International Journal of Multidisciplinary Studies and Innovative Technologies

Volume : 5
Number: 1
Year: 2021
Pages: 36 - 45

EVALUATION OF JEPPESEN AND GARMIN ELECTRONIC FLIGHT BAGS (EFBs) APPLICATIONS IN TERMS OF COGNITIVE WORKLOAD AND AVAILABILITY

N.Firat Ozkan[1*], Emre Aksoy [2] and Gokberk Sensoy [3]

[1]*Gradute School of natural and Applied Science, Osmangazi University, Eskisehir, Turkey*
[2] *Gradute School of natural and Applied Science, Osmangazi University, Eskisehir, Turkey*
[3] *Gradute School of natural and Applied Science, Osmangazi University, Eskisehir, Turkey*

Abstract – As the dimensions and capabilities of tablet computers are suitable for use in aircraft, software has begun to be developed. The approach plans and route information used on printed paper about the airports in flight operations have been transformed into a paperless flight deck concept [1] by means of the development of appropriate software. With the introduction of the paperless cockpit concept, pilots can access the information they need through the tablet and the applications installed in it. Together with the paperless cockpit concept, the physical workload of the pilots on the physical document handling and document up-to-date works has virtually disappeared. However, the disappearance of the physical workload was replaced by the cognitive workload. To use the capabilities of EFBs applications practically, it is important to develop applications to minimize the cognitive workload. Cognitive load analysis and usability studies were conducted by using NASA-TLX and IBM CSUQ questionnaires between JEPPESEN FD PRO application used by participating pilots and GARMIN PILOT application which pilots had not used before. User tests were performed on two different groups. The first group consists of 15 pilot working on light-middle class transport aircraft. The second group consists of 15 flight technicians working in light-middle class transport aircraft. NASA-TLX and IBM CSUQ questionnaires were applied to both groups and the results were analyzed in SPSS program.

Keywords – User test, Cognitive workload, EFBs (Electronic Flight Bags), NASA-TLX, IBM CSUQ

I. INTRODUCTION

Workload is a term that represents the cost of the user to perform tasks.[2] The user can be successful according to his / her ability to use the device or tool he / she has used to perform the given tasks. In this study, the contribution of Electronic Flight Bags applications to cognitive workload, which considerably reduces the workload of pilots, was evaluated. Pilots must know a lot of documents and information according to the content of the flight. Some of these documents must be reachable physically during the flights. These physically mandatory documents mainly include aircraft operating manuals and emergency checklists, performance, and weight-balance guidelines. In addition, there are flight documents that vary according to the route to be flown and the airports to be landed. These flight documents also include airline maps, descent plans for landing and take-off areas, as well as descent plans for spare and emergency descent areas. All of this results in lots of folders that pilots should be able to take with them and use. As a result of the development of tablet technologies and their use by civil aviation authorities, Electronic Flight Bags (EFBs) have become available for flight operations. The companies that produce airway maps and descent ones as printed paper are now switched to the paperless cockpit concept with the software they have designed. With the transition to this concept, the main question is which human factors bring about.[3] In the traditional cockpit environment, pilots provided the information they needed on paper.

Today, these information's have been completely transferred to the electronic environment. To achieve these, tablets which are embedded or portable in aircraft systems have begun to be used. Thanks to these systems, the physical workload factor on pilots and flight crews is considerably reduced. Pilots prefer to fly with electronic flight bags instead of printed documents.

A. EFBs (Electronic Flight Bags)

Electronic flight bags are more cost-effective, safer, and more user-friendly.[4] Documents used in aviation are called living documents. The documents used in accordance with new developments and rules are constantly updated. The pilots must make sure that all changes to the documents they will use are processed before going to the flight. Due to the large number of documents, in some cases it may be necessary to hire a staff member for updating. In addition, updating activity, the pilots operate in a physically narrow space during flight. Accessing all documents quickly is a separate challenge.

