

An Evaluation Instrument (Q-U) for Measuring the usability of Business Intelligence Application

Ayad Hameed Mousa*¹, Heba Adnan Raheem², Nibras Talib Mohammed³

^{1,2,3}University of Kerbala, Karbala, Iraq

*Corresponding Author E-mail: ayad.h@uokerbala.edu.my

Abstract

The main purpose of Business Intelligence (BI) applications is to focus on supporting organizations' strategic, operational and tactical decisions by providing comprehensive, accurate and vivid data to the decision makers. Usability testing is an important part of software development. It focuses on how well users can understand and utilize a product/software in fulfilling their intended goals. In the software development, most software tends to be complex. Hence, the way to ensure that such software will satisfy users will be done by measuring the usability of such software. Due to using many data integration techniques in developing BI application, thus, the different software have different instruments to measure the usability of BI software. In this paper, the new instrument (Q-U) for measuring BI software was developed. A systematic approach was adopted which suggested by many researchers in developing the intended Q-U instrument. The literature has been reviewed in a systematic manner to elicit Q-U instrument attributes. Multi-methods were used in validating Q-U instrument such as face validity, a pilot study for testing the goodness of Q-U consistency, and factor analysis. Multi-methods were used in validating Q-U instrument such as face validity, a pilot study for testing the goodness of Q-U consistency, and factor analysis as well as Bartlett's test for measuring the reliability of Q_U. The finding obtained indicates the workability of Q_U in measuring the usability OF BI applications. Ultimately, Q_U was used in measuring BI application in two different sectors.

Keywords: BI Application, Usability, Software Usability, BI Usability, Pilot Study.

1. Introduction

Business Intelligence (BI) is the mechanism to provide insights for most of the operations and performance of organizations, in addition to identifying strategic business opportunities. BI from a technical point is a set of techniques, tools and methodologies that work together to transform the information and data belonging to the organizations into meaningful and actionable information and making this information available to decision makers in an organization [1-10]. Usability testing is an extremely important element in software development. It focuses on how well users can understand and utilize a product in achieving their intended goals [11, 12]. In the era of software development, most software functions tend to be complex and to ensure that such software will satisfy users, it is extremely important for this software is to have the high degree of usability [13, 14]. In the same aspect, the optimal use of BI applications depends on various factors including the usability of the product [15, 16].

2. Software Testing in BI Environment

Measuring the usability of software is a significant element in software development. It helps to know how users can comprehend and utilize a software to fulfill the objectives which designed for [11,

12, 17]. According to [14], usability, to some extent, is the question of "whether the system is good enough to satisfy all the needs and requirements of the users and other potential stakeholders, such as the users, clients and managers". On the other hand, usability can define as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency. In the same aspects, [18] states that the process of usability evaluation includes three main goals; i) to measure the extent and accessibility of the system's functionality, ii) to measure users' experience of the interaction, and iii) to identify any particular issues within the system.

3. Usability in Bi Applications

Usability can be considered as one of the factors in terms of determining of the best of use and ultimate benefit obtain from BI application. The main aim of BI is to help and support the massive warehouses and flow business data in, out, and around the organizations by identifying, processing the information into meaningful information and valuable managerial knowledge and intelligence [19, 20]. Besides, to gain from the actual users about the prototype usability and according to [21], a usability testing should be conducted.

4. Q-U Instrument Development

To develop the evaluation instrument, a systematic approach was adapted as suggested by authors[22-24]. the rigor process of evaluation instrument development is illustrated in Figure 1.

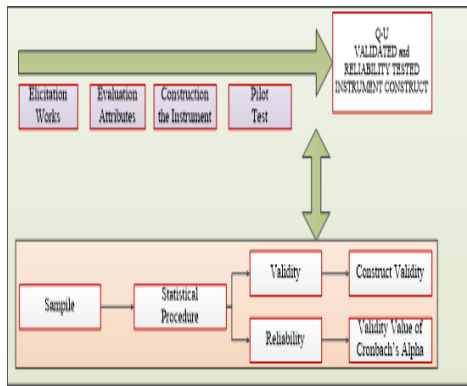


Fig. 1: Summary of Instrument Design

As illustrated in Figure 1, the design of the instrument began with the elicit works to determine the dimensions and items of the evaluation instrument. Then the drafted instrument was piloted for a test of validity and reliability. In this pilot study, 150 respondents were obtained. The respondent numbers who have participated in the pilot study is enough to achieve a reliable outcome in the statistical test as described by[24]. in the next paragraphs, the brief explanation for each stage in instrument design approach was conducted.

