

A User Experience Study of Airline Websites

Mahmut Ekşioğlu^{1,*}, Esin Kırış², Tuğba Çakır¹,
Merve Güvendik¹, Efsane D. Koyutürk¹, and Merve Yılmaz¹

¹ Department of Industrial Engineering, Boğaziçi University, Istanbul-Turkey
mahmut.eksioglu@boun.edu.tr

² Syncsort Incorporated Woodcliff Lake, NJ 07677 USA

Abstract. The aim of this study was to identify and compare user experience issues and provide recommendations for the three airline websites. The websites were evaluated through heuristic evaluation and usability testing methods. Based on the heuristic evaluation findings, two relatively difficult task scenarios were generated for usability testing: A total of 168 student users participated in the study. The participants' performances for user testing were evaluated in terms of task completion time, success rate, and number of page views by a remote usability testing tool, Loop11. Additionally, a post-test questionnaire was administered online. Statistical analysis results indicated a significant 'website-design' effect on task completion times. Overall, the task completion time significantly varied among the three airlines. Notably, 'gender', 'education level' and 'internet use frequency' had no significant effect on task completion times for both tasks. Re-design recommendations are provided for a sample of cases.

Keywords: Heuristic Evaluation, Usability and User Experience, Web Usability, Airline Companies.

1 Introduction

Improving the design of an airline website in terms of user experience builds a competitive advantage by satisfying and improving the customer experience in flight related matters. Even though website design is not a very complex task that anyone with a basic know-how of computer can carry out, creating a good web site in terms of usability, user experience, and HCI principles is a rather difficult task. Due to a number of deficiencies in the website designs, customers are facing problems and are either unnecessarily spending so much of their valuable time, or giving up the use of the sites. The common website design deficiencies as identified by Nielsen (2005) [1] are still prevalent today.

Successful website design, among other factors, should consider two important factors [2][3][4]: (1) Characteristics of users: Needs, abilities and limitations of different user groups (e.g.; gender, education level, age); and (2) Learning factor, which is related to use frequency of a specific website and internet use experience.

* Corresponding author.

The customers do not visit the airline websites frequently unless they want to schedule their flights. The websites are possibly used several times per year. Thus there is no learning factor of using an airline site. Hence, the learning factor cannot be taken into account in the design of these websites.

Current literature on usability and user experience focuses on general guidelines for the Internet. However, there is a need of usability and user experience guidelines applicable to specific types of websites [5].

Case studies are needed to identify the research needs and user experience issues related to the current websites so that specific design guidelines can be developed for the specific websites, such as airline websites.

In this paper, a usability case study was performed to identify various usability problems and develop recommendations that would improve websites of the three airline companies in Turkey: Atlasjet (atlasjet.com), Pegasus Airlines (flypgs.com) and Turkish Airlines (turkishairlines.com). Methodology and findings from heuristic evaluation and user testing are presented. Some recommendations that enhance user experience were provided as well.

2 Objectives

The objectives of this study were:

- Identify high-level UI design issues about the airline websites by conducting a heuristic evaluation
- Conduct a usability study to identify the details about the high-level issues found by the heuristic evaluation
- Provide solutions for the identified usability issues discovered via usability tests sessions.

3 Methodology

The study consisted of the following steps that ultimately prepared a means to execute the user experience study on the three airline websites:

- Determining the web design heuristics that were used for the evaluation of airline websites
- Performing a heuristic evaluation: identify high-level UI design issues with the airlines
- Developing task questions for user testing
- Obtaining permission for user testing from the Institutional Review Board at the university
- Recruiting the users: graduate, undergraduate and high school students
- Reading and signing of the informed consent
- Reading of the task instructions
- Filling out a pre-test questionnaire
- Pilot study

- Performing the tasks
- Filling out the post-test questionnaire
- Debriefing
- Statistical analysis of data and results
- Recommendations
- Conclusion

Web Design Heuristics. For the heuristic evaluation of the websites, the following sources were utilized: Research-Based Web Design & Usability Guideline by U.S. Department of Health and Human Services [3], Shneiderman's Eight Golden Rules of Interface Design [6] and Nielsen's [4] heuristic evaluation principles. However, [3] was used as the primary reference.

Heuristic Evaluation. Every single window of the websites of the three airline companies were evaluated according to the heuristics mentioned above by three evaluators who were somewhat knowledgeable of user experience principles. User experience issues with the designs were identified. Based on the identified major issues, two somewhat difficult task questions were tailored for user testing.

