A Model for Information Quality in the Banking Industry – the Case of the Public Banks in Brazil

(Academic Paper - completed)

Luís Francisco Ramos Lima
Universidade Federal do Rio Grande do Sul, Escola de Administração, Brazil
luisfranciscolima@gmail.com

Antonio Carlos Gastaud Maçada
Universidade Federal do Rio Grande do Sul, Escola de Administração, Brazil
agcmacada@ea.ufrgs.br

Xenophon Koufteros
Texas A&M University, Mays Business School, USA
xkoufteros@mays.tamu.edu

Abstract: Poor Information Quality (IQ) has a significant impact on the general performance of an organization. Investments in IT have had a great impact on the banking industry in recent years, producing a large increase in the volume of information; these investments, however, have not ensured a corresponding return in terms of the quality of information produced. The aim of the present study is to validate a specific IQ model for the Banking Industry, and specifically for the public banks at a federal and state level in Brazil. A previously proposed IQ measurement instrument consisting of 15 IQ dimensions and 65 items was employed. Statistical techniques such as Exploratory Factor Analysis and Reliability Analysis were first applied in order to get a preliminary assessment of the psychometric properties of the constructs. Subsequently, Confirmatory Factorial Analysis using Structural Equations Modeling (first and second order Measurement Models) was used to refine the instrument and examine validity and reliability in a more stringent sense. The final model has four quality factors: *Accessibility*, *Contextuality*, *Believability* and *Comprehensiveness*, arranged in 12 items.

Key words: Information Quality, measurement model, IQ model, public banks.

1. Introduction

While there has been a great deal of investment in Information Technology, the returns have not come up to expectations in terms of the quality of information. In the previous decade, [12] highlighted the importance of the role of IT in aiding organizations to survive and prosper in the market; for this, information is one of the crucial elements and they need to have supporting IT. IT *per se* ensures neither the quality of the information applied nor its good use, as "even those firms that are famous for the application of specific Information Systems (IS) frequently have poor internal informational environments"[13]. This issue remains current, "organizational Information Quality reveals more than IT processes – it reveals how the organizational processes are organized and implemented"[17].

Although it is a relatively new area of study, there has been a gradual advance in the field of Information Quality (IQ). Since the 1990s, debate on this topic has largely come to focus on two lines of study: one that focuses on the administrative/strategic aspects of IS, as in the work of [22] and [13], and another that highlights technological/operational aspects, related to data quality, led by Richard Wang, at MIT – Massachusetts Institute of Technology [35,36], among others. Nevertheless, great efforts have been made to solve the IQ problems in the

academic and business worlds, as there is a critical need for a methodology that can be used to determine to what degree organizations develop information products and services of quality for their users [16, 26].

Investments in IT have had a significant impact on the banking industry [21]; today, technology plays an increasingly important role in the development of new services and more efficient management of the institutional structures of banks [9]. These investments in IT, however, do not ensure the good quality of the information: "banking corporations simply forget to check the quality of their data, directing all their attention to the identification, extraction and information load, the result of which has been dramatic" [8]; and companies have come to question the quality of the information produced by the IT in which they are investing or intending to invest [6].

The banking industry was chosen as the object of this study because of its heavy investment in, and extensive use of IT, as well as the fact that information represents a key element influencing the performance and success of organizations in the sector. Within the context of the industry, it was decided to analyze the public banks due to their relevance: though they represent a small number of institutions within the total number of banks in the country (14 out of 161 banking institutions, approximately 9% in total), the public banks have the largest amount of assets (The *Banco do Brasil* – BB – and *Caixa Econômica Federal* – CEF – retain approximately 30% of the total volume of assets), while also being among the institutions with greatest penetration (number and coverage of branches) in the country [10].

It is noted that the study of Information Quality, while important, is, as stated above, a little explored subject, for which reason the following research question was put: Which dimensions represent Information Quality in the Banking Industry, particularly the public banks? This article seeks to answer this question, while developing and validating a specific IQ model for the Banking Industry, in the federal and state public banks.

2. Information Quality

In their discussion on IQ, [33] state that "bad Information Quality can lead to chaos" in an organization. The same authors, on another occasion, suggest that "those that consume information should avoid problems of quality, whatever their dimension, in their search for high quality information" [32]. An example of this can be seen in the manufacturing industries, where, despite more than a decade of research and practice, it is rare to find a technique for measuring, analyzing and enhancing the quality of the informational products in their business processes [34].

Problems of Information Quality encompass much more than incorrect values. They can also include problems and errors relating to production, technical problems relating to the storage of, and access to data, as well as those caused by changes in the informational needs of the consumers [32]. Also points out some impacts resulting from problems of information quality, such as client dissatisfaction, increased operational costs, less effective decision making and a reduced capacity to produce and execute organizational strategy [31].

A number of studies have analyzed Information Quality as an adjacent topic; as the information utility. Other studies [3, 39, 4] present information quality as a means of assessing the quality of Information Systems. As yet, no studies have been found that look at IQ from sectorial or organizational point of view.