4.1. Elicitation Work

six attributes of usability are considered, which are visibility, flexibility, learn-ability, application behavior, error control, and near real-time decision making. These attributes have been proposed to form the q-u instrument attributes, and to be used later in measuring the usability of bi applications. In the same aspect, the proposed attributes were elicited from previous studies related to usability evaluation of bi applications. A total of 20 evaluation works were reviewed in a systematic manner to elicit Q-U instrument attributes[11, 14, 25-32].

4.2. Instrument Validation via Face Validity

According to [29], the first draft of the developed instrument should be validated through Content Validity and Inter-Item Consistency Analysis; the author states that face validity is the fundamental measuring approach for Content Validity. Consequently, this paper engaged four experts in various expertise which are: BI Developers, Software Engineering, Information Technology, and Multimedia, through e-mails as well as face to face consultation to review the items in terms of Content Validity. This is in line with the suggestion of [33]; where three to five experts were employed for their content validation. From the feedback of the experts, it was found that some of the items were not good enough to use and did not fit well with the intended constructs. This led to some modifications to the first draft. In the context of this study, the instrument has 5 scales. Hence, a 5-point Likert scale ranging from strongly disagree (denoted by 1), to strongly agree (denoted by 5) is used in the study.[34] reported that using response options beyond 5-point do significantly alter the scale reliability.

4.3. Pilot Study: Testing Goodness of Measures of Q-U Consistency

In order to ascertain that the instrument indeed measures the required variables or constructs, a pilot study has been conducted to measure their consistency[24]. In the context of this study, 150 respondents were involved among the postgraduate students who have either previously been a BI developer, lecturer in computer science, IT, or software engineering. The respondent numbers who have participated in the pilot study are enough to achieve a reliable outcome in the statistical test as described by [35]. According to [36]; the sample size for construct validity test should at least have 100 responses to get a reliable significant outcome.

4.4. Factor Analysis (Validity)

The objective of determining factor analysis was to verify the degree of significance of each item and which are most suitable for each dimension[24]. Therefore, the test was run and guided for accepting each item based on utilizing Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity, Measure of Sampling Adequacy (MSA), and Factor loading. In preparing the data for factor loading analysis, KMO test was conducted and the results are tabulated in Table 1. It can be noticed that all the values for KMO test satisfy the condition of KMO test ≥ 0.50 .

Table 1: KMO Test and Significant Values

	Attributes	KMO	Significant value of Bartlett's Test of Sphericity
1	Dim1	0.644	0.000
2	Dim2	0.619	0.000
3	Dim 3	0.621	0.000
4	Dim 4	0.661	0.000
5	Dim 5	0.623	0.000
6	Dim 6	0.653	0.000

Where (dim1...dim6) represent the instrument dimensions.

In addition, the Bartlett's test of sphericity gave the significant value of 0.000 for all constructs, which shows the second condition also met and satisfied (significant value of $p \leq 0.05$). Hence, this evidences that the data are ready for factor loading analysis test. Result of Consistency Analysis (Reliability Test).

Reliability of a measure is an indication of consistency. In the pilot study, the measure of consistency is examined through the interim consistency reliability test. The value of Cronbach's coefficient alpha was computed and should indicate the value of alpha to be accepted as reliable [35]. Thus, the reliability of a measure signifies the level at which the measure is without bias and therefore offers dependable measurement across different items of the instrument [36]. This study ran the Cronbach's alpha test and set ($\alpha > 0.6$) to be significant. Table 2 shows the results of reliability test of measurement items, they were found consistent and significant, hence, this measurement items can be used for data collection in the main study.

Table 2: Reliability Test Outcomes

Attributes	Cronbach's Alpha	Items before factor loading	Items after factor loading
Dim 1	0.709	11	9
Dim 2	0.719	6	5
Dim 3	0.725	7	4
Dim 4	0.771	8	6
Dim 5	0.723	7	6
Dim 6	0.745	7	6

As seen in Table 2, all the items in Q-U are found valid and can be used to represent respective dimensions. As stated earlier, factor loadings ≥ 0.50 are considered practically significant and well-defined structure [36]. Thus, all the items that show loading values less than 0.50 are not included in the usability test for BI applications. Since the items proposed in questionnaire were elicited from various previous works, hence it is important to seek confirmation (through factor analysis) to see if these items underlie that proposed dimensions in questionnaire.

