©2025 Published in 11th International Symposium on Intelligent Technologies in Engineering and Science 16-18 May 2025 (ISITES2025 Diyarbakır - Turkey)

Comparative Usability Testing Using CogTool: A Case Study Approach

[1]Bayram Alkan, [2*]Zeynep Behrin Guven Aydin and [3]Ruya Samli
[1]Monster Notebook, Istanbul, Turkey
*[2]Faculty of Engineering, Department of Software Engineering Dogus University, Turkey
[3]Faculty of Engineering, Department of Computer Engineering Istanbul University-Cerrahpasa, Turkey

Abstract

In this study, the usability modeling and usability test comparison of the mobile applications of different users and 2 phone providers were made. This modeling and comparison test was carried out with the CogTool tool for usability testing. In the study, 3 different levels of users were expected to perform a given task on the mobile applications of two different telephone line providers. As a duty, balance top-ups are given to the phone lines in the mobile applications of these providers. Looking at the results, the usability test with the CogTool tool took 14.2 seconds for one provider and 7.3 seconds for the other provider. Despite being given the same task, the time has nearly doubled because a provider is missing a step. At the same time, the completion times of 3 different levels of users are compared and given in tables

Key words: Human Computer Interaction, Cognitive Modelling, Usability, CogTool

1. Introduction

Considering that people and computers interact in almost every area today, it is very important that each interface, which is the point of contact, is easy to use [1]. Users are in constant communication with computers or computer systems through various interfaces. This interface can be a graphical interface, as well as various input and output devices such as keyboard, mouse, speaker, microphone. Human-Computer Interactions (HCI) designers take into account usability and accessibility, highlighting the user experience in order for the interaction between these interfaces and users to be effective. Because a poorly designed interface causes users to have difficulties in performing the operations they want to perform. Therefore, the purpose of HCI is to significantly increase and improve the usability and accessibility of computers or similar systems [2]. Usability, which is one of the critical points of HCI, has become one of the most important quality criteria for websites [3]. Usability has various definitions according to different disciplines. Usability according to the standard International Organization for Standardization (ISO) 9241-11: 2018 determined by the ISO; it is defined as the effectiveness, efficiency and user satisfaction of a

*Corresponding author: Zeynep Behrin GUVEN AYDIN Address: Faculty of Engineering, Department of Software Engineering, Dogus University, Dudullu Campus 265/ 1, 34775 Umraniye / Istanbul TURKEY. E-mail address: zguven@dogus.edu.tr

product by a certain user group, within the context and purposes determined (ISO 9241-11:2018, 2018). It is emphasized that a usable design is possible with a good understanding of based interaction [4]. Computer users generally aim to achieve a certain purpose, whether it is a job application or decoration. It can be for scientific studies or for social media. Computer users cannot concentrate on a certain point for a long time, except in very rare cases, they cannot stay in front of an application for hours. Experienced and inexperienced users show different behaviors [5]. While HCI examines the interaction between social and behavioral sciences and information technologies, the most frequent focus is usability [6].

Taşar's study of the scope, 11 questions were asked to 59 participants who have received or are studying in the field of genetics and biology, and the processes and preferences of people working in the field of genetics when interacting with the two most used global genetic data environments. Based on the results of this survey, the usability of these data environments for similar operations was also analyzed with the CogTool tool [7].

Jorritsima et al aimed to assess the accuracy of cognitive modeling tools such as KLM (Keystroke-Level Model), GOMS (Goals, Operators, Methods, and Selection rules), and CogTool in predicting human performance differences across various interface alternatives. The researchers modeled three tasks across three interface options using these tools and compared the predictions to the actual performance data collected from 20 expert users completing these tasks on the interfaces. The findings revealed a discrepancy between the predictions made by the models and the actual human performance. Specifically, in cases where the models predicted faster performance on Interface A compared to Interface B, humans performed faster on Interface B instead in about one-third of the instances. This outcome challenges the reliability and validity of these cognitive modeling tools in the context of interface design practice. It suggests that the predictions made by these models might not consistently align with real-world human performance, highlighting limitations or potential gaps in their ability to accurately forecast user behavior across different interfaces. The implications of these findings point toward the need for further investigation and refinement of these modeling tools to enhance their accuracy and applicability in predicting human performance in interface design. It highlights the complexity of human-computer interaction and the challenges in developing models that precisely capture the intricacies of user behavior across diverse interfaces. Cognitive modeling tools like KLM, GOMS, and CogTool are valuable in interface design, but this study's results caution against relying solely on their predictions without considering real user testing and empirical data to validate their accuracy in specific contexts [8].

