Proceedings of the of the 17th Conference on Computer Science and Intelligence Systems pp. 213–216
DOI: 10.15439/2022F252
ISSN 2300-5963 ACSIS, Vol. 30

User Experience and Multimodal Usability for Navigation Systems

Lumbardha Hasimi
Institute of Information Technology,
Lodz University of Technology, Lodz, Poland
ORCID: 0000-0003-3001-350X
Email: lumbardha.hasimi@dokt.p.lodz.pl

Aneta Poniszewska-Marańda
Institute of Information Technology,
Lodz University of Technology, Lodz, Poland
ORCID: 0000-0001-7596-0813
Email: aneta.poniszewska-maranda@p.lodz.pl

***Abstract*—User experience as a concept of human-computer interaction is crucial for the evaluation of systems and applications. Every new generation of navigational systems provides new features and extended functionality, which has additional functions that can oftentimes confuse the primary information on the system's functionality. The conducted experiment analyses and observes available versions of navigation systems such as Garmin Drive 50 and/or TomTom Go, which are offered with certain advantages on features. The paper presents the selected aspects regarding the implementation, design, environment, recruitment, tests, and evaluation of the navigation systems, concluding from the results that there is difference between audio and visual mode.**

***Index Terms*—navigation systems, human-computer interaction methods, user experience, multimodal usability**

I. Introduction

New generation of navigational systems provide new features and extended functionality, which can sometimes be source of confusion on the primary information of the system's functionality. This often times lead to poor usability, especially when it comes to operations of the primary functions [1], [5]. In order to design a proper environment with effect on user satisfaction, the usability aspect should be taken into account. This is to ensure successful performance on the primary operations of the system, as well as on the upgraded versions, regardless of the users' experience and ability with it. User experience as a concept of human-computer interaction is a fundamental concept in the evaluation of systems and applications, particularly because it affects issues such as usability, cognitive load, affective experiences, mental demand, efficiency, etc.

The evaluation of usability is quite diverse. As an area, it is under uninterrupted and active development, meaning new approaches and procedures come as continuous expansion of the practice into several contexts [2]. Hence, the testing methodologies can be different based on aim, place, and moderating style. However, in general there are two evaluation methods on usability, namely qualitative and quantitative [2], [6]. The quantitative evaluation is required to generally get numerical values that represent the level of usability, whilst qualitative evaluation is required to identify and examine issues and problems. In our work, we focus mainly on the quantitative method as a primary evaluation, namely surveys designed for specific tasks.

The concept of cognitive load, in the other hand, referring to the amount of effort required while performing an action, is an aspect taken into consideration mainly because of potential interference with the processes in the use of a system [17]. Considering that an ideal application or system ensures an interface that keeps user cognitive load to a minimum, what we have considered as an important indicator [18].

Additionally, mental workload, considered as a demand placed on the user by the system, depends on many parameters, which lead to the fact that the result of a task performed is subjective to the users' experience and the ability [7]. For this reason, it was determined that during the recruitment process the experience, initial capacities, reaction time and fatigue will be taken into account. The conducted experiment analyses and observes available navigation systems such as Garmin Drive 50 and/or TomTom Go, which are available and with certain advantages on features, although not as widely used as the most traditional available Maps. The compactible devices used were Camper/RV 890 and/or TomTom GO 60S Automotive GPS Navigation Device. The main reasons for carrying out this experiment are to find out the satisfaction and efficiency of the applications used and how much of the aimed aspects are satisfactory. This experiment is carried out aiming to analyse and compare user experience on aspects as efficiency, usability, and cognitive load on system's multimodal mode.

The defined null hypothesis assumes that there is slightly no difference between visual or audio mode of feedback in the user experience. Hence, it was decided on three defined dependent variables, namely efficiency, usability, and cognitive load, where the efficiency regards how fast the user finished the task, taking into account its completion time. Furthermore, usability as the second variable is defined as how easy the feedback is learned by the user. Finally, the mental and/or physical load for the tasks being taken by the user and the difference between audio and visual feedback. The independent variable is the feedback type with the given conditions on the audio and visual feedback treatments, shown in table I. The environment of the design is certainly basic, as it contains only one independent variable with two possible conditions based on the mode of the feedback. The design would be affected by different user features, such as background, accessibility, previous experience which affects efficiency and other minor variables [5], [14], [15].