As mentioned earlier, the instrument is partially used as a measure of outcome; hence, scale sensitivity becomes an important concern [37]. The overall (Q-U) instrument after refinement is ready to use in testing the usability of any BI application. When concerns with scale reliability, [34] reported that using response options beyond 5-point do not significantly alter the scale reliability. However, difficulties might arise in generating categorical names as the scales expanded [34, 37]. In line with the above situations, the overall instrument consists of six dimensions and 36 items spread over those dimensions. As clearly visualized in Figure 2.

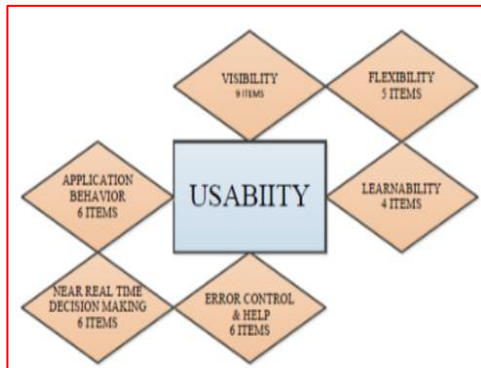


Fig.2: Q-U Instrument Final Draft

5. Use Q-U in Testing Bi Prototype

The usability testing for the two BI applications spread over two organizations was conducted. The first organization under education sector, while the second organization under the business sector. The measurements were made through an instrument named Q-U, which comprises of six main dimensions, visibility, flexibility, learnability, application behavior, error control and help, and near real-time decision making as a Q-U usability attributes. During the usability test, 30 respondents with computer science, software engineering, and IT background participated. Descriptive statistics was used. To describe the basic features of the data in this study, to provide summaries about the sample and the measures, together with graphical analysis, and to form the basis of virtually every quantitative analysis of data.

6. Overall Usability Finding

To obtain the overall finding, the average of results for each usability attribute was calculated. The overall usability for measuring the usability of BI application in organization 1 is made of six dimensions as depicted in Table 3 and Figure 3. Calculating the average agreement value for each dimension, the figure shows that for each dimension, the majority of respondents agree that the BI provide data visibility, is flexible, easy to learn, behave as expected, provide necessary error control and help, and data to allow near real time decision making. This leads that the Q-U instrument is workable in practice.

Table 3: Overall Usability Finding (Case 1)

	Usability Dimension	Strongly Agree & Agree
1	Dim 1	92.98 %
2	Dim 2	85.83 %
3	Dim 3	91.96 %
4	Dim 4	95.55 %
5	Dim 5	88.33 %
6	Dim 6	93.88 %

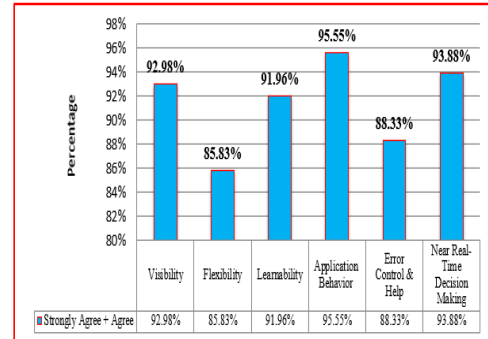


Fig.3: Overall Usability Finding (Case 1)

While the overall usability for testing BI application in these second organization illustrate in Figure 4 and Table 4. The overall usability for measuring the usability of BI application in organization 1 is made of six dimensions as depicted in Table 3 and Figure 3. Calculating the average agreement value for each dimension, the figure shows that for each dimension, the majority of respondents agree that the BI provide data visibility, is flexible, easy to learn, behave as expected, provide necessary error control and help, and data to allow near real time decision making.