2.1 Definitions of Information Quality and the State of the Art

In little over a decade of studies on IQ, a number of definitions have been developed. Table 1 shows the main definitions found. With a close reading of the literature concerned with definitions, it is possible to note a shift from a technological and operational concept (the simple use of data, according to [37] towards a broader definition, in which information, processes,

business effectiveness, consumer and information products and services are used according to [7], so suggesting a migration from mere concern with measuring IQ towards a strategic point of view.

WANG; KON; MADNICK ,1993	Easily usable and understandable data that reflect real conditions
McGEE; PRUSAK, 1994, p. 166	Detailed care with integrity, accuracy, freshness, interpretability and general value of information, as judged by the client.
STRONG; LEE; WANG, 1997[a]	Data that is suited to the use of consumers of information.
O'BRIEN, 1999	Characteristics of information products, whose quality or attributes aid in making them valuable.
ENGLISH, 2002, p. 208	Elimination of the waste of dirty information and reworking, unnecessary processes and increment to the effectiveness of the business by increasing consumer satisfaction among those that consume information products and services.

Table 1 – Definitions of Information Quality

In the field of Information Quality, however, because it is a recent area of study, there are still problems with the definition and organization of its body of theoretical work. For this reason, [30], contributed a proposed theoretical model based on the use of Conceptual Maps, following the guidelines of [23], in an attempt to consolidate the studies available in the main IQ events and IS journals. It is posited that the field of "Quality of Information" is divided into three large conceptual groups, or denominated visions: (1) Operational Vision: is concerned with the impact of IQ on the organization, its management, influence on structure and production, strategic approach and tactics; (2) Behavioral Vision: analyses the influence of the human aspect on IQ, in the performance of everyday tasks, in the insertion of IQ in professional skills and capacities and above all on the view of the organization's internal and external clients; (3) Processual View: looks at the technical and methodological aspects of IQ, such as measurement units, development of applications, data control tools and information systems; denoting an operational approach.

However some authors conceive of the production and distribution of information as the 'manufacture of information' [33]. There are two forms of products in firms: products in a physical form and in an informational form, denominated *information products*, which are generated from the organizational processes [34]. Within this vision of informational product, in the same way as quality is a dimension of a manufactured product (such as color, size, height, caliber, etc), information also has its dimensions [36].

2.2 Research on the Dimensions of IQ

Researchers have continuously made efforts to define attributes or dimensions of IQ. Table 2 shows the 15 dimensions of IQ as proposed by Pipino, Lee, and Wang which we adopt for the purposes of our study. The list is quite comprehensive and covers the content domain of IQ. To support the validity of the 15 dimensions, we summarize other literature and the proposed dimensions on Table 3 [22, 32, 31, 24, 25, 28, 38, 18]. These are arranged to facilitate the understanding of their relation to the dimensions adopted in the present study.

Dimension	**Description**
Accessibility	How available data the are, or how fast and easily they can be recovered
Amount of Data	How suitable the volume of data is in relation to the task
Believability	How believable the data are thought to be
Completeness	The degree to which there is no lack of data and the sufficiency of their depth and width for the task

Concise Representation	How compact the representation of the data is
Consistent Representation	How frequently the data are presented in the same format
Ease of Manipulation	How easy the data are to manipulate and use in different tasks
Free-of-Error	How correct and reliable the data are
Interpretability	How suitable the language, symbol and unit is and the clarity of its definition
Objectivity	How non-dispersed and impartial the data are
Relevancy	How applicable and collaborative the data are
Reputation	How valuable the data are in accordance with their source or content
Security	How sufficiently restricted the data are in order to maintain security
Timeliness	How current the data are for the performance of the task
Understandability	How easily understood the data are

Table 2 – Dimensions used for Research

Source: Pipino, Lee and Wang, 2002

PIPINO; LEE; WANG (2002)	*McGEE; PRUSAK (1994)*	*STRONG; LEE; WANG (1997a, p. 104)**	*REDMAN (1998)*	*O'BRIEN, (1999, p. 49)*	*OZ (2000, p. 10)*	*WIXOM; TODD (2005, p. 88)*	*KIM; KISHORE; SANDERS (2005)*
Accessibility		X	X (Privacy)				X (2)
Amount of Data		X					X (1)
Believability		X					X (3)
Completeness		X	X	X	X	X	X
Concise Representation		X		X (Detailing)			X (1)
Consistent Representation		X	X	X		X	X
Ease of Manipulation							
Free-of-Error	X (Precision)	X (Accuracy)	X (Accuracy)	X (Accuracy)	X (Accuracy)		
Interpretability	X	X	X				X (1)
Objectivity		X		X (Scope)			X (3)
Relevancy		X	X	X	X		X
Reputation		X					X (3)
Security		X (of access)	X				X (2)
Timeliness	X	X (Temporality)	X	X	X (Temporality)	X (Temporality)	X
Understandability or Comprehensiveness		X		X		X	X (1)
Other dimensions	General value	Value	Granularity Possessing Detail level	Frequency Performance Order Representation Media	Cost		Mantaining Delivery