B. Skill Reduction

Electronic flight bags have negative effects as well as the positive effects on pilots' workload. One of them is the decrease in skill. Pilots can perform their calculations or reach the documents by means of electronic flight bags very easily and quickly. However, in the event of a failure or inaccessibility in electronic flight bags, it is observed that their skills are reduced.[5] This indicates that there should be a procedure about electronic flight bags failure in the pilot emergency checklists. If Electronic Flight Bags are used in a portable manner, they should be considered as part of the aircraft systems and included in the system failure trainings given in the simulators.

C. EFBs Caused Aircraft Incidents

No matter how perfectly designed a new system used in the cockpit or on the aircraft, error factors will come into play because the user of this new system is still human. Due to the errors caused by human factors, wrong data are entered into the application and flights are performed with incorrect information. According to the EASA (European Aviation Safety Agency) research project [6], aircraft incidents occur as a result of pilots entering incorrect information in their electronic flight bags or using them incorrectly.

The occurrence of aircraft incidents due to incorrect calculations by the pilots is also the case before the electronic flight bags. Because in both cases, the human factor exists as a factor. Pilots should be aware that data entry errors can be made even if they use electronic flight bags.

It is not possible for printed documents to assume the function of controlling human errors. This deficiency may be prevented by covering Electronic Flight Bags in software designs. Potential mistakes can be avoided by offering suggestions or reminders to the pilot.

II. MATERIALS AND METHOD

The participants were informed about the tasks to be performed before the test and the test scenario.[7] Using the JEPPESEN FD PRO and GARMIN PILOT applications, participants were asked to complete the tasks assigned to them and then complete the NASA-TLX and IBM CSUQ questionnaires. After the tests were completed with each participant, general evaluations about the practices, characteristics that they were satisfied or disturbed were recorded.

A. Participants

User tests were performed on two different groups. The first group consists of 15 pilot working on light-middle class transport aircraft. The second group consists of 15 flight technicians working in light-middle class transport aircraft.

The user group of pilots have previously used JEPPESEN FD PRO. However, they have never used GARMIN PILOT. Flight technicians have never used both applications before.

B. Apparatus

IPad mini 4 tablet was used in the user test. By installing the applications on the same tablet, it is provided to perform the specified tasks. 5 (five) tasks are defined for each user. Training videos of both applications were made to help users understand how to do the tasks.

Level-2 and level-3 headings can be used to detail main headings.

C. Experimental Design

User tests were conducted with two different groups. The first group consists of 15 pilots of light-middle class transport aircraft. The second group consists of 15 participants who are working as flight technicians in light-middle class transport aircraft. At the beginning of the test, the test scenario was explained to all participants and they were asked to perform the tasks shown in “JEPPESEN – GARMIN EFB USER TEST” via JEPPESEN FD PRO and GARMIN PILOT applications.

JEPPESEN GARMIN EFB USER TEST TASKS

1. After watching the application videos.

- **1.1.** Plan **“LTBI – TOKER G8 SRT – LTCL”** route with using JEPPESEN ve GARMIN applications.
- **1.2.** Find the **total flight distance** by using the apps and write.
- **1.3.** Find Eskişehir ve Siirt Airports **tower frequencies** by using the apps and write.
- **1.4.** Find Eskişehir ve Siirt Airports **runway elevations** by using the apps and write.
- **1.5. DIRECT-TO “HAY”** point.

2. IBM and NASA questionnaires that will be given to you after you finish each application.

3. Which one would you choose between JEPPESEN ve GARMIN? Please explain shortly.

After completing the tasks, NASA-TLX and IBM CSUQ questionnaires were completed. Completion times and number of errors are recorded for each application. Finally, the participants' opinions were taken on the most important stage of the two applications or what might be the most important feature that distinguishes between the two applications.

Our first goal when performing the user test is to determine which application creates less cognitive load than the other. We conducted these assessments within and between groups separately. The hypothesis we put forward for usability testing is as follows:

Ho: Considering the variables of NASA-TLX, IBM CSUQ, completion times and number of errors, the application that creates less cognitive workload will be preferred by the participants.