Table 4: Overall Usability Finding (Case 2)

	Usability Dimension	Strongly Agree & Agree
1	Dim 1	98.251 %
2	Dim 2	94.66 %
3	Dim 3	93.49 %
4	Dim 4	95.55 %
5	Dim 5	90.00 %
6	Dim 6	95.55 %

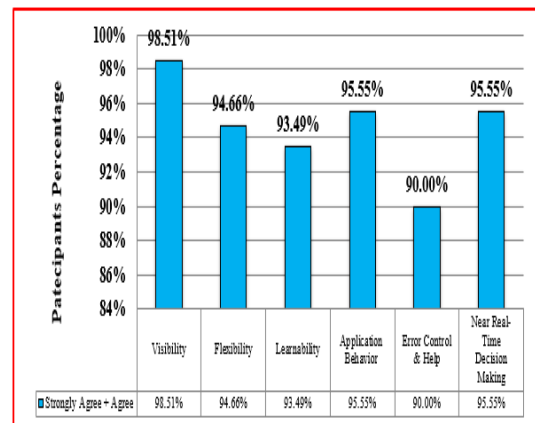


Fig. 5: Overall Usability Finding (Case 2)

The findings also showed that the majority of the participants (more than 90%) in both of education and business sectors, have affirmed the workability of the Q-U instrument in measuring the usability of the BI prototype in particular able to test near real-time decision-making which provided by such BI prototype.

7. Conclusion

Two overall contributions can be obtained from this paper are: (1) the proposed instrument can adapt and use from BI development researchers ;(2) this paper can use as a guideline to other researchers in supporting and guiding them in developed their own instruments.

Acknowledgment

Not forgetting to express my grateful and thanks to the Ministry of Higher Education and Scientific Research (Iraq) and University of Kerbala for supporting and funding my studies. Without their continuous support, this research study would not have materialized.