Aspects regarding the implementation, design, environment, recruitment, tests and evaluation are thoroughly discussed in the next sections. The paper is structured as follows: section 2 presents the recruitment process, selected profile and co-founding variables together with the methods, tasks and setup. Section 3 deals with the result, statistics and significance of each test performed. Finally, section 4 gives an overall conclusion and discussion on final findings.

II. Methodology of the study

The recruitment design is selected considering different factors, concentrating especially on the depended variable – the user experience, which should consider that has a wide range of variations. In the view of certain complex tasks that require information retrieval from the system, problem-solving skill of a certain level is necessary for multimodal feedback. Certainly, some individual users might have limitations, disabilities, or accessibility issues, hence within group design would be best to eliminate such differences [3], [4], [16].

Participant's profile is selected deciding in various factors in order to minimize the confounding factors:

- Users with similar experience in using such systems.
- User of similar age group.
- User with/without IT background.
- Accessibility or level of disability (visual or audio issues/impairments).

Training and overview of the proposed navigation system is given to the user and an orientation session conducted to make participants familiar with the tasks required and how to tackle it. Moderators are not allowed to interfere during any ongoing activity, only in case of error issues, to allow completion of the activity.

In this setting the dependent variable is User Experience for the following categories, i.e., Efficiency, Usability and Cognitive Load. Given it with only one independent variable which is the type of feedback for the chosen systems, taking into consideration two conditions/states: Audio and Visual feedback for the respective system.

Whereas the environment of the carried experiment consisted of several elements. The first session was conducted to assess the navigation system usability, first by being introduced to the use of the system. Secondly, specific tasks for a specific period of time and path. It was followed by a session with the test environment where participants were asked to explore the virtual landscape freely and navigate from one place to another. The participants were instructed to try all of the navigation functions provided.

A short questionnaire was handed out after the road/path completion and the participants answered the questions within d minutes. The moderator followed a set of predefined questions during the later session. Preparation required before the study starts includes the consent form, demographic questionnaire, participants coding and preparing all other questionnaires.

The tasks described in table I shall be assigned to the user walking or riding a car. This was not given, and in order to avoid any confounding variables, it could be provided.

Conditions intended are audio and visual feedback. The order of the condition to which the participant was assigned is the key to avoiding other factors or confounding variables. To avoid variables or other factors on the evaluation of dependent variable, it is very important to use randomization in order to lessen the effects. Such variables can be fatigue, or the learning curve while trying the tasks of audio than visual mode in order.

The conditions are counterbalanced among 20 participants using a simple randomization technique. The certain sequencing considering only two options of sequencing: $C1ßC2$ or $C2ßC1$, having in mind the counterbalance of the order effect, for $n = 20$, gives times 10 for each sequence, hence it was assigned which participant gets what sequence in order, having randomly "1" or "0" assigned and counted equal, followed by code of sequences.

After conducting the test, every participant filled in two questionnaires, one for the SUS (System Usability Scale) and second for TLX (Task Load Index). Using the adopted version of the questionnaire for system usability scale to measure the user experience for both audio and visual feedback [6], [7] as presented in the tables II and III. Similarly, some questions were adopted from user satisfaction questionnaire GUESS [9].

The questions defined in table III are required as a post questionnaire for participants who completed the tasks that had been discussed to measure the usability. For the efficiency measurement [8], the evaluation was done taking into consideration two features: the task completion time and the speed of conducting it, taken into units; time in seconds and speed per minute respectively. After the measurement of efficiency for both modes of the system, using T-Test was possible to compare the two means of speed and task completion for audio and visual feedback.

