). "A Semiotic Information Quality Framework," *Proceedings of the IFIP International Conference on Decision Support Systems (DSS2004)*, Prato, Italy, pp. 658-672.

Price, R. and G. Shanks. (2005a). "Empirical Refinement of a Semiotic Information Quality Framework," *Proceedings of Hawaii International Conference on System Sciences (HICSS38)*, Big Island, Hawaii, USA: IEEE Computer Society, pp. 1-10.

Price, R., and G. Shanks. (2005b). "A Semiotic Information Quality Framework: Development and Comparative Analysis," *Journal of Information Technology* (20) 2, pp. 88-102.

Redman, T. C. (1996). *Data Quality for the Information Age,* Boston: Artech House.

Segars, A. H. (1997). "Assessing the Unidimensionality of Measurement: A Paradigm and Illustration within the Context of Information Systems Research," *Omega, International Journal of Management Science* (25) 1, pp. 107-121.

Schwab, A. J. (2003). "Principal Component Analysis: Additional Topics," http://www.utexas.edu/courses/schwab/sw388r7/SolvingProblems/PrincipalComponentAnalysis_Outliers_Validation_Reliability.ppt#457,115,Requesting%20output%20for%20the%20first%20validation%20sample (Last accessed 3/8/2005).

SPSS Inc. (2005). "Welcome to SPSS.com," http://www.spss.com (Last accessed 3/8/2005).

Straub, D., M.-C. Boudreau, and D. Gefen. (2004). "Validation Guidelines of IS Positivist Research," *Communications of the Association for Information Systems* 14, pp. 380-426.

Wand, Y. and R. Y. Wang. (1996). "Anchoring Data Quality Dimensions in Ontological Foundations," *Communications of the Association for Information Systems* (39) 11, pp. 86-95.

Wang, R. Y. and D. M. Strong. (1996). "Beyond Accuracy: What Data Quality Means to Data Consumers," *Journal of Management Information Systems* (12) 4, pp. 5-34.

APPENDIX I. INITIAL QUESTIONNAIRE (BEFORE FACTOR ANALYSIS)

Pragmatic criterion (level 1)	***Pragmatic Criterion (level 2)***	***Survey question number and wording***
Quickly accessible		3. Information is quick to retrieve.
		27. You can get information quickly.
		65. [W]Information is quickly accessible when needed.
Easily accessible		21. Information is easy to find and retrieve.
		38. [W] Information is easily accessible.
		60. [W]Information is easily obtainable.
Allowing access to relevant metadata		7. There are no definitions available for the terminology or codes used for different types of information.
		41. Information is appropriately documented.
		44. It is easy to find explanations of terminology, abbreviations, codes, or formatting conventions used in presenting information.
Flexibly presented	*Flexible field type selection*	8. The types of information presented can be easily changed as needed.
		15. It is not easy to customize the types of information shown.
		32. It is easy to modify the types of fields displayed.
	Flexible level of detail/precision	25. The level of detail presented can be easily changed as needed.
		40. The level of detail or precision for information can be modified to suit your needs.
		45. The precision of numeric fields cannot be customized.
	Flexibly aggregated	12. Information can easily be collated.
		39. [W]Information is easy to combine with other information.
		42. [W]Information is difficult to aggregate.
	Flexibly formatted	9. The format of retrieved information can easily be changed as needed.
		17.Information layout can easily be modified as required.
		58. It is not possible to customize information format.
	Flexibly measured	19. Measurement units can be customized.
		31. The units of measurement used for retrieved information can be easily changed as needed.
		46. It is not easy to change the measurement units used to display information.
		6. The display of information can easily be changed as needed.
		43. [W]The information is easy to manipulate to meet our needs.
		51. Information output can be customized as required.
		64. It is difficult to modify the presentation of information.
Secure		13. Information can only be modified by people who should be able to modify it.
		18. Information is appropriately protected from damage.
		24. [W]Information is not protected with adequate security.
		55. [W]Information is protected against unauthorized access.
Suitably presented	*Suitable field types*	29. The types of information presented are suitable for your needs.
		50. Some types of information that you need are missing. (Ex. tutorial duration not given)
		54. The information includes appropriate fields. (Ex. tutorial duration given)
		61. The types of fields presented are not useful.

