MECHATRONIC SYSTEM USABILITY EVALUATION

Apostolos Tsagaris (ID)

International Hellenic University (Greece)

tsagaris@ihu.gr

Abstract

This study aimed to refine and validate a Mechatronic System Usability Evaluation (MSUE) questionnaire. A total of 626 users were selected using random sampling, from the area of West Thessaloniki, Greece. The validity of the questionnaire were tested with the content and construction validity method. The reliability of the MSUE questionnaire instrument were tested using test-retest and internal consistency method. Factor analysis resulted a 25 questions questionnaire divided into five axes of, namely efficiency, effectiveness, Satisfaction, ease of use, and ease of learn. The Cronbach Alpha coefficient for the entire scale was 0.819. The questionnaire was tested and the results was shown a suitable instrument to measure usability on mechatronic systems.

To cite this article:

----

1. Introduction

The purpose of the present research is to create and validate a weighted questionnaire that measures the usability of the interaction with mechatronic systems. Mechatronics is the integration of mechanical, electrical, and computer technologies into the design of complex products (Figure 1). It is a combination of precision mechanical engineering, electronic, control engineering and computers for the intelligent control of machines. Is the process where an integrated development to the engineering design is purchased: The form of modern automation systems.

Mechatronics system is any electromechanical system that combines elements from electronic devices and the control is done with the help of software. It usually includes a microcontroller, sensors, actuators and the necessary software to control them. Many times it has a specially designed interface for control either through a PC or through special consoles. The microcontroller undertakes the processing of the signals collected by the various sensors. The collection and the acquisition of the signals through the sensors is a key part of the system. They are then processed to give the appropriate information to the system by performing various control actions.
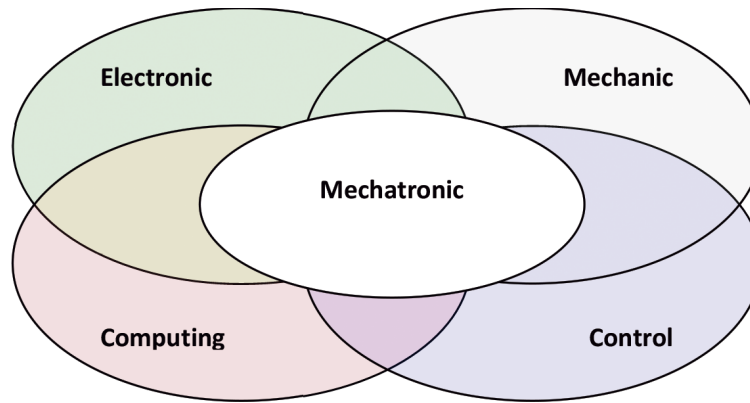
Figure 1. Mechatronic diagram

A mechatronic system is useful when it has a high degree of usability. When the interface is not easy to use, the user does not focus on the content of the interaction, but on the way to interact with the system and that is not good for the result of the process (Luo, Liu, Kuo & Yuan, 2014).

Searching in the literature one will find numerous definitions of usability. As is mentioned in various research and in international standards, usability refers to efficiency and effectiveness in achieving specific goals and user satisfaction (Chatzikyrkou, 2020). Note that the definitions of usability most often refer to Human Computer Interaction or Human Robotic Interaction or Human Machine Interaction. It is the first time speaking for Human Mechatronic System Interaction. This can be done with a combination of the above techniques, because mechatronic systems are a combination of the above areas.

According to various definitions:

> *"Usability is the degree to which a product or system can be used by specific users to achieve specific goals with efficiency, effectiveness and satisfaction in a particular context of use"* (ISO 9241)

The main dimensions of usability examined in the literature for interaction with computers, machines or robotics include ease of learning (Nielsen, 1990; Guillemette, 1995), ease of use (Nielsen, 1990; Guillemette, 1995), utility (Guillemette, 1995), user satisfaction (Nielsen, 1993), efficiency (Parsazadeh et al., 2018) and efficiency (Parsazadeh, Ali, Mehran & Tehrani, 2018).

According to ISO 9241 (Ergonomics of Human-System Interaction), usability is defined as "the extent to which a product or system can be used by designated users to achieve specific goals with efficiency, effectiveness and satisfaction, in a specific context of use" (ISO 9241-210). The individual criteria mentioned in it, are analyzed in the ease of learning, in the high efficiency of execution, in the low frequency of user errors, in the ease of retaining the knowledge of its use and in the subjective user satisfaction (Avouris, Katsanos, Tselios & Moustakas, 2015).