III. STATISTICAL ANALYSIS

When exploring the statistical data, the defined variables are as follows:

PARTICIPANT TRAINING LEVEL: Pilot / Technician status of the participants

NUMBER OF JEPPESEN ERRORS: Number of errors when the Jeppesen task is complete (0 … 4)

JEPPESEN DUTY PERIOD: Jeppesen task completion time (hh:mm:sn)

NUMBER OF GARMIN ERRORS: Number of errors when the Garmin task is completed. (0 … 4)

GARMIN DUTY PERIOD: Garmin task completion time (hh:mm:sn)

JEPPESEN NASA WEIGHT SCORE: Weighted score with NASA TLX questionnaire after Jeppesen task completion. (0 … 120)

GARMIN NASA WEIGHT SCORE: Weighted score with NASA TLX questionnaire after Garmin task completion. (0 … 120)

JEPPESEN IBM OVE: Average of answer scores (0… 9) for all questions (19), which was the first of the variables reduced to four by the IBM CSUQ questionnaire after Jeppesen task was completed.

JEPPESEN IBM SYS: Average of the answer points (0… 9) of the first eight questions, which is the second of the variables reduced to four by the IBM CSUQ questionnaire after Jeppesen task is completed.

JEPPESEN IBM INF: The average of answer points (0… 9) of questions from 9 to 15, which is the third variable, which is reduced to four by the IBM CSUQ questionnaire after Jeppesen task is completed.

JEPPESEN IBM INTER: After Jeppesen task is completed, the average of answer points (0… 9) of questions from 16 to 18, which is the last variable reduced to four by the IBM CSUQ questionnaire.

GARMIN IBM OVE: Average of answer scores (0… 9) for all questions (19), which is the first of the variables reduced to four by the IBM CSUQ questionnaire after the Garmin task is completed.

GARMIN IBM SYS: Average of the answer scores (0… 9) of the first eight questions, which is the second of the variables reduced to four by the IBM CSUQ questionnaire after the Garmin task is completed.

GARMIN IBM INF: After the Garmin task is completed, the average of the answer points (0… 9) of the questions from 9 to 15, which is the third variable, which is reduced to four by the IBM CSUQ questionnaire.

GARMIN IBM INTER: After the Garmin task is completed, the average of the answer points (0… 9) of the questions from 16 to 18, which is the last variable reduced to four by the IBM CSUQ survey

.

A. Frequencies and Normality Tests

The frequencies of the mentioned variables are given in Table 1.

Table 1. Frequency Table of Variables

		Total Number Of Participants	Participant Training Level	Number of Jeppesen Errors	Jeppesen Duty Perıod	Number of Garmin Errors	Garmin Duty Period	Jeppesen Nasa Weight Score	Garmin Nasa Weight Score
N	**Valid**	30	30	30	30	30	30	30	30
	Missing	0	0	0	0	0	0	0	0
Mean		1,50	1,50	1,37	0:02:24	1,03	0:01:54	26,1090	19,4273
Std. Error of Mean		,093	,093	,176	0:00:07	,148	0:00:05	3,51188	2,55084
Median		1,50	1,50	1,00	0:02:25	1,00	0:01:59	14,5150	11,8500
Mode		1[a]	1[a]	1	0:02:48[a]	1	0:02:20	38,20	9,60
Std. Deviation		,509	,509	,964	0:00:41	,809	0:00:31	19,23535	13,97151
Variance		,259	,259	,930	1688,809	,654	962,616	369,999	195,203
Skewness		0,000	0,000	,159	-,248	,356	-,321	1,032	1,044
Std. Error of Skewness		,427	,427	,427	,427	,427	,427	,427	,427
Kurtosis		-2,148	-2,148	-,833	-1,341	-,343	-1,515	-,189	-,699
Std. Error of Kurtosis		,833	,833	,833	,833	,833	,833	,833	,833
Range		1	1	3	0:02:06	3	0:01:37	60,45	39,05
Sum		45	45	41	1:12:04	31	0:57:22	783,27	582,82

Table 1. Frequency Table of Variables (Cont.)