Participants. A total of 168 (81 males and 87 females) users participated in the study. The distribution of the participants with respect to gender and education level is shown in Table 1. Participant internet use frequency data stratified by education level is depicted in Table 2. It can be said that most of the participants were undergraduate students. A large group of undergraduate students and high school students mostly use internet more than 1 hour up to 4 hours a day, whereas most graduate students use internet 4 to 8 or more than 8 hours a day.

Table 1. User Groups and number of participants

	Undergraduate	Graduate	High School	Total
Male	57	16	8	81
Female	57	17	13	87
Total	114	33	21	168

Table 2. Internet usage of participants by education level

Internet Usage	Undergraduate	Graduate	High School	Total
More than 8 hours a day	21	13	2	36
More than 4 up to 8 hours a day	33	12	3	48
More than 1 up to 4 hours a day	51	4	10	65
Less than 1 hour a day	7	4	6	17

Pre-Test Questionnaire. The survey asked the participants to provide information on the following: gender, education level, and daily internet use frequency.

Task. Each participant visited three airline websites, Atlasjet, Pegasus and Turkish Airlines, randomly and one at a time while performing the following tasks at random:

- Task 1: Find the information about flight cancellation, and
- Task 2: Find the information about extra baggage prices.

Pilot Study. A pilot study was performed with 2 participants from each of three education group (a total of 6: 3 males and 3 females) to determine whether there was an issue with the tasks, post-test questionnaire or test procedure so the necessary improvement would be made before the actual test. The pilot study also provided an opportunity for the experimenter to become familiar with the test procedure.

User Testing. The subjects performed Task 1 and Task 2 randomly on each company’s website (a total of 6 tasks per participant) by using Loop 11 usability software. A 5-minute break is given after each task completion. The descriptive statistics of the values of each user performance variable were recorded by Loop 11 usability software. The performance variables were as follows: *task completion time, number of page views, success rate (or failure or abandon)*. In order to overcome the negative effect of probable inadequacies in English, users were all tested in Turkish. However, in this paper, tasks are described in English.

Post-Test Questionnaire. The survey contained a questionnaire for the participants to fill out upon completion of each task of the usability test. After the completion of each task, the participants rated the ease of each task using a Likert scale (-3 -2-1 0 1 2 3; where -3 corresponds to “very difficult”, 3 corresponds to “very easy” and 0 corresponds to “neutral”) (Table 3). Additionally, the participants were asked five subjective evaluation questions to rate the usability of the website on a Likert scale (-3 -2-1 0 1 2 3; where -3 corresponds to “strongly disagree”, 3 corresponds to “strongly agree” and 0 corresponds to “neutral”) and were asked open-ended questions about their experience with the interface. Loop11 is used to administer the questionnaire online and also obtain the descriptive statistics of the collected task data, while Google Docs is used to obtain the collected usability rating data of the website.

Table 3. Post-test questionnaire with 7-point Likert scale

This task was easy to complete.	() -3	() -2	() -1	() 0	() 1	() 2	() 3
The website is visually appealing.	() -3	() -2	() -1	() 0	() 1	() 2	() 3
The content of the website meets my expectations.	() -3	() -2	() -1	() 0	() 1	() 2	() 3
It is easy to go around the website and it is user-friendly.	() -3	() -2	() -1	() 0	() 1	() 2	() 3
The overall organization of the website provides easy access to task.	() -3	() -2	() -1	() 0	() 1	() 2	() 3
Terminology used in the website is clear.	() -3	() -2	() -1	() 0	() 1	() 2	() 3

Debriefing. The participants were asked open-ended questions about their experience with the website interfaces.

Statistical Analysis of the Data. Descriptive and inferential statistical analysis (ANOVA and Tukey tests) were performed to summarize data and determine the factor effects on the response variables. The four factors with their levels are shown in

Table 4. The effects of four factors (web-design of company, gender, internet use frequency and education level) on the task completion time (the response variable) for each of two tasks were investigated with Analysis of Variance (ANOVA) and Tukey's multiple comparison tests using Minitab (v.16) statistical package. For all comparisons, $p \leq 0.05$ was accepted as statistically significant.