(1) - Grouped as Packaging
(2) - Grouped as Accessibility
(3) - Grouped as Accuracy

Table 3 – Literature Review of Dimensions

The approach put forward by [28] was adopted in the present study for the following reasons: (1) it is comprehensive as it satisfactorily encompasses the dimensions of previous authors, and (2) the associated items measure the dimensions in a perceptive manner, which reflects the needs and experiences of the people involved, and so do not depend on other sources of organizational information that may not be made available. The authors also grouped the dimensions obtained into four categories of information [36, 28]. These categories (Table 4) synthesize the concepts related to each dimension, and are used to define the main profiles to be assessed in IQ:

Category	Concept	Related Dimensions
Intrinsic	The information should possess in its own condition	Credibility, Objectivity, Reputation, Free-of-error
Contextual	The quality should be considered within the context of the task that uses it, in order to aggregate value	Completeness, Amount of Data, Relevance, Timeliness
Representative	The information should be representative, emphasizing the importance of the ISs that use it	Conciseness, Consistency, Understandability, Interpretability
Accessibility	The information should be freely accessible to whoever it may be attributed to, while also emphasizing the importance of the systems used to manage it.	Accessibility, Ease of Use, Security

Table 4 – Information Categories
Source: adapted from [36] and [28]

The instrument selected for use in the present study, developed by [20], is composed of the IQ dimensions (Table 2), based on the work carried out by [28].

3. Model of the Study

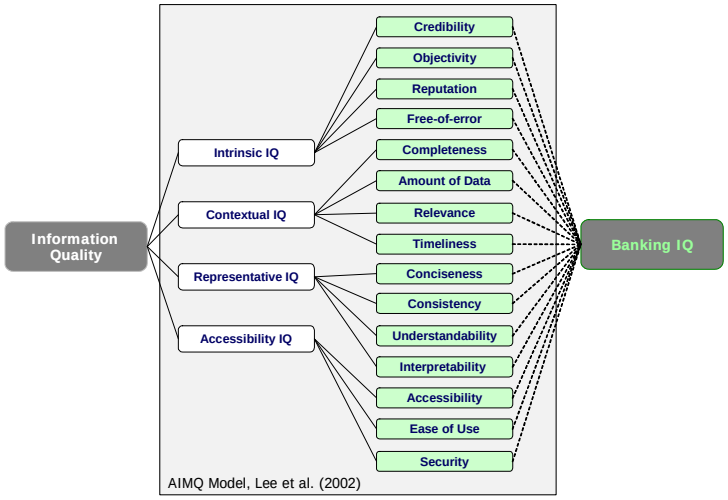

Figure 1 – Model of the Study

The review of the studies carried out on the topic made it possible to restructure the initial research model of this article, proposed in Figure 1. Of particular note in the proposed model (as outlined), is the customization of the construct *Information Quality* for use in the specific industry to be studied.

4. Research Methods

This section deals with the methods and techniques used in carrying out this study. Essentially, we carried out a pilot study to gain an initial understanding of the measurement properties of the scales. This was primarily exploratory work. Subsequently, we collected additional data and applied more confirmatory techniques, firstly to develop and test a first order measurement model and then, to test a second order measurement model.

The items contained in the original instrument were operationalized using a Likert-type scale with 11 points (0 = completely disagree; 10 = completely agree). Those taking part in the study responded to 65 items following the question: "What is the Quality of the Information that you use in your firm?" The 65 items were randomly placed within the instrument to avoid potential issues with sequencing. Fifteen items were reverse scaled but accommodated appropriately during data analysis.

4.1 Unit of Analysis

The target population of the present study is composed of branch-based management level executives who work with intensive information. As they directly manipulate the information necessary for the execution of their tasks, they are considered the most suitable to supply a perceptive diagnosis of the quality of that information. The samples are defined by the availability of the firms. The Banco do Brasil made several groups available from an internal training program; Banrisul indicated the five largest branches in Porto Alegre/RS and the CEF indicated the ten largest in the same geographic location.

The Pilot step was carried out in the Banco do Brasil (BB), as it was considered to be the institution of greatest penetration in the country with more than 14,500 customer attendance points and the highest volume of investments in technology, with an estimated investment in the order of 1.6 billion for the year 2006 [1]. The final large scale study was carried out in two banks: Banrisul (The State bank of Rio Grande do Sul) was contacted as it represents the public bank with the highest penetration in the state of Rio Grande do Sul and the second largest state public bank, and the Caixa Econômica Federal (CEF), as it is the third largest bank in Brazil in terms of total assets.

4.2 Stage 1 – Pilot Study Methods

The validation and refinement of the instrument was performed in accordance with the model provided by [19]. The steps taken in developing the instrument were: analysis of the Instrument (back-translation) – face and content validation; data collection by convenience (bank workers); analysis of reliability (1) - Cronbach's alpha technique; factor analysis (within block); factor analysis across dimensions (Discriminant validity) and Analysis of Reliability (2) full instrument.