		JEPPESEN IBM OVE	JEPPESEN IBM SYS	JEPPESEN IBM INF	JEPPESEN IBM INTER	GARMIN IBM OVE	GARMIN IBM SYS	GARMIN IBM INF	GARMIN IBM INTER
N	**Valid**	30	30	30	30	30	30	30	30
	Missing	0	0	0	0	0	0	0	0
Mean		6,4300	6,0980	6,3770	6,6890	7,4983	7,2817	7,4290	7,9783
Std. Error of Mean		,21345	,27171	,21370	,19181	,15225	,20649	,14905	,13753
Median		6,1250	5,7500	6,0000	6,3300	7,3750	7,2500	7,2900	7,8350
Mode		5,35a	4,75a	6,00	6,00	7,10a	8,63	7,29	7,67
Std. Deviation		1,16911	1,48822	1,17049	1,05061	,83392	1,13097	,81639	,75327
Variance		1,367	2,215	1,370	1,104	,695	1,279	,666	,567
Skewness		,721	,397	,943	,737	,093	-,140	,256	-,126
Std. Error of Skewness		,427	,427	,427	,427	,427	,427	,427	,427
Kurtosis		-,599	-,146	-,071	-,346	-1,117	-,510	-,706	-,749
Std. Error of Kurtosis		,833	,833	,833	,833	,833	,833	,833	,833
Range		3,95	6,00	4,14	4,00	2,55	4,37	3,00	2,67
Sum		192,90	182,94	191,31	200,67	224,95	218,45	222,87	239,35

We use skewness and kurtosis as rough indicators of the degree of normality of distributions or the lack thereof. Unlike test statistics from normality testing procedures like the Kolmogorov–Smirnov or the Shapiro–Wilk, skewness and kurrtosis are used here like an effect size, to communicate the degree of nonnormality, rather than statistical significance under some null hypothesis of normality. The use of skewness and kurtosis to describe distributions dates to Pearson (1895) and has been reviewed more recently by Moors (1986), D'Agostino, Belanger, and D'Agostino (1990), and DeCarlo (1997). Skewness is a rough index of the asymmetry of a distribution, where positive skewness in unimodal distributions suggests relatively plentiful and/or extreme positive values, and negative skewness suggests the same for negative values. Skewness is an estimate of the third standardized moment of the population distribution.[8]

Skewness can range from -∞ to +∞ and symmetric distributions like the normal distribution have a skewness of 0. The n-based bias correction term for the population estimate is negligible for the large samples that we have here, but we include it for completeness. Kurtosis is an estimate of the fourth standardized moment of the population distribution.[8]

Kurtosis can range from 1 to +N. The kurtosis of a normal distribution is 3. Although it is common to subtract 3 from k and describe this as "excess kurtosis" —beyond that expected from a normal distribution— we use the definition above, where k\3 is "platykurtic (less peakedness, weaker "tails", heavy "shoulders") and k.3 is leptokurtic (more peakedness, heavy tails, weak shoulders). As references, a uniform distribution has a kurtosis of 1.8 (platykurtic), and a logistic distribution has a kurtosis of 4.2 (leptokurtic).

As the standard errors get smaller when the sample size increases, z-tests under null hypothesis of normal distribution tend to be easily rejected in large samples with distribution which may not substantially differ from normality, while in small samples null hypothesis of normality tends to be more easily accepted than necessary. Therefore, critical values for rejecting the null hypothesis need to be different according to the sample size as follow one: For small samples ($n < 50$), if absolute z-scores for either skewness or kurtosis are larger than 1.96, which corresponds with a alpha level 0.05, then reject the null hypothesis and conclude the distribution of the sample is non-normal.[9]

Table 1 shows that the skewness value in the 95% confidence interval is less than 3 because it is platycurtic (fewer peaks, weaker tail, and stronger shoulders), and is close to normal distribution because the kurtosis values are less than 1.96.