References

- [1] W. F. Cody, J. T. Kreulen, V. Krishna, and W. S. Spangler, "The integration of business intelligence and knowledge management," *IBM systems journal*, vol. 41, pp. 697-713, 2002.
- [2] M. Castellanos, A. Simitsis, K. Wilkinson, and U. Dayal, "Automating the loading of business process data warehouses," in *Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology*, 2009, pp. 612-623.
- [3] R. Kimball, M. Ross, W. Thorthwaite, B. Becker, and J. Mundy, *The data warehouse lifecycle toolkit*: John Wiley & Sons, 2008.
- [4] R. Kimball and M. Ross, *The data warehouse toolkit: the complete guide to dimensional modeling*: John Wiley & Sons, 2011.
- [5] A. H. Mousa, N. Shiratuddin, and M. S. A. Bakar, "RGMDV: An approach to requirements gathering and the management of data virtualization projects," in *INNOVATION AND ANALYTICS CONFERENCE AND EXHIBITION (IACE 2015): Proceedings of the 2nd Innovation and Analytics Conference & Exhibition*, 2015, p. 030024.
- [6] A. H. Mousa, N. Shiratuddin, and M. S. A. Bakar, "Process Oriented Data Virtualization Design Model for Business Processes Evaluation (PODVDM) Research in Progress," *Jurnal Teknologi*, vol. 72, 2015.
- [7] A. H. Mousa, N. Shiratuddin, and M. S. A. Bakar, "Evaluation Framework for Business Process Evaluation Approaches," *evaluation*, vol. 18, p. 29, 2014.
- [8] A. H. Mousa, N. Shiratuddin, and M. S. A. Bakar, "Virtual Data Mart for Measuring Organizational Achievement Using Data Virtualization Technique (KPIVDM)," *Jurnal Teknologi*, vol. 68, 2014.
- [9] A. H. Mousa and N. Shiratuddin, "Data Warehouse and Data Virtualization Comparative Study," in *Developments of E-Systems Engineering (DeSE), 2015 International Conference on*, 2015, pp. 369-372.
- [10] H. N. Khraibet, A. H. Mousa, A. Bakar, and M. Shahbani, "Intelligent Iraqi Health System (IIHS) using Online Analytical Process (OLAP) model," 2013.
- [11] T. Clemmensen, M. Hertzum, K. Hornbæk, Q. Shi, and P. Yammiyavar, "Cultural cognition in usability evaluation," *Interacting with computers*, vol. 21, pp. 212-220, 2009.
- [12] [12]O. Frandsen-Thorlacius, K. Hornbæk, M. Hertzum, and T. Clemmensen, "Non-universal usability?: a survey of how usability is understood by Chinese and Danish users," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009, pp. 41-50.
- [13] R. Benbunan-Fich, "Using protocol analysis to evaluate the usability of a commercial web site," *Information & management*, vol. 39, pp. 151-163, 2001.
- [14] J. Nielsen, "Guerrilla HCI: Using discount usability engineering to penetrate the intimidation barrier," *Cost-justifying usability*, pp. 245-272, 1994.
- [15] N. Bevan and M. Macleod, "Usability measurement in context," *Behaviour & information technology*, vol. 13, pp. 132-145, 1994.
- [16] M. Z. Elbashir, P. A. Collier, and M. J. Davern, "Measuring the effects of business intelligence systems: The relationship between business process and organizational performance," *International Journal of Accounting Information Systems*, vol. 9, pp. 135-153, 2008.
- [17] A. Lönnqvist and V. Pirttimäki, "The measurement of business intelligence," *Information Systems Management*, vol. 23, p. 32, 2006.
- [18] G. Ellis and A. Dix, "An explorative analysis of user evaluation studies in information visualisation," in *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization*, 2006, pp. 1-7.
- [19] R. Van der Lans, *Data Virtualization for Business Intelligence Systems: Revolutionizing Data Integration for Data Warehouses*, 2012.
- [20] A. H. Mousa, "DATA VIRTUALIZATION DESIGN MODEL FOR NEAR REA TIME DECISION MAKING IN BUSINESS INTELLIGENCE ENVIRONMENT," PhD Thesis, Computer Science Department, Universiti Utara Malaysia, Malaysia, 2017.
- [21] C. M. Barnum and S. Dragga, *Usability testing and research*: Allyn & Bacon, Inc., 2001.
- [22] A. M. Ariffin, "Conceptual design of reality learning media (RLM) model based on entertaining and fun constructs," Universiti Utara Malaysia, 2009.
- [23] S. Siti Mahfuzah, "Conceptual Design Model of Computerized Personal-Decision AID (CompDA)," Universiti Utara Malaysia, 2011.
- [24] U. Sekaran and R. Bougie, *Research methods for business: A skill building approach*: John Wiley & Sons, 2016.
- [25] P.-Y. Yen and S. Bakken, "Review of health information technology usability study methodologies," *Journal of the American Medical Informatics Association*, vol. 19, pp. 413-422, 2012.
- [26] S. Ssemugabi and R. De Villiers, "A comparative study of two usability evaluation methods using a web-based e-learning application," in *Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, 2007, pp. 132-142.
- [27] J. Nielsen, "How many test users in a usability study," *Nielsen Norman Group*, vol. 4, 2012.
- [28] A. F. Newell, A. Carmichael, M. Morgan, and A. Dickinson, "The use of theatre in requirements gathering and usability studies," *Interacting with computers*, vol. 18, pp. 996-1011, 2006.
- [29] M. Mantel, "A basic framework for cost-justifying usability engineering," *Cost-justifying usability*, p. 9, 1994.
- [30] H. X. Lin, Y.-Y. Choong, and G. Salvendy, "A proposed index of usability: a method for comparing the relative usability of different software systems," *Behaviour & information technology*, vol. 16, pp. 267-277, 1997.
- [31] J. D. Gould and C. Lewis, "Designing for usability: key principles and what designers think," *Communications of the ACM*, vol. 28, pp. 300-311, 1985.
- [32] W. Albert and T. Tullis, *Measuring the user experience: collecting, analyzing, and presenting usability metrics*: Newnes, 2013.
- [33] W. Maneesriwongul and J. K. Dixon, "Instrument translation process: a methods review," *Journal of advanced nursing*, vol. 48, pp. 175-186, 2004.
- [34] D. V. Cicchetti, D. Shoinralter, and P. J. Tyrer, "The effect of number of rating scale categories on levels of interrater reliability: A Monte Carlo investigation," *Applied Psychological Measurement*, vol. 9, pp. 31-36, 1985.
- [35] K. H. Esbensen, D. Guyot, F. Westad, and L. P. Houmoller, *Multivariate data analysis: in practice: an introduction to multivariate data analysis and experimental design*: Multivariate Data Analysis, 2002.
- [36] J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson, and R. L. Tatham, *Multivariate data analysis* vol. 5: Prentice hall Upper Saddle River, NJ, 1998.
- [37] R. A. Cummins and E. Gullone, "Why we should not use 5-point Likert scales: The case for subjective quality of life measurement," in *Proceedings, second international conference on quality of life in cities*, 2000, pp. 74-93.