4.3 Stage2 – Large Scale Study

A new set of data was collected from two banking institutions Banrisul and the Caixa Econômica Federal. A new analysis of Reliability was also carried out to ensure the quality of the final instrument.

Later, Confirmatory Factor Analysis using Structural Equation Modeling, was performed in order to assess the posited measurement model. The measurement model specifies how the latent variables (factors) can be measured in function of observable variables (items) and allows reliability to be estimated [5]. The second order measurement model makes it possible to verify whether the estimated first order factors are in reality sub-dimensions of a wider construct [14]: in this case, the IQ. Table 6 contains detailed descriptions of each step.

Step	Description
Examination of the revised Instrument	The instrument was produced from the results obtained in the previous Stage and from items added during this step. The Face and Content Validities were also noted, especially in any possible new additional items.
Data Collection	A new sample of employees was obtained from the banks BANRISUL and CAIXA ECONÔMICA FEDERAL.
Analysis of Reliability 1	The relationship of the constructs in the instrument was assessed. The Reliability of the instrument was tested using Cronbach's Alpha.
Confirmatory Factor Analysis	Structural Equation Modeling using AMOS was applied. Also, the Validity of the Construct was checked, that is, the link between the measurements and the IQ construct were checked.
Analysis of Reliability 2	Composite Reliability was calculated to the full instrument, considering only the items and dimensions selected in the previous step.

Table 6 – Details of Large Scale Study

During the stages of the study, the SPSS (Statistical Package for Social Sciences) was used for the exploratory work, while AMOS was used for the Confirmatory Factor Analysis.

5. Results

This section describes the findings from the various stages of the process of refining the model and the instrument. The descriptions presented below refer to Stage 1 and are relative to the article presented by [29], which shows the details necessary for the development of the final research instrument.

5.1 Stage 1 – Pilot Study

Firstly, the instrument from [20] was translated and adapted to the Portuguese language. The face and content validity was checked to assure the integrity of the instrument. One hundred and seventy (170) of these questionnaires were then sent to executives from the Banco do Brasil who were taking part in a training program organized by the bank. The rate of return of the survey was 88.8%. The questionnaires that were considered valid for use represented 75.3% of the total sent and 84.8% of those that were returned. Based on Corrected Item Total Correlation (CITC) criteria the following dimensions were eliminated: Ease of Use, Relevance and Security. Dimensions with a Cronbach's Alpha of less than 0.7 [14] were also excluded – Quantity (0.592) and Error-Free (0.680) – resulting in an instrument with 10 dimensions and 34 items, with reliability of 0.971 for the entire instrument.

Using Exploratory Factorial Analysis (EFA) on the retained items, five factors were identified: Accessibility, Conciseness, Believability, Contextuality and Comprehensiveness. In order to confirm the result, the same analysis was repeated, this time specifying the desired factors (which, in this case, were 10 factors so as to correspond to the dimensions retained after CITC and reliability analysis), and only the first five obtained were significant for the analysis, reinforcing the prior findings of 5 factors. This set of five factors resulted in an explained variance of 75.4%, of the total, a value superior to the lower limit of 60% for Social Sciences proposed by [14]. Seven more items were eliminated at this step due to low loadings.

In order to validate the final instrument with the five factors and 27 items, a new reliability analysis was carried out. The Cronbach's Alpha score for the final instrument is 0.964; the reliability coefficients for each factor (new dimension) and number of items are: Accessibility (0.874; 4); Conciseness (0.843; 3); Believability (0.903; 5); Contextuality (0.925; 7); Comprehensiveness (0.936; 8).

5.2 Stage 2 – Large Scale study

Once the instrument had been revised, a meeting was held with two representatives of the bank from which the pilot study had come, during which the findings were presented. An assessment of the application of the pilot questionnaire was made: the response percentages were given, along with the findings – new factors obtained and their respective items.

In this new stage 200 questionnaires were sent – in this new application executives from Banrisul (five branches) and Caixa Econômica Federal (10 branches) took part. The reponse rate – out of the total delivered questionnaires - was 90.8% for the Caixa and 91.4% for Banrisul, an overall total of 91%.

Using a pre-defined model, Confirmatory Factorial Analysis (CFA) was performed in order to identify latent relationships between the observed variables. The diagram presented in Figure 2 shows five correlated latent variables (factors represented byξ) corresponding to the 27 items included in the final instrument, represented by **x**. The measurement errors for each item are represented by δ. The factorial loads of each item-factor pair are represented by λ, and the correlations obtained for the factor-factor pair are represented by ϕ.