B. Independent Sample T-Test

JEPPESEN FD PRO and GARMIN PILOT Tests for pilot and flight technicians; Completion time, number of errors The Independent Sample T-Test Group Statistics applied to the weighted data of NASA TLX and IBM CSUQ survey results are given in Table 2.

Table 2. Independent Sample Participant Type T-Test Group Statistics

TOTAL NUMBER OF PARTICIPANTS		**N**	**Mean**	**Std. Deviation**	**Std. Error Mean**
NUMBER OF JEPPESEN ERRORS	PILOT	15	1,07	,961	,248
	TECHNICIANS	15	1,67	,900	,232
NUMBER OF GARMIN ERRORS	PILOT	15	,67	,724	,187
	TECHNICIANS	15	1,40	,737	,190
JEPPESEN DUTY PERIOD	PILOT	15	0:01:46	0:00:19	0:00:05
	TECHNICIANS	15	0:03:01	0:00:11	0:00:03
GARMIN DUTY PERIOD	PILOT	15	0:01:26	0:00:16	0:00:04
	TECHNICIANS	15	0:02:22	0:00:06	0:00:01
JEPPESEN NASA WEIGHT SCORE	PILOT	15	39,8733	18,90665	4,88168
	TECHNICIANS	15	12,3447	1,72795	,44615
GARMIN NASA WEIGHT SCORE	PILOT	15	28,4333	15,12599	3,90551
	TECHNICIANS	15	10,4213	1,32318	,34164
JEPPESEN IBM OVE	PILOT	15	7,2233	1,09689	,28322
	TECHNICIANS	15	5,6367	,52863	,13649
GARMIN IBM OVE	PILOT	15	8,1267	,59458	,15352
	TECHNICIANS	15	6,8700	,49092	,12675
JEPPESEN IBM SYS	PILOT	15	7,1440	1,24756	,32212
	TECHNICIANS	15	5,0520	,82876	,21399
GARMIN IBM SYS	PILOT	15	8,1860	,65785	,16986
	TECHNICIANS	15	6,3773	,68139	,17593
JEPPESEN IBM INF	PILOT	15	7,1240	1,19177	,30771
	TECHNICIANS	15	5,6300	,47103	,12162
GARMIN IBM INF	PILOT	15	7,9727	,69107	,17843
	TECHNICIANS	15	6,8853	,51928	,13408
JEPPESEN IBM INTER	PILOT	15	7,3567	1,07211	,27682
	TECHNICIANS	15	6,0213	,42631	,11007
GARMIN IBM INTER	PILOT	15	8,3567	,65965	,17032
	TECHNICIANS	15	7,6000	,65841	,17000

Here is a summary of the Independent Sample T-Test data we compared pilot and flight technicians:

a) Considering the number of errors.

For the JEPPESEN FD PRO test, it was found that flight technicians made more mistakes. Although the flight technicians made more mistakes for the GARMIN PILOT Test, both groups of participants made fewer errors in the GARMIN PILOT test than the JEPPESEN FD PRO test.

b) When the test completion times are observed.

It has been observed that the duration of flight technicians who are not accustomed to using the JEPPESEN FD PRO application is almost twice that of the pilots who are relatively familiar with the application.

However, for both pilots and flight technicians who are not familiar with GARMIN PILOT, the completion times are shorter than the JEPPESEN FD PRO completion times. This difference is remarkable with flight technicians watching for around a minute.

c) Considering the NASA TLX weighted averages.

For both JEPPESEN FD PRO and GARMIN PILOT applications, the weighted averages of pilots are strongly negatively differentiated from that of flight technicians.

d) Considering the IBM CSUQ weighted averages.

According to the flight technicians, the pilot scores were high for both tests.

Independent sample T-Test applied to the weighted data of NASA TLX and IBM CSUQ survey results of JEPPESEN FD PRO and GARMIN PILOT Tests for pilot and flight technicians is given in Table 3.