III. System Usability Scale (SUS) Tests of Navigation systems

Usability studies are well suited for gathering qualitative or quantitative behavioural data and for answering design-related questions. Considering that our setup is rather focused on collecting feedback on the features rather than attitudinal feedback [10], [11], [12], it is well fit for our purpose. Nevertheless, data collected through methods mentioned, was later on described and analysed with Paired Sample T-Test. System usability score was calculated based on the questioners filled by participants [13].

Table IV contains the results from preliminary questionnaire, calculated and elaborated accordingly to the System Usability Scale (SUS) test score. In a general view the Visual mode tends to draw higher values than the Audio one (Fig. 1). Similarly, as seen in the results from boxplot the mean SUS scores of visual values are higher than the audio values.

TABLE I
TASKS AND SETUP.

Scheduled routes	Schedule a route in advance, either by given destination, decided route according to the path and time given	$\approx 45min$
Specific destination	Type in an address from a location outside of the city center (e.g., Suburban area in proximity to the city) you want to travel and follow a direction on the navigation system using the feedback type assigned.	$\approx 90min$
Find a nearby medical center	Find closest hospital/clinic specialized in surgery and follow feedback from a navigation system	$\approx 45min$
Completion	Wrap-up, archiving logs and documents, collection of results.	$\approx 45min$

TABLE II
SYSTEM USABILITY SCALE QUESTIONS.

Sentence	**Score (1-5)**
1. I think that I would like to use this system frequently.	Score (1-5)
2. I found the system unnecessarily complex.	Score (1-5)
3. I thought the system was easy to use.	Score (1-5)
4. I think that I would need the support of a technical person to be able to use this system.	Score (1-5)
5. I found the various functions in this system were well integrated.	Score (1-5)
6. I thought there was too much inconsistency in this system.	Score (1-5)
7. I would imagine that most people would learn to use this system very quickly.	Score (1-5)
8. I found the system very cumbersome to use.	Score (1-5)
9. I felt very confident using the system.	Score (1-5)
10. I needed to learn a lot of things before I could get going with this system.	Score (1-5)

TABLE III
TASK LOAD INDEX QUESTIONNAIRE TO ASSESS THE WORKLOAD OF TASKS FOR GIVEN CASES.

Subject	**Question**	**Scale elaboration**
Mental Demand (MD)	How mentally demanding was the task?	Very low to Very high
Physical Demand (PD)	How physically demanding was the task?	Very low to Very high
Temporal Demand (TD)	How hurried or rushed was the pace of the task?	Very low to Very high
Performance (P)	How successful were you in accomplishing what you were asked to do?	Failure to Perfect
Effort (E)	How hard did you have to work to accomplish your level of performance?	Very low to Very high
Frustration (F)	How insecure, discouraged, irritated, stressed, and annoyed were you?	Very low to Very high

TABLE IV
CALCULATED SUS SCORES FOR VISUAL AND AUDIO FEEDBACK.

PID	**Visual mode**	**Audio mode**
1	55	27.5
2	80	37.5
3	30	12.5
4	70	100
5	85	32.5
6	72.5	85
7	95	100
8	50	45
9	97.5	50
10	80	95
11	67.5	42.5
12	100	100
13	57.5	7.5
14	62.5	82.5
15	82.5	45
16	57.5	22.5
17	42.5	30
18	95	95
19	40	37.5
20	82.5	45

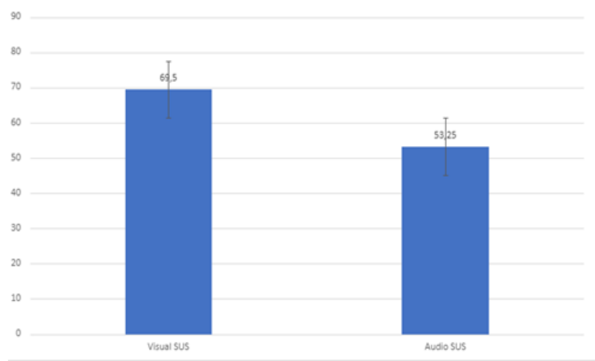

Fig. 1. Boxplot showing mean SUS scores of visual and audio values.