In Jeremy Ludwig's study, the keypads on the keypad programmable interfaces of different helicopters were examined, and CogTool and Omia cognitive modeling tools were utilized. The study focused on comparing the MH-60S (Sierra) and MH-60R (Romeo) helicopters, which have different sets of private keys on their interfaces, particularly concerning alphanumeric data entry

methods. The significant difference between the two sets of private keys lies in the method of alphanumeric data entry. Sierra provides a complete set of more than 26 keys, each containing one letter, whereas Romeo offers a condensed mobile phone-style arrangement where a single key includes three letters and one number. The study compared the performance results of these two different interfaces using alphabetical data entry tasks, which are currently part of the training for Sierra helicopters. According to the findings presented in the document, the initial Romeo model takes approximately twice as long to input data compared to the original Sierra model [9].

Demirci aimed to provide an understanding between the findings by analyzing the relationship between the perceived usability of the interface designs of shopping sites on users in Turkey and the measured usability of websites with the help of CogTool. To achieve this aim, firstly, 25 women were surveyed via e-mail to determine the most preferred online shopping sites. Then to the participants, in order to measure the usability of the websites, they were given the task of purchasing the same shoes on the website. Usability testing was performed on the selected websites with CogTool and as a continuation of the first phase of the study, user testing was conducted with ten of 25 female users [10].

Yamira and Kara described CogTool models for two pilot crews performing two different types of data link clearance acceptance missions and on two different simulation platforms time estimates for accepting and executing CogTool's required arrival and interval management approvals were compared with empirical data observed on videotapes and recorded in simulation files. The results show that there is no statistically significant difference between empirical data and CogTool estimates [11].

2. METHOD

2.1 Problem

The aim of this study is to determine how close the CogTool tool is to the truth because of the participants performing the tasks in line with some tasks. For this, the participants who participated in the study read some instructions given to them so that the given tasks were completely correct and error-free. Then they performed the given tasks via smartphones. As a result of the study, the people fulfilled their basic duties, and in case of any mistake while performing the tasks, it was decided to do the tasks again and a path was followed accordingly.

2.2 Participants

The research group consists of 3 people, 2 female and 1 male user, who actively use the mobile applications of two different phone service providers. This group is the people who have been performing mobile applications with a smartphone for a while. The ages of the participants are between 22-45.

2.3. Tasks

Within the scope of the study, the participants are asked to perform the necessary steps to top up their account using the mobile application of two different telephone service providers. The names of this task given to the participants are given in the table, and these steps are shown in the pictures as screenshots. Users are expected to be logged into the application for the task to be performed. While selecting the task, the most used feature in the application was selected. Task results are also detailed in the Findings section in tabular form.

Table 1. Task Steps

Task Order	Tasks
1	Entering the Lira (TL) Loading Page
2	Choosing Lira (TL) Loading Process
3	Choosing a Number
4	Choosing Among
5	Making Pay

2.4 CogTool Tool

CogTool is indeed a cognitive modeling tool developed by Carnegie Mellon University, but it doesn't use the ACT-R architecture directly. Instead, it employs a simpler version of cognitive modeling based on keystroke-level modeling (KLM) to predict user task performance. CogTool allows users to create predictive models of user interactions with interfaces by estimating the time required to complete tasks based on cognitive principles. It focuses on predicting the time it takes for an experienced user to perform tasks on an interface by simulating their actions using predefined models and parameters. By utilizing information such as mouse clicks, keyboard inputs, and other interactions, CogTool estimates task completion times without involving actual user testing. It's crucial to note that while CogTool provides quick estimations, its accuracy might vary depending on the complexity of the tasks and the interface being analyzed. Overall, CogTool is a useful tool for quickly estimating task completion times and assessing the usability of interfaces, but its predictions are based on predefined models and assumptions and may not fully capture the complexities of actual user behavior in all scenarios [12].