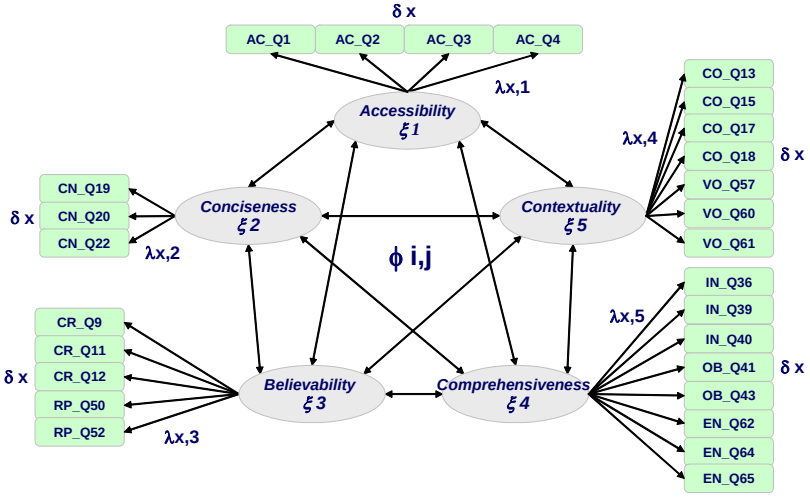

Figure 2 – Diagram of the Paths of the Initial Measurement Model

The model shown here was included in the AMOS statistical software. There was a sufficient number of respondents as recommended by [14]. The responses were five times greater than the quantity of items. One hundred and fifty-nine (159) valid occurrences were included.

The measurement model was also assessed by examining factor to item loadings composite reliability measures and Average Variance Extracted (AVE). In addition we examined Discriminant Validity analysis. In this first process, of the 27 items that composed the final instrument 12 were eliminated (15 items remained). Items with a standardized factorial loading below 0.7 (value recommended for confirmatory studies) were excluded, as they reveal a percentage of explained variance of the item lower than 50% [14].

This refinement, however, proved to be insufficient. The fit indices of the measurement model, though approximate, were not at the levels recommended [14] for a model with 15 items, indicating that further refinement of the model was necessary. Conciseness was dropped as it did not meet acceptable criteria. Table 7 presents the model fit values of the Final Model.

Adjusted model values and recomended values		
	Recomended Values	*Definitive Values*
Chi-Square / Degrees of Freedom (χ^2/df)	≤ 3,00	2,32
GFI - Goodness of Fit Index	≥ 0,90	0,90
AGFI - Adjusted Goodness of Fit Index	≥ 0,80	0,83
NFI - Normed Fit Index	≥ 0,90	0,91
NNFI - Non-Normed Fit Index (Tucker-Lewis)	≥ 0,90	0,93
CFI - Comparative Fit Index	≥ 0,90	0,95
RMSEA - Root Mean Square Error of Approximation	≤ 0,10	0,09

Table 7 – Model Fit Evaluation
Source: Adapted from [14, 21]

In the final model, the values for Composite Reliability (CR) are above the desired level of 0.7 [11, 14]. The Average Variance Extracted (AVE) values also exceeded the recommended 0.50 [14], as shown in Table 8. In all the factors there was a significant improvement in the AVE values: the factor *Comprehensiveness* was the most significant, increasing from 51% to 69%.

	Before purification (27 items)		After purification (12 items)	
	CC	AVE	CC	AVE
Accessibility	0,75	0,57	0,81	0,68
Conciseness	0,71	0,59		
Believability	0,84	0,62	0,81	0,68
Contextuality	0,81	0,53	0,81	0,68
Comprehensiveness	0,81	0,51	0,82	0,69

Table 8 – Composite Reliability and Average Variance Explained

The Discriminant Validity analysis was conducted by comparing the variance shared between the different pairs of constructs (squared correlation) with the variance extracted (AVE) in each one of them [11], as shown in Table 9. The correlation between the factors varied from 0.53 (Accessibility) to 0.69 (Comprehensiveness). The Discriminant Validity is ensured as there was no shared variance between two factors greater than the variances explained by each of these factors.

	Accessibility	Believability	Contextuality	Comprehensiveness
Accessibility	**0,68**	0,53	0,58	0,61
Believability	0,28	**0,68**	0,63	0,69
Contextuality	0,33	0,39	**0,68**	0,68
Comprehensiveness	0,37	0,33	0,47	**0,69**

Chart 9 – Discriminant Validity Test and Correlation Matrix

Note: The values above the main diagonal are the correlations (All $p < 0.001$) and the values below are the squared correlations (i.e. shared variance between the factors). The main diagonal indicates the explained variance for each factor (in bold).

Table 10 presents the final indicators of the Confirmatory Factorial Analysis. The statistical *t* values, associated to each one of the factor loadings, indicate the significance of the items. In general, if the *t* values are greater than 2 or 2.576, they are considered significant at the level 0.05 or 0.01 [19]. The R^2 values are all above 0.5. The lowest *t* value found in the investigation is 9.99 (CO_Q17), thus being highly significant. Hence, all the items are

significantly related with their factors, confirming the positive relationship between them. The total variance explained by the instrument is 68.21%.