Table 3. Independent Sample T-Test Group Statistics

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
NUMBER OF JEPPESEN ERRORS	Equal variances assumed	,000	1,000	-1,765	28	,088	-,600	,340	-1,296	,096
	Equal variances not assumed			-1,765	27,879	,089	-,600	,340	-1,296	,096
NUMBER OF JEPPESEN ERRORS	Equal variances assumed	,005	,945	-2,750	28	,010	-,733	,267	-1,280	-,187
	Equal variances not assumed			-2,750	27,991	,010	-,733	,267	-1,280	-,187
JEPPESEN DUTY PERIOD	Equal variances assumed	1,133	,296	-12,632	28	,000	-0:01:14	0:00:05	-0:01:26	-0:01:02
	Equal variances not assumed			-12,632	23,060	,000	-0:01:14	0:00:05	-0:01:26	-0:01:02
JEPPESEN DUTY PERIOD	Equal variances assumed	7,901	,009	-12,240	28	,000	-0:00:55	0:00:04	-0:01:05	-0:00:46
	Equal variances not assumed			-12,240	18,357	,000	-0:00:55	0:00:04	-0:01:05	-0:00:46
JEPPESEN NASA WEIGHT SCORE	Equal variances assumed	18,078	,000	5,616	28	,000	27,52867	4,90202	17,48733	37,57000
	Equal variances not assumed			5,616	14,234	,000	27,52867	4,90202	17,03106	38,02627
GARMIN NASA WEIGHT SCORE	Equal variances assumed	77,325	,000	4,594	28	,000	18,01200	3,92043	9,98137	26,04263
	Equal variances not assumed			4,594	14,214	,000	18,01200	3,92043	9,61539	26,40861

Table 3. Independent Sample T-Test Group Statistics (Cont.)

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
JEPPESEN IBM OVE	Equal variances assumed	13,178	,001	5,047	28	,000	1,58667	,31439	,94267	2,23067
	Equal variances not assumed			5,047	20,171	,000	1,58667	,31439	,93122	2,24212
GARMIN IBM OVE	Equal variances assumed	2,162	,153	6,312	28	,000	1,25667	,19909	,84886	1,66447
	Equal variances not assumed			6,312	27,032	,000	1,25667	,19909	,84820	1,66513
JEPPESEN IBM SYS	Equal variances assumed	4,141	,051	5,410	28	,000	2,09200	,38672	1,29984	2,88416
	Equal variances not assumed			5,410	24,342	,000	2,09200	,38672	1,29445	2,88955
GARMIN IBM SYS	Equal variances assumed	,925	,344	7,396	28	,000	1,80867	,24455	1,30773	2,30960
	Equal variances not assumed			7,396	27,965	,000	1,80867	,24455	1,30770	2,30963
JEPPESEN IBM INF	Equal variances assumed	14,315	,001	4,515	28	,000	1,49400	,33088	,81623	2,17177
	Equal variances not assumed			4,515	18,270	,000	1,49400	,33088	,79959	2,18841
GARMIN IBM INF	Equal variances assumed	3,473	,073	4,872	28	,000	1,08733	,22319	,63014	1,54453
	Equal variances not assumed			4,872	25,988	,000	1,08733	,22319	,62854	1,54613
JEPPESEN IBM INTER	Equal variances assumed	11,538	,002	4,482	28	,000	1,33533	,29790	,72511	1,94555
	Equal variances not assumed			4,482	18,319	,000	1,33533	,29790	,71025	1,96042
GARMIN IBM INTER	Equal variances assumed	,495	,488	3,144	28	,004	,75667	,24064	,26373	1,24960
	Equal variances not assumed			3,144	28,000	,004	,75667	,24064	,26373	1,24960

C. Paired Sample T-Test

Paired Sample Statistics are applied to all participants and applied to the weighted data of JEPPESEN FD PRO and GARMIN PILOT Tests, completion time, number of errors and NASA TLX and IBM CSUQ survey results.