2.5 Results

Mobile Service Provider-1 (TTelekom)

Table 2. Mobile Servise Provider 1 Results

Participant	Task Steps					Time		
	Entering the Lira (TL) Loading Page	Choosing Lira (TL) Loading Process	Choosing a Number	Choosing Among	Making Pay	Total Time Spent(sn)	Time taken by CogTool(sn)	Difference(sn)
1	1.74	1.59	1.64	2.18	2.54	9.69	14.2	4.51
2	4.37	5.13	4.69	5.7	5.26	25.15	14.2	-10.95
3	2.36	3.72	4.07	2.46	2.91	15.52	14.2	-1.32

In Table 2, the duration of the tasks performed by the participants using the application in the number 1 mobile service provider, the duration of the CogTool and the difference are shown. While the 1st participant's time to complete all tasks was 9.69, the execution time of the Cogtool tool was 14.2. For the second participant, this period was calculated as 25.15. The 3rd participant completed the given tasks in a total of 15.52 seconds. The difference between the task durations of the 3 participants and the Cog Tool was calculated in seconds. Screenshots showing the step-by-step progress to complete the task are shown in Figure 1. A visualization of the CogTool tool's estimation time for the first mobile provider to complete the task is shown in Figure 2.

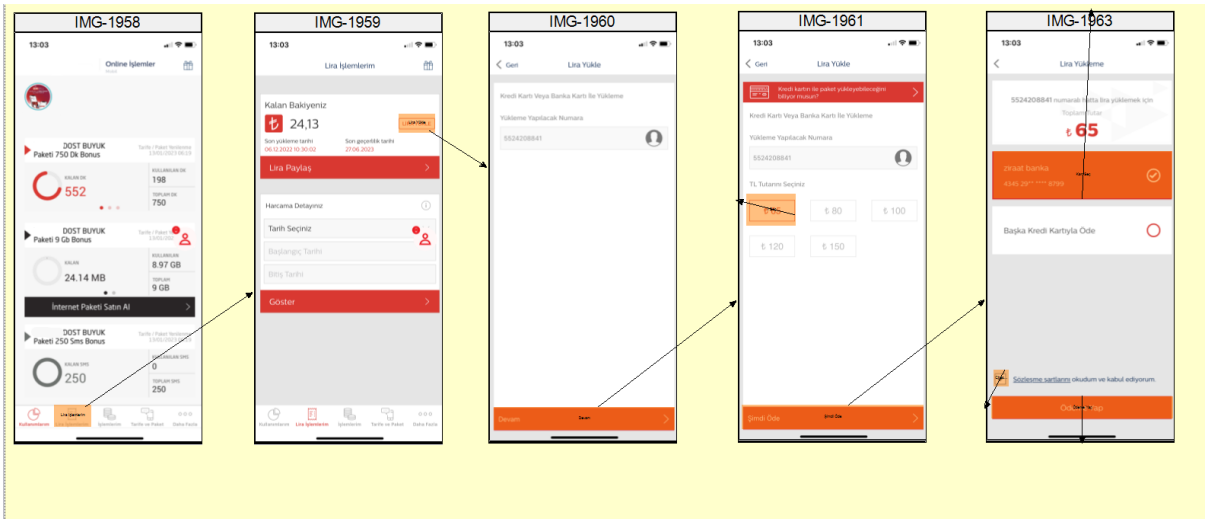

Fig 1. Participants' follow steps while performing tasks for Mobile Service Provider1