	Suitably formatted	4. The layout of information output is not suitable for your needs.
		33. [B]Information is presented in an appropriate format.
		49. [C]The information is presented in a useful format
	Suitably measured in units	5. Presented information use suitable units of measurement. (Ex. tutorial duration expressed in minutes or hours not seconds)
		16. The measurement units used for displayed information are not suitable for your needs. (Ex. tutorial duration given in seconds)
		28. The measurement units used for information are appropriate for your needs (Ex. tutorial duration is in minutes or hours not seconds)
	Suitably precise	37. Information is not presented at the appropriate level of detail or precision.
		53. The precision of numeric information is suitable for your needs.
		56. [B]Information is provided at the right level of detail.
		22. Information output is displayed in an appropriate manner.
		36. The display of retrieved information is suitable for your needs.
		66. Information is not well-presented.
Timely		10. The information presented is too old or too recent to be useful.
		34. The currency (age or date) of the information is suitable for your needs.
		52. [B]The information provided is timely.
Understandable		11. [W]It is easy to interpret what this information means.
		30. [W]The meaning of the information is difficult to understand.
		62. [W]Information is easy to comprehend.
Semantic perception, mapped completely		20. [W]The information includes all necessary values.
		35. All the information values you need are available.
		59. The information is missing some required values.
Semantic perception, reliable (i.e. individual phenomenon /DB units have correct 1:1 mapping, mapped unambiguously & phenomena/properties mapped correctly, mapped consistently, mapped meaningfully)		2. [W]The information is reliable.
		23. [W]The information is free of errors.
		48. [C]The information is dependable.
		57. There is duplicate or inconsistent information.
		63. [W]The information is incorrect.
Syntactic, conforms to metadata		1. The information *content* is consistent with organizational or common-sense rules. (Ex. tutorial times are between 8am and 8pm).
		14. Information does not follow organizational rules and standards. (Ex. a tutorial is scheduled for midnight).
		26. Presented information follows standard rules and conventions. (Ex. tutorial dates consist of a day from 1 to 31 followed by a month from 1 to 12).
		47. The information *format* is consistent with organizational or common-sense standards and conventions. (Ex. subject codes begin with letters describing the school)

B - sourced from [1]
C - sourced from [5]
W - sourced from [20]

APPENDIX II. FINAL FACTOR SOLUTION

Note questions that were originally negatively worded are marked as *(cp)*.

Factor 1 *Useful*: Information is useful (Cronbach's α=0.846)

Original pragmatic Category	Question Number & Wording	Factor loading
Timely	34. The currency (age or date) of the information is suitable for your needs.	0.644
Understandable	11. It is easy to interpret what this information means.	0.617
Suitably presented, suitably formatted	33. Information is presented in an appropriate format.	0.586
Understandable	62. Information is easy to comprehend.	0.566
Flexibly presented, flexibly aggregated	12. Information can easily be collated.	0.485
Suitably presented, suitable precise	37 (cp). Information is not presented at the appropriate level of detail or precision.	0.416
Suitably presented	36. The display of retrieved information is suitable for your needs.	0.367
Suitably presented, suitable field types	29. The types of information presented are suitable for your needs.	0.357

Factor 2 *Flex content*: Information content (e.g. measurement units, level of detail, level of precision) can be customised (Cronbach's α=0.846)

Original pragmatic Category	Question Number & Wording	Factor loading
Flexibly presented, flexibly measured	31. The units of measurement used for retrieved information can be easily changed as needed.	0.698
Flexibly presented, flexibly measured	19. Measurement units can be customized.	0.675
Flexibly presented, flexible level of detail/precision	40. The level of detail or precision for information can be modified to suit your needs.	0.674
Flexibly presented, flexible level of detail/precision	25. The level of detail presented can be easily changed as needed.	0.624
Flexibly presented, flexibly measured	46 (cp). It is not easy to change the measurement units used to display information.	0.557
Flexibly presented, flexible field type selection	32. It is easy to modify the types of fields displayed.	0.552
Flexibly presented, flexible level of detail/precision	45 (cp). The precision of numeric fields cannot be customized.	0.495
Allowing access to relevant metadata	44. It is easy to find explanations of terminology, abbreviations, codes, or formatting conventions used in presenting information.	0.399
Flexibly presented	51. Information output can be customized as required.	0.397

Factor 3 *Complete*: Information is complete (Cronbach's α=0.669)

Original pragmatic Category	Question Number & Wording	Factor loading
Semantic perception, mapped completely	59 (cp). The information is missing some required values.	0.881
Suitably presented, suitable field types	50 (cp). Some types of information that you need are missing. (Ex. tutorial duration not given)	0.635
Semantic perception, reliable	57 (cp). There is duplicate or inconsistent information.	0.398

