Research   **Open Access**

Psychometric Evaluation of the Persian Version of Measuring the Usability of Multi-Media Software Questionnaire

Soghra Moshtaghi[1], Seyed Abolfazl Zakerian[2], Reza Osqueizadeh[3*], Pourya Rezasoltani[4], Elahe Amouzadeh[1], Sara Shahedi Aliabadi[5] and Maryam Jamshidzad[1]

[1]Department of Ergonomic, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[2]Department of Occupational Health, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran

[3]Department of Ergonomics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[4]Department of Biostatistics and Computer, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran

[5]Center of Excellence for Occupational Health, Research Center for Health Science, School of Public Health, Hamadan University of Medical Sciences, Hamadan, Iran

*Corresponding author: Reza Osqueizadeh, Department of Ergonomics, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran, Tel: (+98) 2122180119/ (+98) 09194448556; E-mail: Re.osqueizadeh@uswr.ac.ir

Abstract

Background: Multimedia systems are of considerable benefit to the current age of information technology, being beneficial for teaching purposes. Valid and reliable tools are required to assess the effectiveness of such systems. The current study aimed to evaluate psychometric properties of the Persian version of a widely held multimedia usability measurement questionnaire.

Methods and findings: The study followed a descriptive-analytical design, in which American Association of Orthopedic Surgeons (AAOS) methodology was firstly consulted to translate the original version into Persian. Content and Face Validity assessments were carried out through Lawshe's method. Construct validity was evaluated applying Exploratory and Confirmatory Factorial Analyses (EFA & CFA). Reliability was evaluated via Test-Retest methodology. A number of 357 medical university students and 10 survey instrument normalization experts were randomly invited for participation.

In measuring stability Pearson coefficient was calculated for all sub-scales as (Attractiveness: 0.598; Control: 0.534; Efficiency: 0.715; Helpfulness: 0.662; Learnability: 0.698; Excitement: 0.692). In Face Validity, Content Validity Index and Ratio were of acceptable value for all 48 questionnaire items as 0.88 and 0.94 correspondingly. Face Validity was also proved acceptable in all dimensions. Intra-Class Correlation Coefficient was calculated as 0.447 for Reliability.

Conclusions: Results show that measuring the Usability of Multi-Media Software questionnaire is valid and reliable in Persian, and has the potential to be employed in measuring software usability in Persian speaking society.

Keywords: Usability; Questionnaire; Psychometric evaluation; Multi-Media software

Introduction

In the current world of technology, knowledge and information are vastly expanded through multimedia systems. That's why this type of marketing in the world is rapidly developing. The use of multimedia in education is also beneficial [1].

Software production is an extensive domain in computer science that requires expertise and knowledge. Due to its nature, his area of proficiency generally requires lower amounts of capital investment, bringing considerable added value [2,3]. Nowadays in business world, poor software design affects customer satisfaction in marketing competition and can influence the pace of work and duty, give the users have less control over their work [4]. Usability is the level of success that has been considered for interactive systems and products [5-7].

According to the International Organization for Standardization (ISO) (1998) 9241-11, Ergonomic requirements for office work with visual display terminals (VDTs Visual Display Terminal), usability can be defined as: "effectiveness, efficiency and satisfaction, with which specified users achieve specified goals in particular environments to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use" [8-12].

Usability evaluation methods determine the problems, and are critical in ensuring quality. Usability relates to all aspects of a product such as hardware, software, visual symbols, messages, instructions, available resources and training. Each product used by human can be considered in terms of usability evaluation, such as software programs, authoring tools, and equipment and information goods [13-15].

Lack of attention to usability and it's rules would likely result in a need for system redesign, wasted user and designer energy and resources [16]. The usability is one of the most important aspects of software products. However, in practice not much attention is being paid to this issue. Because usually the knowledge, tools or time are not suitable for testing usability [17].