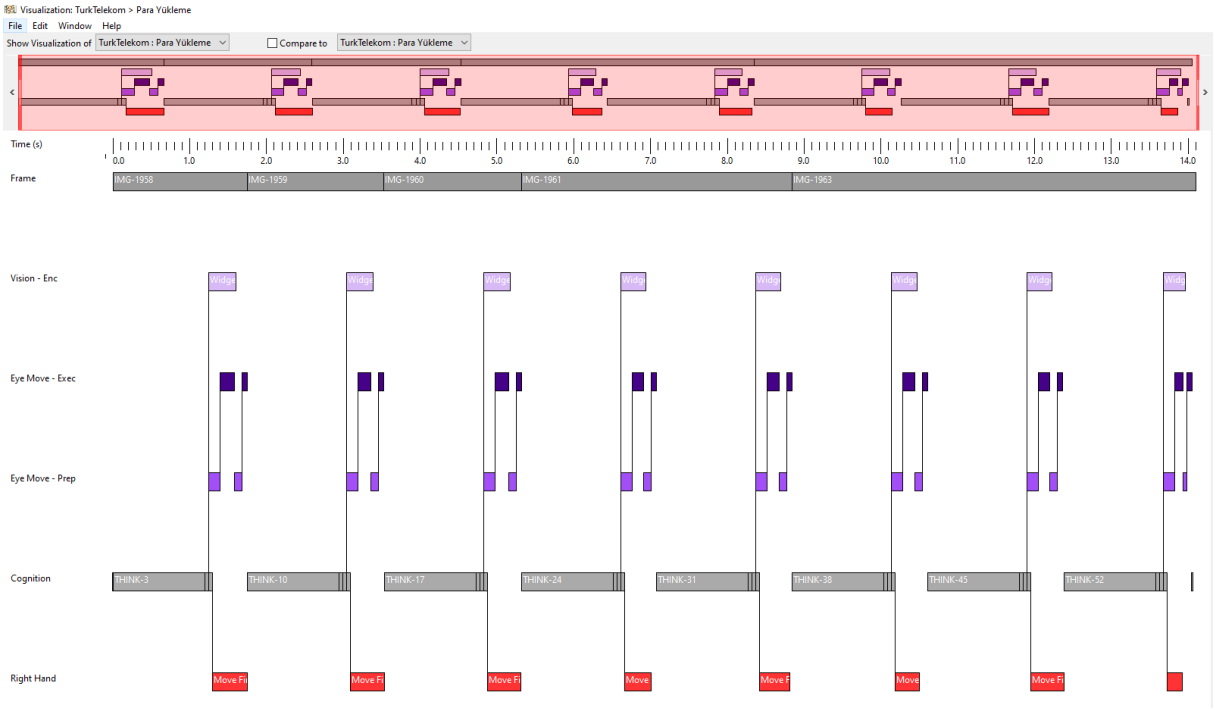

Fig 2. CogTool Tool Prediction Visualization for tasks 1nd Mobile Service Provider

Mobile Service Provider 2 (Vodafone)

Table 3. Mobile Service Provider 2 Results

Participants	Task Steps					Time		
	Entering the Lira (TL) Loading Page	Choosing Lira (TL) Loading Process	Choosing a Number	Choosing Among	Making Pay	Total Time Spent (sn)	Time taken by CogTool (sn)	Difference (sn)
1	2.54	2.2	-	1.87	3.32	9.93	7.3	-2.63
2	5.59	4.75	-	3.81	5.21	19.96	7.3	-12.06
3	1.98	2.13	-	1.84	3.06	9.01	7.3	-1.71

In Table 3, the duration of the tasks performed by the participants using the application in the number 1 mobile service provider, the duration of the CogTool and the difference are shown. While the 1st participant's time to complete all tasks was 9.93, the execution time of the CogTool tool was 7.3. For the second participant, this time was calculated as 19.96. The third participant completed the given tasks in a total of 9.01 seconds. The difference between the task durations of the 3 participants and the Cog Tool was calculated in seconds. Screenshots showing step-by-step progress to complete the task are shown in Figure 3. A visualization of the CogTool tool's estimation time for the second mobile provider to complete the task is shown in Figure 4.