Measure 1		Measure 2	t	df	p
Visual SUS	-	Audio SUS	2.778	19	0.012

Fig. 2. Paired sample t-test.

Considering that only statistics are not enough to shows the significance in the system usability, paired samples T-Test shall be applied. Furthermore, Standard Error of Mean (SEM) and Standard Deviation (SD) are higher for audio over the visual feedback for which the values in audio are not very close to mean values, compared to ones in visual feedback (Table V).

Conducted Paired Samples T-Test on calculated SUS scores resulted in significant effect, $t(19) = 2.778, p = 0.012$

TABLE V
DESCRIPTIVE STATISTIC FOR THE DUAL MODE.

Descriptive Statistics		
System Usability Scale	**AUDIO**	**VISUAL**
Valid	20	20
Missing	0	0
Mean	53.25	69.50
Standard Error of Mean (SEM)	7.30	4.51
Standard Deviation (SD)	32.66	20.19
Variance	1067.17	407.63
Minimum	7.50	30.00
Maximum	100.00	100.00

TABLE VI
DESCRIPTION STATISTICS FOR VISUAL AND AUDIO FEEDBACK.

	N	**Mean**	**Standard Deviation**	**Standard Error**
MDVIS	20	6.857	4.214	0.796
MDAUD	20	9.250	5.803	1.097
PDVIS	20	7.893	4.653	0.879
PDAUD	20	6.964	4.647	0.878
TDVIS	20	4.821	3.389	0.640
TDAUD	20	5.857	4.461	0.843
PVIS	20	8.750	4.436	0.838
PAUD	20	9.000	4.431	0.837
EVIS	20	8.500	5.037	0.952
EAUD	20	9.679	4.877	0.922
FVIS	20	5.750	4.024	0.761
FAUD	20	9.821	5.976	1.129

(Fig. 2). Results of Paired Samples T-Test conclude that the differences between is significant which means that the population (SUS scores for different feedback systems) have intrinsic differences. The p-value of 0.012 is much smaller than 0.05, so it is possible to reject the null hypothesis of no difference between type of feedback in system usability and consider with a high degree of confidence 98.8% that visual type of feedback is giving better usability. It is shown in table VI, that Fatigue and Mental Demand have significance, results have *p* value smaller than 0.05, so we can say high confidence that there is difference in Mental Demand and Fatigue factor effected user experience for both Audio and Visual feedback.

The observed result were $t(27) = -3.170$ and $p = 0.004$ for the Mental Demand, whilst for the Fatigue $t(27) = -3.663$ and $p = 0.001$. So null hypothesis that there is no difference in user experience between audio and visual feedback is rejected. Due to the significance test results applied.

IV. CONCLUSIONS

The main goal of this paper was to find out whether there is a difference in user experience and system usability, in case of audio and visual feedback of navigation systems. The paper further investigated how using audio or visual mode impacts the variables user experience, usability and cognitive load in given setup. The results of the experiment show that there is a considerable difference in user experience for the System Usability and task cognitive load for the two types of audio and visual feedback for navigation systems, as shown in the results of the tests engaged. As for system usability dependent variable Visual Feedback was preferred among all participants with 98.8% confidence, whereas the degree of freedom showed the result: $t(19) = 2.778, p = 0.012$. On the other hand, for the Task Cognitive Load as dependent variable, the results showed that only two factors have significant effect in its Task Load, namely Fatigue and Mental Demand where *p* resulted smaller than 0.05, what leads to confidence of 99% that Audio type feedback requires more mental demand and causes more Tiredness to the users. These data resulted on given scores: $t(27) = -3.170, p = 0.004$ for Mental Demand and $t(27) = -3.663, p = 0.001$ for Fatigue. Consequently, the results lead to the rejection of null hypothesis that there is no difference in user experience between audio and visual feedback.