	Item	Non-standardized Factor loading	Standardized Factor loading	St Error	*t* value	R^2	Explained Variance
Accessibility	AC_Q2	*1,00*	0,76			0,57	68%
	AC_Q3	1,25	0,93	0,11	11,31	0,86	
	AC_Q4	1,26	0,81	0,12	10,41	0,66	
Believability	CR_Q11	*1,00*	0,88			0,78	68%
	CR_Q12	0,91	0,80	0,08	11,74	0,64	
	RP_Q52	0,81	0,79	0,07	11,44	0,62	
Contextuality	CO_Q13	*1,00*	0,84			0,70	68%
	CO_Q15	1,16	0,90	0,09	12,75	0,81	
	CO_Q17	0,86	0,72	0,09	9,99	0,52	
Comprehensiveness	EN_Q62	*1,00*	0,82			0,67	69%
	EN_Q64	0,97	0,85	0,08	11,81	0,72	
	EN_Q65	0,92	0,83	0,08	11,43	0,68	

Table 10 – Final Indicators of the Confirmatory Factorial Analysis

Note: the values in italics are the loads inferred by the AMOS

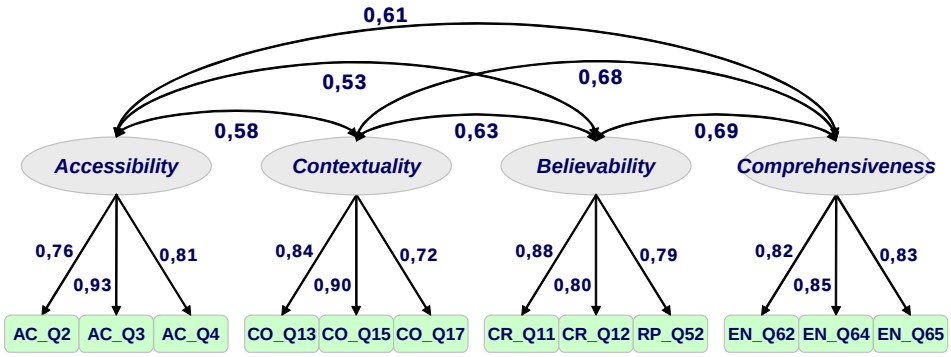

Figure 3 – Diagram of the paths of the Final Measurement Model

Note: For the sake of clarity, the measurement errors inherent to the items have been were excluded from the diagram

The items resulting from the CFA, their respective factor loadings and the correlations found between the factors are presented in Figure 3, within the context of Paths of the Final Measurement Model.

In order to confirm the Final Model, the goodness of fit index was re-checked with all the survey respondents (pre-test and final surveys). This resulted in a database of 287. Table 11 presents the values of the Final Model together with the values of the complete database.

Adjusted model values and recomended values			
	Recomended Values	*Definitive Model (sample)*	*Definitive Modell (whole database)*
Chi-Square / Degrees of Freedom (χ^2/df)	≤ 3,00	2,32	2,82
GFI - Goodness of Fit Index	≥ 0,90	0,90	0,93
AGFI - Adjusted Goodness of Fit Index	≥ 0,80	0,83	0,88
NFI - Normed Fit Index	≥ 0,90	0,91	0,94
NNFI - Non-Normed Fit Index (Tucker-Lewis)	≥ 0,90	0,93	0,95
CFI - Comparative Fit Index	≥ 0,90	0,95	0,96
RMSEA - Root Mean Square Error of Approximation	≤ 0,10	0,09	0,08

Table 11 – Adjustment of the Final Model for the Complete Data Base

As can be seen in the above table, the indicators confirm the fitness of the proposed Model. The extended database provided an improvement in these indices, as indicated by [15].

Reliability Analysis was then carried out of the factors and of the instrument following the CFA. The Cronbach's Alpha scores are between 0.854 and 0.869 for the factors; the Alpha for the full instrument was 0.910. Due to the elimination of items, there was a slight reduction in Instrument Alpha, from 0.952 (Final Survey Instrument with 27 items) to 0.91 (Post-AFC Instrument). Nevertheless, the value continued to be high and within the acceptable levels of Reliability. The process of refining the research instrument resulted in a questionnaire consisting of four factors and 12 items: the reliability and the number of items for each variable are: Accessibility (0.862; 3), Believability (0.861; 3); Contextuality (0.854; 3) and Comprehensiveness (0.869; 3).

The CFA ensured that the model relates the four resulting factors. A Second Order Measurement Model made it possible to check whether the previously estimated first order factors are in reality sub-dimensions of a wider and more general construct: in this case, the Quality of the Information. Figure 4 presents a Diagram of the Paths of the second order model. The tested relationship was found to be valid by the factorial loads of the first order constructs. Analysis of this second order measurement model makes it possible to state that the four studied factors can compose a single construct - Information Quality.