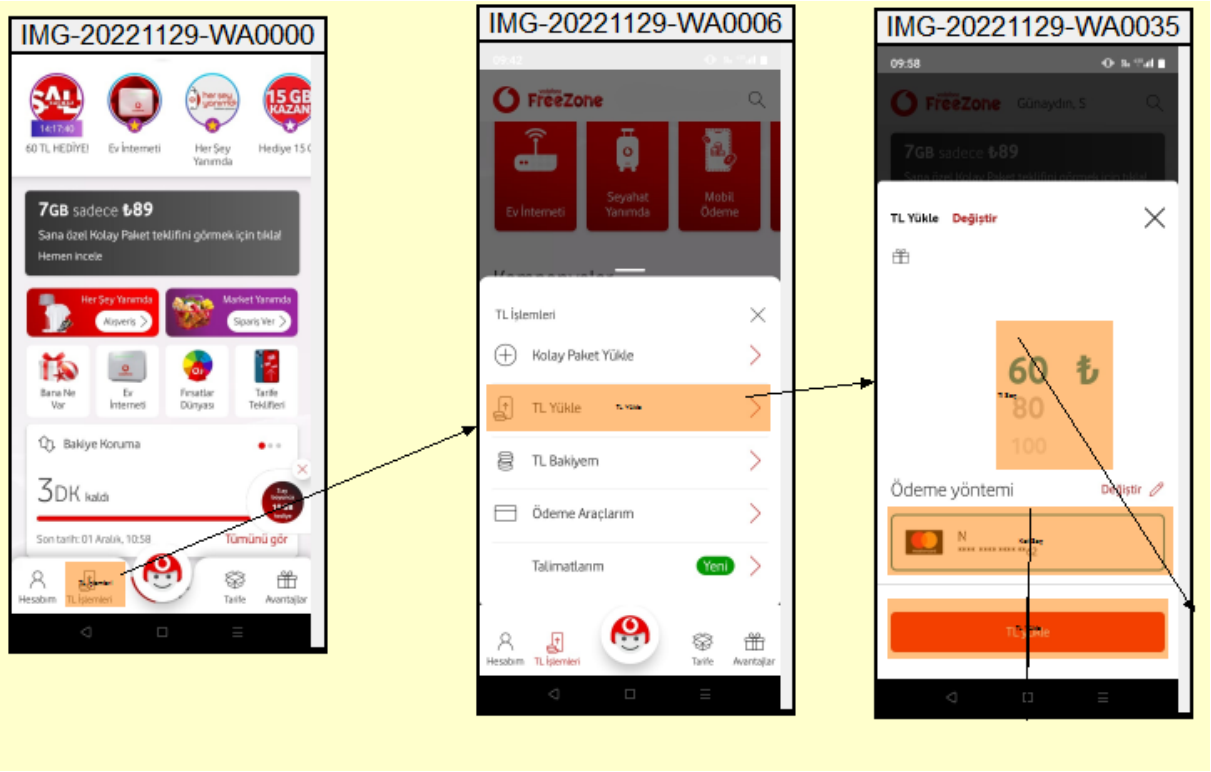

Fig 3. Participants' follow steps while performing tasks for Mobile Service Provider2