Table 4. Paired Sample Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	**NUMBER OF JEPPESEN ERRORS**	1,37	30	,964	,176
	NUMBER OF GARMIN ERRORS	1,03	30	,809	,148
Pair 2	**JEPPESEN DUTY PERIOD**	0:02:24	30	0:00:41	0:00:07
	GARMIN DUTY PERIOD	0:01:54	30	0:00:31	0:00:05
Pair 3	**JEPPESEN NASA WEIGHT SCORE**	26,1090	30	19,23535	3,51188
	JEPPESEN NASA WEIGHT SCORE	19,4273	30	13,97151	2,55084
Pair 4	**JEPPESEN IBM OVE**	6,4300	30	1,16911	,21345
	GARMIN IBM OVE	7,4983	30	,83392	,15225
Pair 5	**JEPPESEN IBM SYS**	6,0980	30	1,48822	,27171
	GARMIN IBM SYS	7,2817	30	1,13097	,20649
Pair 6	**JEPPESEN IBM INF**	6,3770	30	1,17049	,21370
	GARMIN IBM INF	7,4290	30	,81639	,14905
Pair 7	**JEPPESEN IBM INTER**	6,6890	30	1,05061	,19181
	JEPPESEN IBM INTER	7,9783	30	,75327	,13753

The short and concise expressions we extracted using the JEPPESEN FD PRO and GARMIN PILOT data from the Paired Sample T-Test data are as follows:

a) Considering the number of errors.

More mistakes were made in JEPPESEN FD PRO than GARMIN Pilot. The standard deviation and the standard error mean are wider.

b) When the task completion periods are taken into consideration.

According to JEPPESEN FD PRO, GARMIN PILOT is applied for 30 seconds. completed in less time.

c)According to the NASA TLX survey weighted average, which is an indicator of cognitive workload.

Regarding to GARMIN Pilot, JEPPESEN FD PRO application caused the participants to load cognitive work with a difference of about 7 points.

d)Availability satisfaction indicator as for the weighted average data of the IBM CSUQ survey.

GARMIN PILOT application is superior to JEPPESEN FD PRO in all four titles.

Table 5 shows the time of completion, number of errors and weighted data of the NASA TLX and IBM CSUQ survey results of the JEPPESEN FD PRO and GARMIN PILOT Tests applied to all participants.

Table 5. Paired Sample T-Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Conf.Interval of the Difference Lower	Upper			
Pair 1	**NUMBER OF JEPPESEN ERRORS NUMBER OF GARMIN ERRORS**	,333	,479	,088	,154	,512	3,808	29	,001
Pair 2	**JEPPESEN DUTY PERIOD GARMIN DUTY PERIOD**	0:00:29	0:00:19	0:00:03	0:00:22	0:00:36	8,278	29	,000
Pair 3	**JEPPESEN NASA WEIGHT SCORE GARMIN NASA WEIGHT SCORE**	6,68167	9,50546	1,73545	3,13227	10,23106	3,850	29	,001
Pair 4	**JEPPESEN IBM OVE GARMIN IBM OVE**	-1,06833	,83175	,15186	-1,37891	-,75775	-7,035	29	,000
Pair 5	**JEPPESEN IBM SYS GARMIN IBM SYS**	-1,18367	1,03667	,18927	-1,57077	-,79657	-6,254	29	,000
Pair 6	**JEPPESEN IBM INF GARMIN IBM INF**	-1,05200	,86699	,15829	-1,37574	-,72826	-6,646	29	,000
Pair 7	**JEPPESEN IBM INTER GARMIN IBM INTER**	-1,28933	,90401	,16505	-1,62690	-,95177	-7,812	29	,000

IV. RESULTS

After the tests with the users, opinions were taken about JEPPESEN FD PRO and GARMIN PILOT applications.

When the opinions of the pilot group of participants were evaluated, almost all of them found GARMIN PILOT application convenient and easy. The recommendations made by the GARMIN PILOT program were especially appreciated by some participants when creating the route. All participants agreed that the GARMIN PILOT application enables faster data entry. However, most of the participants stated that they would choose JEPPESEN FD PRO when it came to preference. When asked about the reasons for this, they stated that it is easier to use GARMIN PILOT than JEPPESEN FD PRO, but the current difficulty level of JEPPESEN FD PRO does not disturb them. It is considered that the familiarity of the participating group of pilots with the brand is also effective in the selection of JEPPESEN FD PRO.