Factor 4 *Flex layout*. Information layout can be customised (Cronbach's α=0.806)

Original pragmatic Category	Question Number & Wording	Factor loading
Flexibly presented, flexibly formatted	9. The format of retrieved information can easily be changed as needed.	0.733
Flexibly presented	6. The display of information can easily be changed as needed.	0.665
Flexibly presented, flexible field type selection	8. The types of information presented can be easily changed as needed.	0.591
Flexibly presented, flexibly formatted	17. Information layout can easily be modified as required.	0.538

Factor 5 *Secure*: Information is secure (Cronbach's α=0.672)

Original pragmatic Category	Question Number & Wording	Factor loading
Secure	18. Information is appropriately protected from damage.	0.631
Secure	55. Information is protected against unauthorized access.	0.629
Secure	24 (cp). Information is not protected with adequate security.	0.527
Secure	13. Information can only be modified by people who should be able to modify it.	0.444

Factor 6 *Accessible*: Information is accessible (Cronbach's α=0.865)

Original pragmatic Category	Question Number & Wording	Factor loading
Quickly accessible	27. You can get information quickly.	0.746
Easily accessible	60. Information is easily obtainable.	0.716
Easily accessible	21. Information is easy to find and retrieve.	0.654
Quickly accessible	65. Information is quickly accessible when needed.	0.621
Easily accessible	38. Information is easily accessible.	0.619
Quickly accessible	3. Information is quick to retrieve.	0.545

APPENDIX III. FINAL QUESTIONNAIRE (AFTER FACTOR ANALYSIS)

Note that the questions are ordered according to the guidelines described in Section III of the paper.

The Subjective Information Quality Questionnaire (SIQQ)

1. The currency (age or date) of the information is suitable for your needs.
2. The units of measurement used for retrieved information can be easily changed as needed.
3. The information is missing some required values.
4. The format of retrieved information can easily be changed as needed.
5. Information is appropriately protected from damage.
6. You can get information quickly.
7. Measurement units can be customized.
8. It is easy to find explanations of terminology, abbreviations, codes, or formatting conventions used in presenting information
9. Information is not protected with adequate security.
10. Information is easily obtainable
11. Information output can be customized as required
12. The types of information presented are suitable for your needs.
13. There is duplicate or inconsistent information.
14. Information can only be modified by people who should be able to modify it.
15. Information is quick to retrieve.
16. The display of retrieved information is suitable for your needs.

17. Information is protected against unauthorized access.
18. Information layout can easily be modified as required.
19. The level of detail presented can be easily changed as needed
20. Information is not presented at the appropriate level of detail or precision.
21. It is easy to modify the types of fields displayed.
22. Information is easily accessible.
23. The display of information can easily be changed as needed.
24. It is not easy to change the measurement units used to display information.
25. Information can easily be collated.
26. Information is easy to find and retrieve.
27. Some types of information that you need are missing. (Ex. tutorial duration not given)
28. The level of detail or precision for information can be modified to suit your needs.
29. Information is presented in an appropriate format.
30. The precision of numeric fields cannot be customized.
31. The types of information presented can be easily changed as needed.
32. Information is quickly accessible when needed.
33. Information is easy to comprehend.
34. It is easy to interpret what this information means.

ABOUT THE AUTHORS

Rosanne J. Price is an Australian Postdoctoral Fellow in the Department of Information Systems in University of Melbourne, and an Honorary Research Associate in the Faculty of Information Technology in Monash University. Her research interests and publications are in the area of data quality, databases, and information systems, and object-oriented and conceptual modeling. She received her Ph.D from Monash University. She has had over 14 years of academic and professional experience, including a senior research fellow position at Monash University and research/lecturing positions at RMIT, and European divisions of Boston University and the University of Maryland.

Dina Neiger (Dina.neiger@infotech.monash.edu.au) has over 15 years of experience with academic and professional distinction in statistics, decision support and business process engineering. She holds a PhD in Business Systems from Monash University.

Graeme Shanks (GShanks@unimelb.edu.au) is a Professorial Fellow in the Department of Information Systems at the University of Melbourne. He has also held academic positions at Monash University and Chisholm Institute of Technology, Australia. Prior to becoming an academic, Graeme worked for a number of years as programmer, programmer-analyst and project leader in several large organizations. Graeme has a number of research and teaching interests including conceptual modeling, data quality, decision support systems, identity management, and the implementation and impact of enterprise and inter-organisational systems. Graeme has published the results of his research in over 100 refereed journals and conference papers.

Communications of the Association for Information Systems

ISSN: 1529-3181