Dix, Finlay, Abowd and Beale (2004) in their book "Human-Computer Interaction" (3rd edition) consider usability to be a complex concept and analyze it in ease of learning, flexibility, and robustness (Koutsambasis, 2015). Shneiderman and Pleasant (2009), in their book "Designing the User Interface" (5th edition) define usability as the sum of the following measures (or metrics): time to learn, speed of performance, rate of errors by users, retention over time and subjective satisfaction (Koutsambasis, 2015). According to Avouris, the evaluation of the usability of a system includes the analysis of the characteristics of the system in relation to a certain context of use, the analysis of the interaction process and the analysis of efficiency, effectiveness and user satisfaction (Avouris et al., 2015).

Although there are many research on the usability of human computer interaction, human machine interaction and human robot interaction, there is no mention on the usability of human mechatronic system interaction, because the term mechatronic system is new, is used the last years and combines different kind of technologies, which includes part of technologies mentioned above. This research

examines the "usability" of a mechatronic system in terms of efficiency, effectiveness, Satisfaction, ease of use, and ease of learn [8, 9, 10]. (Avouris et al., 2015; Koutsambasis, 2015; Lund, 2001).

The effectiveness of a mechatronics system depends on the degree to which a system achieves its goals. User effectiveness is defined as the accuracy and completeness with which users achieve their goals using the system. It reflects the user's ability to use the system effectively and is related to his desire to use it. This factor is influenced by other factors, such as the support and management of the system, the ease of use of the system and the quality of the services it produces. It is a powerful factor of satisfaction. In other words, effectiveness refers to the extent to which users' objective goals are met. The term effectiveness is defined as the property, or ability to deliver an expected result. Efficiency mainly concerns the internal function of the interaction and expresses the resources consumed to achieve a result. How easily or difficultly one can achieve the desired result and how many resources will be consumed to achieve this. The satisfaction concerns the subjective feeling that the user derives from its use. Ease of use is the ability of inexperienced users to be able to use the system without special knowledge. The intuition of the system that allows them to make use of it without being particularly knowledgeable also plays an important role in achieving this goal. Learnability is the ability of novice users to understand how to use the system and how to get an initial level of good performance. It includes the individual properties: predictability, synthesizability, familiarity, generalizability and consistency (Koutsambasis, 2015).

The structure of the paper includes the section 2 with the methodology of questionnaire design, the section 3 with the questionnaire development, section 4 with the evaluation of the section 5 with the construction of the questionnaire and finally the conclusions where the researcher summarizes the results of the work described in the paper.

2. Methodology of Questionnaire Design

The development of the questionnaire includes several stages. As described in figure 2 includes the design of the questionnaire, the development of the questionnaire, the evaluation of the questionnaire and the final construction of the questionnaire (Figure 2).
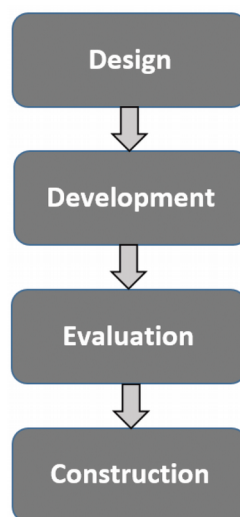

Figure 2. Questionnaire Development Methodology

Questionnaire weighting includes pilot application, validation and reliability check and related adjustments / improvements. It is necessary to conduct a pilot study with a relatively small number of participants (approximately 30-50), in order to make a preliminary assessment of the validity and reliability of the questionnaire and to correct as many errors and omissions as possible. The corrected questionnaire is then subjected to a final validity and reliability test, through a new pilot study.

Validity is the fidelity with which the attribute we want to measure counts, while reliability is the accuracy with which a questionnaire measures an attribute. The validity of a measurement scale, in simple words, refers to the degree to which what it was made to measure actually counts. There are many types of validity. In the present study, content validity and construct validity will be examined (Galanis, 2013).