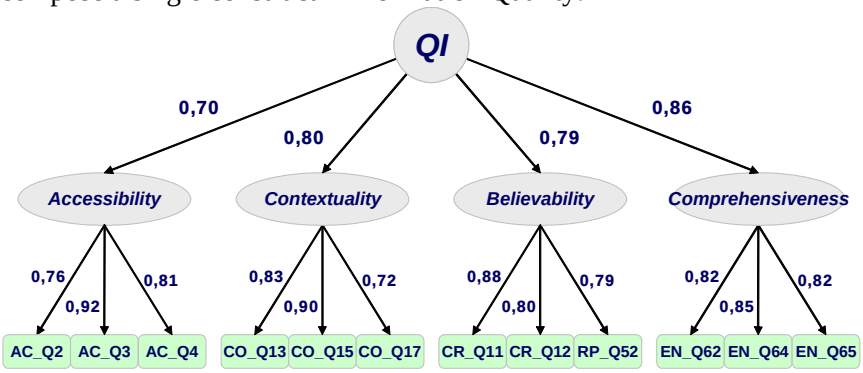

Figure 4 – Diagram of the Paths of the Second Order Measurement Model

Note: For the sake of clarity, the measurement errors inherent to the items have been were excluded from the diagram.

6. Conclusion

The present article has presented the results of research that has led to the development and validation of a model of Information Quality (IQ), composed of dimensions significant to the banking industry, specifically the public federal and state banks in Brazil.

The initial research model was based on a review of the literature and specifically on the original IQ model developed by [20], with 15 dimensions and 65 items. This model was chosen because it represents all the dimensions found in the literature on IQ and IS. The process of refinement and validation of the survey instrument used in the present study was carried out in accordance with the recommendations found in the literature [19].

The resulting model is of use in the process of evaluating IQ in the Brazilian public federal and state banks. It is composed of four dimensions: *Accessibility*, *Believability*, Contextuality and Comprehensiveness, distributed in 12 items.

It can be seen that dimensions that at first would appear important – especially for the sector under study – such as Security, Free of Errors and Ease of Use, came to be eliminated from the model. It was found that such items were not considered important by the respondents, as they are already part of the informational culture of the banks – a great deal of their investments and costs are related to failures in these dimensions (for example, attacks by hackers via the web, rework, client dissatisfaction and investment in training and IS development), so justifying their exclusion. A degree of semantic redundancy was perceived in some of the original dimensions – examples being *Understandability* and *Interpretability*. The elimination of some dimensions with little semantic differentiation during the validation process produced greater clarity in the proposed model. Thus it can be stated that the four dimensions in the model are quantitatively different and qualitatively distinct from one another. This was identified to be the case and proven during the interviews held with executives from the public banks in which the research findings were presented. This procedure of returning the findings to the respondents contributed greatly towards the analysis.

The rigor and care taken in the use of statistical techniques aided in the validation of the First and Second Order Measurement Models, demonstrating how robust the validated IQ model for the public federal and state banks is.

The contributions made by this study are the model and instrument that, if replicated, will aid in identifying the IQ perceived by executives. Their findings can be of use in the elaboration of IQ strategies related to the information products that constantly circulate within the banks and are shared with the clients.

References

[1] BANCO DO BRASIL SA. Relatório Anual 2004. Disponível em http://www.bb.com.br/portal/ri/dce/dwn/relan2004.pdf. Acessado em 15/07/2005.

[2] BOUDREAU, Marie-Claude; GEFEN, David; STRAUB, Detmar W. Validation in Information Systems research: a state-of-the-art assessment. MIS Quarterly. March 2001, v. 25, n. 1, p. 1-16.

[3] BRODBECK, Angela Freitag. Avaliação da Qualidade da Informação nos Sistemas de Informação e de Apoio à Decisão: um Estudo Introdutório. In: Anais do XIX Congresso Nacional da ANPAD. João Pessoa: 1995.

[4] CECCAGNO, Charlene. Qualidade das Informações Geradas Pelos Sistemas ERP. Trabalho de Conclusão do Curso de Especialização em Gestão Empresarial, EA/UFRGS. Porto Alegre, 2001.

[5] CHAU, P. Y. K. Reexamining a Model for Evaluating Information Center Success Using a Structural Equation Modeling Approach. Decision Sciences. 1997, v. 28, n. 2, p. 309-334.

[6] COHAN, Peter S. CFO's to Tech: "I'll Spend for the Right Technology". Financial Executive. April 2005, p. 30-34.

[7] ENGLISH, Larry P. Process management and Information Quality: how improving Information Product Processes improves Information (product) Quality. In: Proceedings of the 7th International Conference on Information Quality ICIQ'2002. 2002, p. 206-209.