Due to the importance of usability, it is important to develop its measurement tools. Variety of questionnaires have widely been used to evaluate usability of interactive systems [11,18,19]. The main advantage of such methods compared to other methods is that they provide feedback from users' point of view. In addition, questionnaires are usually quick and affordable to score and administer. A large amount of data can be collected in this way and data can be used as a reliable source to check whether the usability targets have been met or not [20,21].

Measuring the Usability of Multi-Media Software (MUMMS) questionnaire was developed in response to the rapidly changing patterns and technology of computing. MUMMS questionnaire, is an exact test and validated questionnaire to measure quality of software, based on users' perspective. This questionnaire has 50 questions and 6 subscales "learnability", "helpfulness", "efficiency", "control", "excitement" and "attractiveness", that was designed in accordance with peoples' overall mentality [22].

The aim of this study was to evaluate the validity and reliability of the Persian version of MUMMS questionnaire.

Methods

The study followed a descriptive-analytical framework, and intended to psychometrically evaluate the Persian version of the MUMMS questionnaire. This questionnaire consists of 2 explanatory statements, and 48 test items, with answers ranging from "strongly agree", "agree", through "Undecided", "disagree" and "strongly disagree", systematically assessing 6 subscales including "learnability", "helpfulness", "efficiency", "control", "excitement" and "attractiveness", each of which containing 8 items.

The translation process followed the AAOS guideline recommendations [23,24]. For the initial conversion into Persian, the questionnaire was translated by two bilingual individuals (one was also a usability expert) whose first language was Persian. Then two bilingual individual whose first language was English did the reverse translation, all working independently. Finally, after completion of translation process, a panel of experts familiar with the instrument made a comparison between original version and reversely translated version [25].

To evaluate the content validity of questionnaire Lawshe's model was used. A group of experts recruited and they were asked to rate their agreements on any question in 4 aspects of judgment (necessity, relevance, clarity and simplicity of all questions). A number of 10 experts with academic and practical experience in ergonomics and software engineering formed the Content Validity evaluation panel [26,27]. Content Validity Index and Ratio were calculated, based on expert panel judgments on every single item of the questionnaire [28-30]. After content confirmation, questionnaire's Face Validity was also tested by 40 university students, following the same pattern of judgment [31]. Construct Validity was assessed through Exploratory and Confirmatory Factor Analyses, based on 240 respondent questionnaire completions. For Reliability, Test-Test was implemented and Pearson Correlation Coefficient and Intra-class Correlation Coefficient were calculated.

Results

A total of 357 participate collaborated in various phases of the study (29.2 ± 3.4 years), (42.02% males and 57.98% females). Demographic information of the participants is summarized in Table 1.

Content Validity evaluation confirmed that all 48 questions were acceptable. CVR was 0.85 and (CVI) was 0.91 for this questionnaire. Face Validity assessments confirmed 70% of participants scoring higher than 4 in the descriptive scale, although, some questions achieved acceptable levels of the score. In measuring Face Validity, 48 items of the questionnaire were considered acceptable.
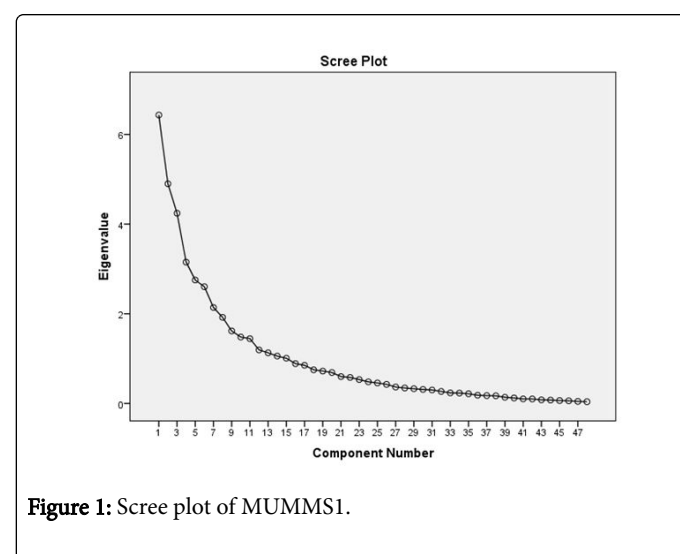
Figure 1: Scree plot of MUMMS1.