Table 4. Considered factors and their levels

Factors	Levels
Gender	(0) Male
	(1) Female
Education Level	(1) High School
	(2) Undergraduate
	(3) Graduate
Internet Use Frequency	(1) Less than one hour
	(2) 1-4 hours
	(3) 4-8 hours
	(4) More than 8 hours
Web Design	(1) Atlas
	(2) Pegasus
	(3) Turkish Airlines

Usability Software. Web-based user experience tool Loop11 (www.Loop11.com) was used to conduct tests on users. Loop 11 tool allows the ability to create online tasks and questionnaires for testing user-experience of a website and also identifying usability issues with the website. According to the results of heuristic evaluation; two user testing tasks and a post-task survey questionnaire were created for these three websites which are common for three of them. Then, these tasks and questionnaires are transferred to the internet environment by Loop 11. Once a user starts a task in Loop 11 environment, he/she can see the task statement that appears at the top of the page with the "Task Complete" and "Abandon Task" icons. As the user tries to accomplish the task given, he/she can click "Task Complete" when he/she believes that task is accomplished successfully, or he/she can "Abandon Task" if he/she feels like accomplishing the task is improbable. Loop 11 records task completion times, success rates, and number of page views as the users attempt to accomplish the tasks.

4 Results

Heuristic Evaluation. A sample of findings through heuristic evaluation performed by three evaluators are shown in Table 5. It is obvious that some main principles of web design for usability are not followed by the web designers of the three airline companies.

User Testing: Descriptive Statistics of Performance Results. Task completion time, task completion success rate (or failure or abandon) and number of page views were recorded for each task and participant. Figure 1, Fig. 2 and Table 6 depicts the

results for the web design of three airline companies. Success rate of Atlasjet and Pegasus are similar (about 60%); however, Turkish Airlines has low success rate (about 7%), indicating severe issues with the design of this website. The highest success rate was around 60%, which is also low, and indicates some issues with the design of these websites as well. For both Task 1 and Task 2, the participants, on average, spent nearly twice as much time on the websites of Atlasjet and Turkish Airlines in comparison to Pegasus website.

Table 5. A sample of usability issues of the websites identified by heuristic evaluation

Atlasjet	Pegasus Airlines	Turkish Airlines
A hazard symbol (warning sign) is used to direct attention of customers to a bargain price.	Contrast level is poor	There are pages that are too crowded with items of information
Not easy to determine which data entry fields are required and which are optional. For example, asterisk is not used in reservation and ticket purchase sections.	Website check-in service is frequently used, but it is hard to find because of the color used and placement.	Not easy to determine which data entry fields are required and which are optional. For example, asterisk is not used in reservation and ticket purchase sections.
Although Istanbul is the hub of the company, it is not listed on the top of the menu (first row) to save user time.	Although Istanbul is the hub of the company, it is not listed on the top of the menu (first row) to save user time.	Although Istanbul is the hub of the company, it is not listed on the top of the menu (first row) to save user time.
Credit card numbers sections, which are considered as long data, are not partitioned for ease of data entry.	Credit card numbers sections, which are considered as long data, are not partitioned for ease of data entry.	Credit card numbers sections, which are considered as long data, are not partitioned for ease of data entry.
Includes many terms and acronyms and abbreviations that are familiar to designers and content writers, but not to users.	Users return to homepage many times during a search. There is no homepage button or company logo which leads the user to the homepage of the website.	Website is not well-prepared in terms of the information organization: it does not consider a clear and logical structure to typical users.
No search option on each page	No search option on each page	No search option on each page
Does not provide feedback to let users know where they are in the website by using bread-crumbs	Does not use color changes to indicate to users when a link has been visited	
Does not use color changes to indicate to users when a link has been visited	Keeps flight search option in irrelevant pages	
Does not contain a site map	Mouseover is used instead of pointing-and-clicking	
Does not convey baggage information in a well structured way	City list is very long, because it is not listed under country	

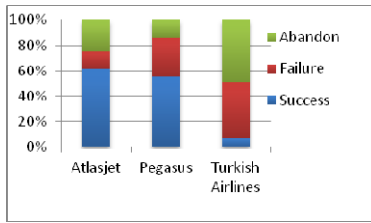


Fig. 1. Performance results for Task 1

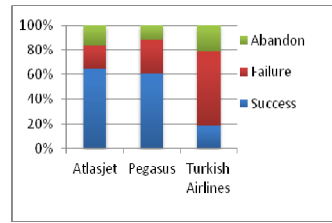


Fig. 2. Performance results for Task 2

Table 6. Overall performance results for the user testing of all participants

Website	Task no.	Success (%)	Fail (%)	Abandon (%)	Avg. time spent (sec)	No. of page views
Atlasjet	Task 1	61.3	14.3	24.4	158.9	3.2
	Task 2	64.3	19.1	16.7	150.4	2.4
Pegasus	Task 1	55.4	30.4	14.3	78.3	4.9
	Task 2	60.7	26.8	14.3	90.4	4.7
Turkish Airlines	Task 1	7.1	44.1	48.8	169.5	5.1
	Task 2	19.1	60.1	20.8	112.3	3.7