When the opinions of the group of flight technicians were evaluated, they concluded that it was easier to use the GARMIN PILOT application. They stated that they could prefer GARMIN PILOT because they did not use both applications unlike the pilots.

V. DISCUSSIONS

In this study, it is aimed to analyze the GARMIN PILOT application in terms of cognitive workload and determine the ease of use with JEPPESEN FD PRO application which is designed for basically the same purposes. In addition to being familiar with the JEPPESEN FD PRO application, the use of printed paper versions in the past was a major factor in their unwillingness to make changes. In the evaluations of JEPPESEN FD PRO in terms of cognitive workload factors, the group of participants from the pilots used their preferences as JEPPESEN FD PRO, although it was behind the GARMIN PILOT application. For this reason, H0 hypothesis could not be confirmed for the pilot group of participants.

The participant group of flight technicians made assessments based on the current capabilities of both applications. The fact that they had not used both applications before shaped the preference processes according to cognitive workload factors. Participant group of flight technicians stated that they could prefer GARMIN PILOT application because of ease of use and lower cognitive workload values. For this reason, H_0 hypothesis is validated for a group of flight technicians.

It was determined that pilots performed better in the number of errors for both tests. While the pilots expect the number of errors made in the JEPPESEN FD PRO application to be less, the fewer mistakes the participants make in the GARMIN PILOT application show that the GARMIN PILOT causes less error.

When we examine the observed task completion times for both applications considering the experience, The GARMIN PILOT application has shorter task completion times. This information indicates that the GARMIN PILOT is more practical.

According to NASA TLX survey data, which is an indicator of cognitive workload, pilots were exposed to more cognitive workload in both applications than flight technicians. The possible reasons for this are that the feeling of trust that the pilots will be able to open will reduce the relationship to the task.

VI. CONCLUSION

JEPPESEN FD PRO and GARMIN PILOT applications are sufficient to perform the given tasks. Can be considered as interchangeable applications. What makes the distinction between applications is the levels of cognitive workload on users. As a result of user tests, the participant group consisting of pilots, it was seen that the cognitive workload provided by GARMIN PILOT did not cause any change in preference. However, flight technicians having the same experience in both applications, determined their preferences according to cognitive workload factors.

REFERENCES

[1.] ICAO, *Doc 10020 Manual of Electronic Flight Bags (EFBs).* 2018.

[2.] Hart, S.G., *Nasa-Task Load Index (Nasa-TLX); 20 Years Later.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2006. 50(9): p. 904-908.

[3.] Divya C. Chandra, M.Y., Vic Riley, Susan J. Mangold, *Human Factors Considerations in the Design and Evaluation of Electronic Flight Bags (EFBs)*. 2003.https://rosap.ntl.bts.gov/view/dot/8815

[4.] Winter, S.R., et al., *Pilot performance comparison between electronic and paper instrument approach charts.* Safety science, 2018. 103: p. 280-286.

[5.] Milner, M.N., et al., *The Effect of Chart Type on Pilots' Response Time.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2017. 61(1): p. 1365-1368.

[6.] EASA, *Research Project: EFB Electronic Flight Bag Aircraft performance calculations and mass & balance Best practices for evaluation and use of EFB.* 2015.

[7.] Krug, S., *Rocket Surgery Made Easy: The Do-It-Yourself Guide to Finding and Fixing Usability Problems.* Berkeley, CA: New Riders. pp. 147-152, 2010.

[8.] Ho, A.D. and C.C. Yu, *Descriptive Statistics for Modern Test Score Distributions: Skewness, Kurtosis, Discreteness, and Ceiling Effects.* Educational and psychological measurement, 2015. 75(3): p. 365-388.

[9.] Kim, H.-Y., *Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis.* Restorative dentistry & endodontics, 2013. 38(1): p. 52.