[8] FÉ, Ana Lucia Moura. BI de Novo na Agenda do CIO. InfoCorporate. Janeiro 2006, n. 28, p. 27-39.
[9] FEBRABAN. Atendimento e Serviços.Disponível em http://www.febraban.org.br/ Arquivo/Servicos/Dadosdosetor/atendimento_2005.asp, Acessado em 14/07/2005. (a)
[10] FEBRABAN. O Setor Bancário em 2005. Disponível em http://www.febraban.org.br/Arquivo/Servicos/Dadosdosetor/2006/item01.asp. Acessado em 20/12/2006.
[11] FORNELL, C., LARCKER, D.F. Evaluating Structural Equations Models with Unobservable Variables and Measurement Error. Journal of Marketing. 1981, v. 18, n. 1, p. 39-50.
[12] FREITAS, Henrique. et al. Informação e Decisão: sistemas de apoio e seu impacto. Porto Alegre: Ortiz, 1997
[13] DAVENPORT, Thomas H. Ecologia da Informação. São Paulo: Futura, 1998.
[14] HAIR, J.F. et al. Análise Multivariada de Dados. Porto Alegre: Bookman, 2005
[15] HOYLE, R.H. Structural Equation Modeling: Concepts, Issues and Applications. Sage, 1995.
[16] KAHN, B.K.; STRONG, D.M.; WANG, R.Y. Information Quality benchmarks: Product and Service performance. Communications of the ACM. Apr 2002, v. 45, n. 4, p. 184-192
[17] KATZ-HAAS, Raïssa; LEE, Yang. Understanding Interdependencies between Information and Organizational Processes. In: Information Quality. New York: M.E.Sharpe, 2005, p. 167-178.
[18] KIM, Yong Jin; KISHORE, Rajiv; SANDERS, G. Lawrence. From DQ to EQ: Understanding Data Quality in the Context of E-Business Systems. Communications of the ACM. October 2005, v. 48, n. 10, p. 75-81.
[19] KOUFTEROS, Xenophon A. Testing a model of pull production: a paradigm for manufacturing research using structural equation modeling. Journal of Operations Management. 1999, v. 17, p. 467-488
[20] LEE, Yang W. et al. AIMQ: a Methodology for Information Quality Assessment. Information & Management. Amsterdam: Elsevier, December 2002, v. 40, n. 2, p. 133-146
[21] MAÇADA, A.C.G. Impacto dos investimentos em Tecnologia da Informação nas variáveis estratégicas e na eficiência dos bancos brasileiros. Tese de Doutorado apresentada no Programa de Pós-graduação em Administração PPGA. Escola de Administração, Universidade Federal do Rio Grande do Sul, 2001
[22] McGEE, James; PRUSAK, Laurence. Gerenciamento Estratégico da Informação. São Paulo: Campus, 1994.
[23] NOVAK, J.D. A Theory of Education. New York: Cornell University Press, 1977.
[24] O'BRIEN, James A. Management Information Systems. McGraw-Hill, 1999.
[25] OZ, Effy. Management Information Systems. Cambridge: Course Technology Thomson Learning, 2000.
[26] PIERCE, Elizabeth M. Introduction. In: Information Quality. New York: M.E.Sharpe, 2005, p. 3-20.
[27] PINSONNEAULT, A.; KRAEMER, K.L. Survey Research in Management Information Systems. Journal of Management Information Systems. Autumn 1993.
[28] PIPINO, Leo L.; LEE, Yang W.; WANG, Richard Y. Data Quality Assessment. Communications of the ACM. April 2002, v. 45, n. 4
[29] RAMOS-LIMA, L.F.; MAÇADA, A.C.G.; BRODBECK, A.F. Métricas para Avaliar a Qualidade da Informação. In: Anais do Congresso Anual de Tecnologia da Informação – CATI 2006. FGV-EAESP: São Paulo, 2006.
[30] RAMOS-LIMA, L.F.; MAÇADA, A.C.G.; VARGAS, L.M. Research into Information Quality: a Study of the State-of-the-Art in IQ and its Consolidation. In: Proceedings of the 11th International Conference on Information Quality ICIQ'2006. 2006, p. 146-158.
[31] REDMAN, Thomas C. The impact of poor data quality on a typical enterprise. Communications of the ACM. February 1998, v. 41, n. 2.
[32] STRONG, Diane M.; LEE, Yang W.; WANG, Richard Y. Data Quality in Context. Communications of the ACM. May 1997, v. 40, n. 5, p. 103-110. (a)
[33] STRONG, Diane M.; LEE, Yang W.; WANG, Richard Y. 10 Potholes in the Road to Information Quality. Computer. Los Alamitos: IEEE Computer Society Press, August 1997, v. 30, n. 8, p. 38-46. (b)
[34] SU, Ying; JIN, Zhanming. A methodology for Information Quality Assessment in the designing and manufacturing processes of mechanical products. In: Proceedings of the 9th International Conference on Information Quality ICIQ'2004. 2004, p. 447-465.
[35] WANG, Richard Y. et al. Manage your information as a product. MIT Sloan Management Review. Summer 1998, v. 39, n. 4, p. 95-105.

[36] WANG, Richard. A product perspective on Total Data Quality Management. Communications of the ACM. February 1998, v. 41, n. 2, p. 58-65.
[37] WANG, Richard Y. KON, Henry B. MADNICK, Stuart E. Data Quality requirements analysis and modeling. In: 9th International Conference on Data Engineering. April, 1993.
[38] WIXOM, Barbara H.; TODD, Peter A. A Theoretical Integration of User Satisfaction and Technology Acceptance. Information Systems Research. March 2005, v. 16, n. 1, p. 85-102.
[39] WIXOM, Barbara H.; WATSON, Hugh J. An Empirical Investigation of the Factors Affecting Data Warehousing Success. MIS Quarterly. Mar 2001, v. 25, n. 1, p. 17-41.