REFERENCES

[1] P. Zalewski and B. Muczynski, Extended Framework for Usability Testing in e-Navigation Systems, *TransNav, the International Journal on Marine Navigation and Safety of Sea Transportation*, Vol. 10, pp. 43–48, 2016.

[2] G. Papatzanis, Evaluating usability evaluation methods for location-aware interactive systems in contextually rich environment, https://qmro.qmul.ac.uk/xmlui/handle/123456789/1275, 2011.

[3] R.L. Charles and J. Nixon, Measuring mental workload using physiological measures: A systematic review, *Applied Ergonomics*, Vol. 74, pp. 221–232, 2019.

[4] C. Marchand, J.B. De Graaf and N. Jarrasse, Measuring mental workload in assistive wearable devices: a review, *J NeuroEngineering Rehabil*, Vol 18(10), pp. 160, 2021.

[5] C.M. Barnum, Preparing for usability testing, in Usability Testing Essentials, *Ed. Morgan Kaufmann, 2nd edition*, 2021.

[6] P. Laubheimer, Beyond the NPS: Measuring Perceived Usability with the SUS, NASA-TLX, and the Single Ease Question After Tasks and Usability Tests, Nielsen Norman Group, https://www.nngroup.com/articles/measuring-perceived-usability/, 2018.

[7] A.J. Lazard, J.S.B. Brennen and S.P. Belina, App Designs and Interactive Features to Increase mHealth Adoption: User Expectation Survey and Experiment, *JMIR Mhealth Uhealth*, Vol. 9(11), pp. e29815, 2021.

[8] R. Minelli, A. Mocci and M. Lanza, Measuring Navigation Efficiency in the IDE, *7th International Workshop on Empirical Software Engineering in Practice (IWESEP)*, pp. 1–6, 2016.

[9] M.H. Phan, J.R. Keebler and B.S. Chaparro, The Development and Validation of the Game User Experience Satisfaction Scale (GUESS), *Human Factors*, Vol. 58(8), pp. 1217–1247, 2021.

[10] J. Sauro, Measuring Usability with the System Usability Scale (SUS), *MeasuringU*, 2011.

[11] S. Radmard, A.J. Moon and E.A. Croft, Interface design and usability analysis for a robotic telepresence platform, *Proceedings of 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 511–516, 2015.

[12] M. Schrepp, A. Hinderks and J. Thomaschewski, User Experience mit Fragebögen evaluieren – Tipps und Tricks für Datenerhebung, Auswertung und Präsentation der Ergebnisse, *UP 2016, Gesellschaft für Informatik e.V. und die German UPA e.V.*, 2016.

[13] K. Tcha-Tokey, O. Christmann, E. Loup-Escande and S. Richir, Proposition and Validation of a Questionnaire to Measure the User Experience in Immersive Virtual Environments, *IJVR*, Vol. 16(1), 2016.

[14] B. Borowska, Learning Competitive Swarm Optimization, *Entropy*, Vol. 24(2), 283, MDPI, ISSN 1099-4300, pp. 1–17, 2022.

[15] T. Galaj, F. Pietrusiak, M. Galewski, R. Ledzion and A. Wojciechowski, Hybrid Integration Method for Sunlight Atmospheric Scattering, *IEEE Access*, Vol. 9, Publisher IEEE, ISSN 2169-3536, pp. 40681–40694, 2021.

[16] S. Zakrzewski, B. Stasiak, T. Klepaczka and A. Wojciechowski, VR-oriented EEG signal classification of motor imagery tasks, *Human Technology*, Vol 18 (1), ISSN 1795-6889, pp. 29–44, 2022.

[17] F. Paas, J. E. Tuovinen, H. Tabbers and P. W. M. Van Gerven, Cognitive Load Measurement as a Means to Advance Cognitive Load Theory, *Educational Psychologist*, vol. 38(1), pp. 63–71, Jan. 2003.

[18] K. Whitenton, Minimize Cognitive Load to Maximize Usability, Nielsen Norman Group. https://www.nngroup.com/articles/minimize-cognitive-load/ (accessed Jul. 14, 2022).