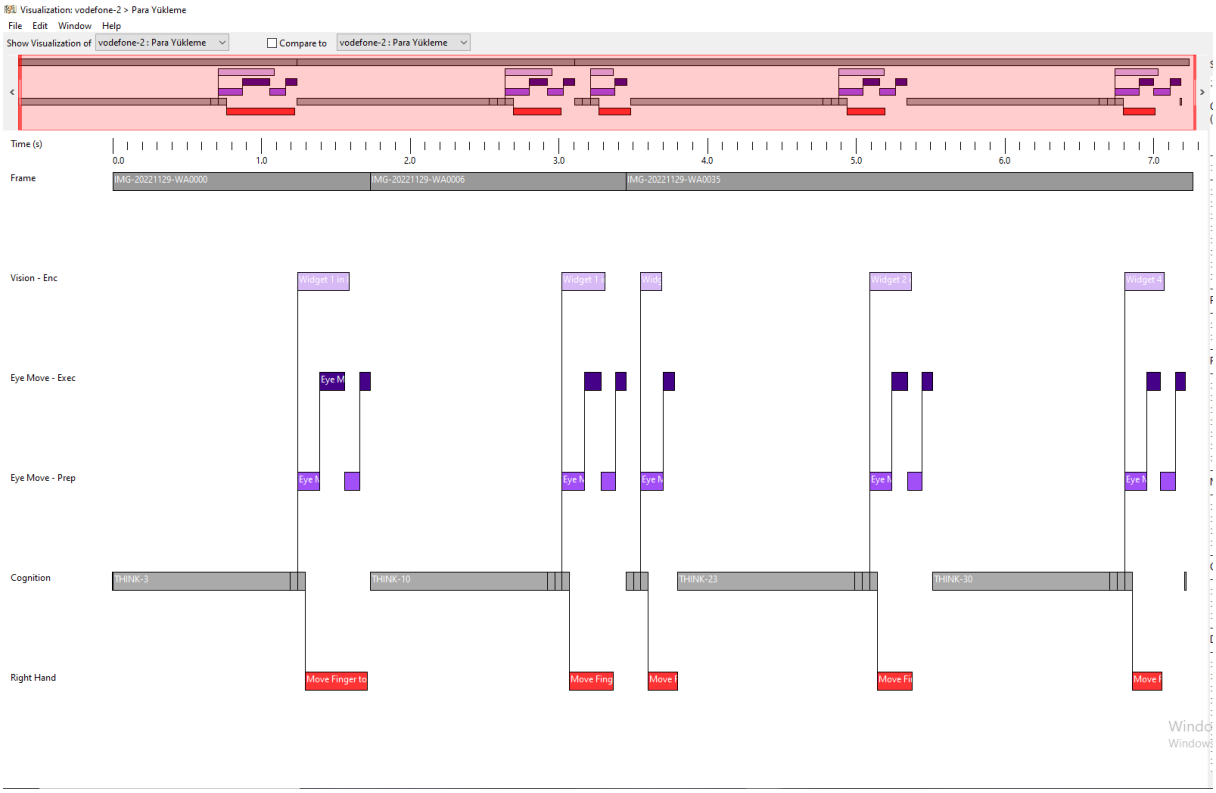

Fig 4. CogTool Tool Prediction Visualization for 2nd Mobile Service Provider

According to the results obtained, it can be said that the Vodafone service provider's mobile application is a more usable application since the transactions are performed in a shorter time after the tasks given to the users are performed via mobile applications. It is thought that the previous experiences or thinking times of the users during the tasks affected these results. It was observed that the mobile application of the Vodafone service provider gave more effective and efficient results. The tasks were completed in a shorter time, since the transactions on Vodafone were shorter and less detailed. As the number of steps required for a task in the mobile application decreases, the time will shorten, it is estimated that this situation will increase the satisfaction rate among the users. According to the satisfaction rate, companies should be more careful and selective in their interface designs and increase the usability of applications and web pages. In the light of this data, it seems that the CogTool tool can predict realistic results with a small error rate.

3.CONCLUSIONS

Tools like CogTool perform a system prediction based on cognitive modeling theories. The goal is to infer how long it will take experienced users to complete a task by estimating.

According to the literature studies, CogTool and users experienced in cognitive modeling is a tool that makes close predictions when it comes to testing interface design for usability. Very small error rate became more important because CogTool made realistic predictions in terms of improving the interface.

Recent studies in HCI have suggested that CogTool makes close estimation in real-time tasks, while in other studies it moves away from real-time tasks.

As a result of the study, if it is desired to draw a conclusion based only on "CogTool", it is seen that Vodafone mobile application is faster with a difference of "6.9" seconds compared to Türk Telekom mobile application. If we look at the overall result of the study, it is seen that Vodafone Mobile application gives faster results for all user types.

REFERENCES

[1] Card, K., et al. (1980). The keystroke-level model for user performance time with interactive systems. Communications of the ACM, 23(7), 396–410.

[2] Fiona, N., et al. (2006). Human–Computer Interaction Research In The Management Information Systems Discipline. International Journal of Human Computer Studies, 64(9), 787–788.

[3] Nielsen, J. (2012). Usability 101: Introduction to Usability. Geliş tarihi 03 Ekim 2018, gönderen https://www. nngroup.com/articles/usability-101-introduction-to-usability/

[4] Nah,F. F.-H., & Davis, S. (2002). HCI research issues in e-commerce. Journal of Electronic Commerce Research, 3(3), 16.

[5] Cagiltay, K., & ODTU, E. F. (2005). E-dönüşümü kullanabilmek? İnsan bilgisayar etkileşimi, kullanılabilirlik ve e-devlet projeleri. 2.Polis Bilişim Sempozyumu, 14-15.

[6] Carroll, M. J. (2003). HCI Models Theories, and Frameworks Toward a Multidisciplinary Science. Morgan Kaufmann Publishers, San Francisco.

[7] Taşar, O. (2020). Genetik Veri Ortamlarının Kullanılabilirlik Analizi [Usability Analysis Of Genetic Data Environments] In Sevinç Gülseçen, Kerem Rızvanoğlu, Nilgün Tosun, Emre Akadal(Ed), İnsan Bilgisayar Etkileşimi Araştırma ve Uygulamalar (pp. 9–25). İstanbul, Turkey: Istanbul Üniversitesi Yayınevi / Istanbul University Press.

[8] Jorritsma W., et al. (2015). Predicting human performance differences on multiple interface alternatives: KLM, GOMS and CogTool are unreliable. Procedia Manufacturing, 3, 3725–3731.

[9] Ludwig, Jeremy & Richards, R.. (2010). MH60S/R helicopter multi-platform & Web-based crew trainer with FLIR. 1 - 7. 10.1109/AERO.2010.5446775.

[10] Demirci H. M, (2021) Online Shopping Web Sites' Perceived Usability: A Case Study with Turkish Shopping Related Web Sites. In Advances in Usability, User Experience, Wearable and Assistive Technology: Proceedings of the AHFE 2021 Virtual Conferences on Usability and User Experience, Human Factors and Wearable Technologies, Human Factors in Virtual Environments and Game Design, and Human Factors and Assistive Technology, July 25-29, 2021, USA (pp. 759-766). Springer International Publishing.

[11] Yamira Santiago-Espada and Kara A. Latorella "Predicting Operator Execution Times Using Cogtool", 17th International Symposium on Aviation Psychology, May 6, 2013.

[12] Arik, K. (2019). A usability Research on Task Prediction For Experienced Users with Cogtool, The Journal of Cognitive Systems, 4 (2), 46-53.