With regard to Construct Validity, the results of both EFA and CFA were presented for the MUMMS questionnaire. In this evaluation KMO (Kaiser-Meyer-Olkin Measure) measures of sampling accuracy were 0.63. The scree plot supported the uni-dimensionality of the MUMMS (Figure 1). Total variance explained of the scale was 77.23%.

In Confirmatory Factor Analysis, questions related to the six dimensions of attractiveness, control, efficiency, learnability, helpfulness and excitement. Each one containing 8 questions. Raw data were processed into AMOS software and analysed. The CFA model showed sensibly good fit indices. Chi-square<0.001, (Chi-square)/df=1.4., GFA (Group Factor Analysis)=0.94, CFI (Comparative Fit Index)=0.83, and NFI (Normed Fit Index) =0.81. Therefore the model fit was confirmed by the indices there was a good support for the one factor structure of the MUMMS.

To ensure the Satiability and Consistency of the questionnaire Reliability testing was performed. The stability of the questionnaire was measured via Test-Retest approach and Pearson's correlation coefficient and Intra-class correlation coefficient were calculated for subscales [32]. Reliability analysis of MUMMS Confirmed high stability of the questionnaire. The results of this analysis are presented in Table 2.

Specification	Abundance	Percent
Gender		
Men	150	42.02
Women	207	57.98

Age group (years)		
20-30	226	63.3
31-40	108	30.25
41-50	23	6.44
Marital Status		
Married	142	39.77
Single	215	60.22
Employment status		
Staff and students	214	59.94
only students	143	40.06
Education		
B.s	121	33.89
M.sc	194	54.34
PhD	42	11.76

Table 1: Demographic information of participants.

Sub scale	Questions	P-value	Pearson Correlation
Attractiveness	1-7-13-19-25-31-37-43	0.007	0.598**
Control	2-8-14-20-26-32-38-44	0.021	0.524*
Efficiency	3-9-15-21-27-33-39-45	0.001	0.715**
Helpfulness	4-10-16-22-28-34-40-46	0.002	0.662**
Learnability	5-11-17-23-29-35-41-47	0.001	0.698**
Excitement	6-12-18-24-30-36-41-48	0.001	0.692**
*Correlation coefficient at significance level 0.05			
**Correlation coefficient at significance level 0.01			

Table 2: Analysis of stability of questionnaire to measure the usability of multimedia software based on the Pearson correlation coefficient (001> P) and interclass correlation coefficient.

Discussion

Usability studies are of great importance, both for product designers and customers. However, few have been conducted in Persian interfaces. Lack of valid and reliable tools in Iran has raised concerns in Persian community to measure users attitudes about usability issues [33]. Information about usability methods show that clearly questionnaire one of the important method in usability studies [33,34]. So designing and development a questionnaire, for all researchers, including researchers in the field of ergonomics is of particular importance Usability evaluation of software based on questionnaire has caused measuring of this equipment expands [11].

Validation of a questionnaire in a new context new society is costly and consuming. However, questionnaire are the best method for self-report that should be considered. Questionnaire cause the information to be collected in one way [23].

One of the most important features that should always be considered when choosing a tool and focusing on it is emphasized by the experts, easy translation and optimal quality of the version provided in the target language [35]. This means that these experts always try to select and use words, phrases and sentences which avoids as much as possible ambiguous, intangible, non-transparent, and multiple meanings and this way facilitate the translation and equivalence of the text of the instrument into another language.

The most important criteria issue in this process is validity. The validity carefully measures what the questionnaire want the questionnaire has validity that accurately measures what it is intended to do [34].

In this research, face, content and construct validity of questionnaire were investigated. Face validity was considered about clarity, simplicity and the comprehensibility of each of the question and the suitability of the translation of the Persian version of questionnaire; The suitability of the questionnaire for the Persian community; Understanding the questionnaire And the property of the questionnaire to evaluate usability [36]. Results show that Persian version of this questionnaire was no apparent problem in this test and the sample group did not have a major problem in understanding the questions and almost the questions were interesting to them.