Content validity refers to the extent to which a scale of measurement measures the total of the request for which it has been made. During the process of assessing the validity of the content, a record is made of the data related to the specific concept examined in a question. From all these data, the most relevant ones are selected. The authors of the questionnaire ask a pre-selected team of experts to judge each item of the questionnaire as "necessary", "useful, but not necessary" or "necessary", and then calculate the "Content Validity Ratio" for each item in the questionnaire, according to the following equality Validity is the fidelity with which the attribute we want to measure counts, while reliability is the accuracy with which a questionnaire measures an attribute.

$$content\ validity\ ratio = \frac{n_e - \frac{N}{2}}{\frac{N}{2}} \tag{1}$$

In the equation, N for an item (eg question) symbolizes the total number of experts who judge the questionnaire items and ne symbolizes the number of experts who characterize the specific item in the questionnaire as "necessary". When the content validity ratio for an element of a questionnaire is equal to 0, then half of the experts consider this element as "necessary". When the content validity ratio for an item in a questionnaire is > 0, then more than half of the experts consider this item to be "require". When the content validity ratio for an item in a questionnaire is <0, then less than half of the experts consider this item to be "required". 10 experts are used in this research. According to Galanis (2013) the minimum content validity ratio that an item must have in order to be included in the questionnaire is 0.62 (Galanis, 2013).

Construct validity refers to the degree to which a scale of measurement accurately measures the meaning we have defined as measuring. It is to some extent subjective and therefore requires a significant number of studies to be conducted in different countries, in different studied populations and at different times. In the present research, the control of the validity of a conceptual construction is achieved by factor analysis (Pereira, Maia, Marques, Bos, Soares, Gomes et al., 2008).

The reliability or, in other words, the accuracy of a questionnaire refers to the stability or, in other words, to the consistency with which the questionnaire measures the meaning or the variable that it claims to measure. Increasing the reliability of a questionnaire means reducing the random error. We must keep in mind that reliability refers to the results of measuring a scale and not to the scale itself. This means that reliability is influenced by the subjects of the research (for example by the respondents) and by the measurement protocol. Therefore a scale can be reliable in one application area and unreliable in another.

There are different types of reliability and its evaluation in this research to check the reliability of repetitive measurements or control - test (retest) and the reliability of internal consistency (internal consistency). In the reliability test with the method of repetitive measurements (test-retest) we apply, at other times, the scale twice to the same people, under the same conditions and finally we check the statistical correlation between the two scores. Calculate the coefficient r (Pearson coefficient) for test-retest.

In the control of reliability with the method of internal consistency (internal consistency) the homogeneity of the questions of the scale is evaluated. We apply the scale once because the answers are not binary but more, we use Cronbach's alpha. If the coefficient is <0.7 the internal consistency reliability is questionable, while if it is> 0.7 it is acceptable. Of course, the higher the reliability (Galanis, 2013).

3. Questionnaire Development

The study was conducted in two stages concerning the development of the standard research tool and data collection. As the work aims to collect data for the validation of the research tool, data were collected

from 626 mechatronics system users. The use involved an ATM withdrawal system that is known to many people. Similar studies have shown that the sample size is satisfactory (Hair, Anderson, Tatham, Black, 1998; Gorsuch, 1983; Kline, 1979)

An ATM is simply a terminal with two input devices (card reader, keyboard) and four output devices (microphone, monitor, receipt printer, slot and cash dispenser). These devices are connected to the processor which is the heart of the ATM. All ATMs worldwide are based on a central database system and for this reason the ATM communicates with the central processor (server) which in turn communicates with the internet service provider (ISP) which is the bridge through which All ATMs are available to the cardholder. When the user wants to make a transaction, he provides the necessary information through his card, which is "read" by the built-in card reader, and through the keyboard, which selects the type of service he needs. The ATM forwards this information to the server where it transfers the user's request to the respective bank. In case of withdrawal or deposit, the bank subtracts or adds the required amount from the user and communicates with the ATM processor to complete the transaction through the cash-receiver.

To pilot the questionnaire, a pilot application was implemented with a relatively small number of participants (25 people) and a preliminary assessment of the validity and reliability of the questionnaire was carried out, correcting as many errors and omissions as possible. Some corrections were made regarding the comprehension and correct wording of the questions, the determination of the appropriate time to complete, the interest of the respondents and the overall appearance. The corrected questionnaire was then subjected to a final validity and reliability check.

4. Evaluation of Questionnaire

Validity included validity of content and validity of conceptual construction. In Content Validity, a pre-selected team of 10 experts was asked to rate each question in the questionnaire as "necessary", "useful but not necessary" or "unnecessary". The "Content Validity Ratio" was then calculated for each question according to equation 1.

Questions answered as required with a validity ratio of > 0.62 were retained and incorporated into the proposed usability assessment questionnaire. Those deemed necessary with a validity ratio of <0.62, or deemed useful but unnecessary or unnecessary were discarded and deleted from the tool (Galanis, 2013).