The responsiveness scale in Content Validity Instrument in this study was Lawshe's 4 scales. That's the amount of people's approval for the subject in four areas, the amount of necessity, relevance, clarity and simplicity was measured on a Likert scale of four options.

The other Actions included in this study were to improve the validity of the questionnaire Selection of Entry Indicators for the study Such as the use of computer science professionals, ergonomics, and occupational health with an ergonomic orientation. In this study was used one of the most reliable methods of content validity, Content Validity Ratio (CVR) and Content Validity Index (CVI).

In this study, factor analysis was performed which is a six-factor model with a number of 48 items to measure the usability of multimedia software it has a good fit And so, 48 items of questionnaire are on the same path [37].

The reliability of a survey instrument has always been one of the most important issues, which make it possible and should be considered by researchers. Reliability is mainly defined through accuracy and consistency [34]. In present study, internal consistency of questionnaire measured by test-retest approach and ICC. The results shows, high satiability of MUMMS questionnaire which indicates the internal consistency of whole questionnaire and its six subscales. This finding Expresses this point that the questionnaire measures same concept and has similar structure and Conceptual dispersion can't be seen in this.

This study evaluated the validity and reliability of MUMMS questionnaire in students of medical universities in Tehran. This study showed that Persian version of questionnaire for students in medical universities in Tehran has validity and reliability.

To sum up, items for the Persian version of questionnaire were evaluated in terms of clarity, simplicity and understand ability, the suitability of the translation, suitable for the Persian community, and finely benefit for usability evaluation. Due to lack of valid and reliable tools, MUMMS questionnaire can be useful to measure the usability of multi-media software; in order to secure in this field. It should be

noted that this questionnaire is not the only tool for measuring usability.

Acknowledgements

Competing and Conflicting Interests

There is no conflict of interest to be declared regarding aspects of this study.

References

1. Abbas RS (2016) Educational multi-media. http://razavi50webscom/emhtm. 17: 47.

2. Club A (2015) Artificial Intelligence and Software in Iran.

3. Mohd Yusof Ahmad Nizam ANL (2012) An Investigation on the Relationship between Online Distance Learning with Learning Usability. Procedia Soc Behav Sci 65.

4. Inger ABC, Jenny P, Mats L (2003) Why usability gets lost or usability in in-house software development. Interacting with Computers 16.

5. Martin M (2001) Methods to support human-centred design. Intern J Human-Comp Studies 48.

6. Anirban CKS, Swathi MR, Subrata G, Debkumar C (2014) Usability is more valuable predictor than product personality for product choice in human-product physical interaction. Intern J Indus Ergonom 44.

7. Muhammad KE, Mark P, Duong T, Krishna KK, Oluwabunmi T, et al. (2014) Are three methods better than one? A comparativeassessment of usability evaluation methods in an EHR. Intern J Med Inform 8: 3.

8. Bendik BGG, Eivind B (2007) Software development methods and usability: Perspectives from a survey in the software industry in Norway. Interacting with Computers 20: 10.

9. Claudia ZHS, Sonja KC, Cornelia H, Andrea K (2013) Brain Painting: Usability testing according to the user-centered designing end users with severe motor paralysis. Art Intelligence Med 59.

10. Yen Po-Yin BS (2011) Review of health information technology usability study methodologies. BMJ J 2011.

11. Jianbo XL, Qun M, Jiajie Z, Yang G (2014) The Current Status of Usability Studies of Information Technologies in China: A Systematic Study. Hindawi Publishing Corporation BioMed Res Intern.

12. Kasper H (2006) Current practice in measuring usability: Challenges to usability studies and research. Int J Human-Computer Studies 64.

13. Christopher PN (2004) Human Factors Methods for Design. CRC PRESS.

14. Azar A (2011) The effective multimedia instruction in remedy spelling disability students specific learning in Iran at year 2009. Procedia Soc Behavioral Sciences 15.