In the next phase, the validity of the conceptual construction was evaluated. Factor analysis was used since the data had many dimensions (scale 1-5). Applying factor analysis to the data of a questionnaire, 5 factors emerged that express individual dimensions of the concept measured by the study questionnaire. These factors emerged based on the correlations presented between the various elements of the questionnaire. Thus, a broad concept was simplified and grouped into individual parts to make it clearer. Table 1 below shows the questions and factors.

By examining each factor item according to the model and literature framework, items under factor 1 can be placed under the aspect of efficiency, factor 2 under the aspect of effectiveness, factor 3 under the aspect of Satisfaction, factor 4 under the aspect of ease of use and factor 5 under ease of learn. Items loaded from each of the five components have strong, clear, and conceptual links.

In order to analyse the valid items for each component, a Kaiser–Meyer–Olkin (KMO) test and Bartlett's test of sphericity have been carried out. Table 2 shows that the KMO test resulted in a value of 0.916. This value exceeded the recommended value of 0.6 (Kaiser, 1970; Kaiser, 1974), indicating that the sample was adequate to test the factor analysis.

The test-retest method was used to check the reliability. Thus, the stability of the answers was evaluated. The same scale was applied with a time difference of 15 days to the same individuals (20 individuals) under the same conditions and the statistical correlation between the values of the responses was checked. The pearson coefficient calculated was 0.923 which means a very large correlation.

	Questions	Factors				
		1	**2**	**3**	**4**	**5**
Q1	The design of interaction is simple with the essentials.	0.847				
Q2	The response speed of the system is satisfactory.	0.783				
Q3	The execute of the command is accurately.	0.821				
Q4	The Interaction is complete.	0.904				
Q5	It works as expected.	0.791				
Q6	The given information by the interface is satisfactory.		0.837			
Q7	The interaction is easy.		0.868			
Q8	A wide range of commands is covered.		0.793			
Q9	There is freedom of movement by the user.		0.862			
Q10	The change of command is easy.			0.848		
Q11	There is flexibility in interaction.			0.786		
Q12	The use provides satisfaction.			0.846		
Q13	It is flexible in the choice of movements.			0.831		
Q14	Interaction is not affected by environmental conditions.				0.871	
Q15	The equipment does not make movements difficult.				0.823	
Q16	Operating the system is not tedious.				0.799	
Q17	Commands are easy to remember.				0.841	
Q18	It is easy to use.				0.787	
Q19	It is not tedious to use.				0.773	
Q20	No mistakes are made during the interaction.				0.826	
Q21	I learned to use it quickly.					0.707
Q22	I easily remember how to use it.					0.841
Q23	I quickly became a capable user.					0.785
Q24	It is easy to learn.					0.837
Q25	It has no difficult functions.					0.773

Table 1. Principal Component Analysis

Test	Value
Kaiser-Meyer-Olkin Adequacy Sampling Measure	0.916
Bartlett's Test of Sphericity	
Approx. Chi-Square	5135.417
df	190
Sig.	.000

Table 2. Kaiser–Meyer–Olkin (KMO) and Bartlett's test.

Questions	Factor	Cronbach Alpha
Q1 - Q5	Efficiency	0.829
Q6 - Q9	Effectiveness	0.840
Q10 - Q13	Satisfaction	0.828
Q14 - Q20	Ease of Use	0.817
Q21 - Q25	Ease of Learn	0.789

Table 3. Cronbach Alpha

The internal consistency reliability was then checked by calculating the Cronbach alpha coefficient. The coefficient was calculated at 0.819 for the whole questionnaire, while for each axis separately the coefficient values are shown in Table 3.

In all cases we see that we have values greater than 0.78 and in some values they reach up to 0.84. This means that the questions have a very high internal coherence both in each axis and in the overall questionnaire.

As can be seen from the analysis of the validity and reliability of the questionnaire, it can be concluded that is a tool that can be used to measure the usability of the interaction.

5. Questionnaire Construction

The questionnaire starts with a report on the details of the survey. It consists of two parts, one referring to the general information of the participants and the other to the details of the usability assessment.

In the general questions the participants were asked about their gender, age and educational level. These data are crucial information for the results of the research, thus identifying the independent variables of the research. The second part questions refer to the details of the evaluation of the usability of mechatronic systems divided into the 5 axes analyzed above. The questionnaire was constructed using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree).