15. Dağ Funda DL, Serpil G (2014) Evaluation of Educational Authoring Tools for Teachers stressing of Perceived Usability Features. Procedia Social and Behav Sci 116.

16. Constance RJT, Jiajie Z (2004) A User-centered framework for redesigning health care interfaces. J Biomed Inform 38: 13.

17. Erik D (1998) Questionnaire Based Usability testing. Conference proceeding european software quality week.

18. Raymond DFD, Chris ND, George M, Daniel GR (2013) A usability evaluation of medical software at anexpert conference setting. Computer methods and programs in biomedicine 113.

19. Sanjiv SA, Nithin R, Marc TK (2011) Effect of font size, italics, and colour count on web usability. Int J Comput Vis Robot.

20. Panagiotis ZPA (2009) Developing a Usability Evaluation Method for e-Learning Applications: Beyond Functional Usability. Intl J Human Comp Interac 25: 75-98.

21. Roy Sharmistha KPP, Mall Rajib (2014) A quantitative approach to evaluate usability of academic websites based on human perception. Egyptian Inform J 15.

22. Measuring the Usability of Multi-Media Software.

23. Beaton E, Dorcas GF (2000) Guidelines for the Process of Cross-Cultural Adaptation of Self-Report Measures. Spine 25: 3186-3191.

24. Wild Diane GA, Mona M, Sonya E, Sandra Mc, Verjee-Lorenz A, et al. (2008) Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. Value in Health.

25. Barbosa FD, Guerreiro MM, de Souza EAP (2008) The Brazilian version of the Quality of Life in Epilepsy Inventory for Adolescents: Translation, validity and reliability. Epilepsy Behavior 13: 218-222.

26. Dianat Iman GZ, Mohammad AJ (2014) Psychometric Properties of the Persian Language Version of the System Usability Scale. Health Prom Persp.

27. Mojtaba TP, Gholamreza G, Robab S, Davood S (2015) Validity and Reliability of Self Efficacy of Health Practice Scale (SRAHPS) in Iranian Elderly. Quarterly J Sabzevar Univ Med Sci.

28. Nahad Homa RM, Farnoush J, Shohreh J, Akram P, Helnaz M, et al. (2014) Translation, Validity, and Reliability of a Persian Version of the Iowa Tinnitus Handicap Questionnaire. Iranian J Otorhinolaryngol.

29. Doretto MGM, de Souza E (2008) The Brazilian version of the Quality of Life in Epilepsy Inventory for Adolescents: Translation, validity, and reliability. Epilepsy Behavior.

30. Mengyang WL, Yu Z, Quan Z, Cheng L (2009) The Chinese QOLIE-AD-48: Translation, validity, and reliability. Epilepsy Behavior 14.

31. Cook DA, Beckman TJ (2006) Current concepts in validity and reliability for psychometric instruments: theory and application. American J Med 119: 166. e7-166. e16.

32. Arip MASM, Saad FM, Rahman AMA, Salim SSS, Bistaman MN (2013) Translation, validity and reliability of Multidimensional Self-Concept Scale (MSCS) questionnaire among Malaysian teenagers. Procedia Soc Behav Sci 84: 1455-1463.

33. Yen PY, Bakken S (2011) Review of health information technology usability study methodologies. J Amer Med Inform Assoc 19: 413-422.

34. Bandari R, Heravi KM, Rejeh N, Zayeri F, Mirmohammadkhani M, et al. Translation and validation of the critical care family needs inventory.

35. Dianat I, Ghanbari Z, AsghariJafarabadi M (2014) Psychometric properties of the persian language version of the system usability scale. Health Prom Perspect 4: 82.

36. Azizi R, Zakerian S, Rahgozar M (2015) Determining Reliability and Validity of the Persian Version of Software Usability Measurements Inventory (SUMI) Questionnaire. Intern J Occup Hygiene 5: 31-34.

37. Wang M, Wu L, Zheng Y, Zhang Q, Li C (2009) The Chinese QOLIE-AD-48: translation, validity, and reliability. Epilepsy Behavior 14: 476-480.