The questions Q1-Q5 refers to Efficiency and Q6-Q9 to the Effectiveness. The Q10-Q13 refers to the Satisfaction, while the questions Q14-Q20 refers to the Ease of Use. Finally the questions Q21-Q25 refers to Ease of Learn.

6. Conclusions

This paper presents the construction and testing of a questionnaire that can be used to assess usability in interaction with mechatronic systems. Combining efficiency, effectiveness, ease of use, ease of learning and satisfaction, the questionnaire was built to detect the overall usability of a system. All checks on the validity and reliability of the questionnaire were performed and the indicators showed that it is an essential tool for evaluating usability. For the test were used among the other 626 users and the value of the Cronbach alpha coefficient was calculated at 0.819 given a very high validity and reliability of the tool. As further work of the research could be mentioned the application of this questionnaire in the interaction with computer integrated systems, such production lines, construction lines, robotic assembly lines, etc. in order to optimize the interaction with them.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest concerning the research, authorship, and/or publication of this article.

Funding

References

Avouris, N., Katsanos, C., Tselios, N., & Moustakas, K. (2015). *Introduction to Human-Computer Interaction*. Heallink, Ed. Athens: SEAB. Available at: http://hdl.handle.net/11419/4224

Chatzikyrkou, M., (2020). Usability evaluation of mechatronic system by primary school children. *7th ICMMEN International Conference*. Thessaloniki, Greece. https://doi.org/10.1051/matecconf/202031801026

Dix, A.J., Finlay, J.E., Abowd, G.D., & Beale, R. (2004). *Human-Computer Interaction* (3rd Edition). Prentice-Hall.

Galanis, P. (2013). Validity and reliability of questionnaires in epidemiological studies. *Archives of Hellenic Medicine*, 30(1), 97-110

Gorsuch, R.L. (1983). *Factor Analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Guillemette, R.A. (1995). The evaluation of usability in interactive information systems. In: Carey, J.M. (Ed.). *Human factors in information systems: Emerging theoretical bases* (207-221). Norwood, NJ: Ablex.

Hair, J.F.., Anderson, R.E., Tatham, R.L., Black, W.C. (1998). Multivariate Data Analysis (5th ed.). Upper Saddle River, NJ: Prentice Hall.

ISO 9241 (2010). *Ergonomics of Human-System Interaction*. International Organization for Standardization.

ISO 9241-210 (2010). *Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems.* International Organization for Standardization. Available at: https://www.iso.org/standard/52075.html (Accessed: August 2018)

Kline, P. (1979). Psychometrics and Psychology. London, UK: Academic Press.

Kaiser, H. (1970). A second-generation Little Jiffy. *Psychometrika* 35, 401-415. https://doi.org/10.1007/BF02291817

Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika* 39, 31-36. https://doi.org/10.1007/BF02291575

Koutsambasis, P. (2015). Evaluation of user-centered interactive systems. Seab Ed. Available at: https://repository.kallipos.gr/handle/11419/2765

Lund, A., (2001). Measuring usability with the USE questionnaire. *Usability Interface,* 8(2), 3-6.

Luo, G.H., Liu, E.Z.F., Kuo, H.W., & Yuan, S.M. (2014). Design and implementation of a simulation-based learning system for international trade. *The International Review of Research in Open and Distance Learning*, 15(1). https://doi.org/10.19173/irrodl.v15i1.1666

Nielsen, J. (1990). *Evaluating hypertext usability*. In Jonassen, D.H., & Mandl, H. (Eds.), *Designing hypermedia for learning* (147-168), Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-642-75945-1_9

Nielsen, J. (1993). *Usability Engineering*. London: Academic Press. https://doi.org/10.1016/B978-0-08-052029-2.50009-7

Parsazadeh, N., Ali, R., Mehran, R., & Tehrani, S.Z., (2018). The construction and validation of a usability evaluation survey for mobile learning environments. *Studies in Educational Evaluation,* 58, 97-111. https://doi.org/10.1016/j.stueduc.2018.06.002

Pereira, A.T., Maia, B.R., Marques, M., Bos, S.C., Soares, M.J., Gomes, A. et al. (2008). Factor structure of the Rutter Teacher Questionnaire in Portuguese children. *Revista Brasileira de Psiquiatria,* 30(4), 322-327. https://doi.org/10.1590/S1516-44462008000400004

Shneiderman, B., & Plaisant C. (2009). *Designing the User Interface* (5th Edition). Addison-Wesley