Factor Effects on Task Completion Times. Effects of classification factors Gender, Education Level, Daily Internet Use Frequency, and Web Design of Company on Task 1 and Task 2 completion times were investigated via analysis of variance (ANOVA). ANOVA results indicated significant ‘web design’ (company) effect on ‘task completion time’ ($p < 0.001$) for Task 1 but not on Task 2. Gender, internet use frequency, and education level did not have significant effect on task completion times of both tasks (Table 7 & Table 8). Although statistical analysis results did not indicate significant difference, the difference of means of Task 2 completion times between three airline websites were 60 s (Atlasjet (-) Pegasus), 38 s (Atlas (-) Turkish Airlines) and 22 s (Turkish Airlines (-) Pegasus), respectively. These indicate that it was taken longer to complete the Task 2 on the website of Atlasjet, which may be considered significant in practical applications. Following the ANOVA, performed Tukey’s multiple comparison tests indicated that Task 1 had taken significantly less time on Pegasus website in comparisons to the websites of Atlasjet and Turkish Airlines ($p < 0.05$); and no significant difference between the sites of Atlasjet and Turkish Airlines (Table 9).

Table 7. ANOVA for Task 1 Completion Time

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Company	2	836865	836865	418432	18.84	<0.001
Internet Usage	3	42925	53109	17703	0.80	0.496
Education	2	26021	23607	11803	0.53	0.588
Gender	1	30854	30854	30854	1.39	0.239
Error	495	10994475	10994475	22211		
Total	503	11931139				

Table 8. ANOVA for Task 2 Completion Time.

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Company	2	309926	309926	154963	2.01	0.135
Internet Usage	3	66482	100917	33639	0.44	0.727
Education	2	60953	49487	24743	0.32	0.725
Gender	1	161607	161607	161607	2.10	0.148
Error	495	38123758	38123758	77018		
Total	503	38722726				

Table 9. Tukey test results for company-group effect on the task completion time of Task1.

*Pairwise comparisons for levels of company-group	Difference of means	t	p	Grouping ($\alpha=0.05$)
Turkish A. vs Atlasjet	10.60	0.652	0.7915	Atlasjet A**
Pegasus vs Atlasjet	-80.65	-4.96	<0.001*	Turkish A. A
Turkish A. vs Pegasus	91.25	5.612	<0.001*	Pegasus B

*Significant ($p<0.05$); ** means that do not share a letter are significantly different

Post-Task Questionnaire Results. Post-task surveys by the participants on Loop 11 after the completion of tasks for each website. The summary results are presented in Table 10. According to the results, except unclear terminology, both Atlasjet's and Pegasus websites were slightly better than neutral. On the other hand, there were problems on every aspect asked about Turkish Airlines website design, except visual appearance. In addition, the tasks are found harder to complete when using Turkish Airlines website in comparison to using the websites of Atlasjet and Pegasus.

Table 10. Summary of post-task results of three airline websites

Post-Task Questions	Atlasjet	Pegasus A.	Turkish A. Airlines
This task was easy to complete.	0.76	0.82	-0.62
The website is visually appealing.	0.70	0.92	0.18
The content of the website meets my expectations.	0.86	1.05	-0.05
It is easy to go around the website and it is user-friendly.	0.83	0.78	-0.34
The overall organization of the website provides easy	0.77	0.65	-0.71
Terminology used in the website is clear.	-0.80	-0.64	-0.09
Total score	3.12	3.58	-1.63

5 Recommendations

To improve the design of the websites so there is a better user experience, some suggestions are made which are listed in Table 11.

Table 11. Recommendations for the design of the websites to improve the user experience

Atlasjet	Pegasus Airlines	Turkish Airlines
<p>Experienced users usually look first at the top item in a menu or list. So place Istanbul, which is most used, in the first row to save user time.</p> <p>Users should be able to easily determine which data entry fields are required and which are optional: Use asterisk in front of the ticket reservation and ticket purchasing indicating that these are required data entry fields.</p> <p>Provide a site map.</p>	<p>Experienced users usually look first at the top item in a menu or list. So place Istanbul, which is most used, in the first row to save user time.</p> <p>Destination city list is very long. Classify cities under country.</p>	<p>Experienced users usually look first at the top item in a menu or list. So place Istanbul, which is most used, in the first row to save user time.</p> <p>Users should be able to easily determine which data entry fields are required and which are optional: Use asterisk in front of the ticket reservation and ticket purchasing indicating that these are required data entry fields.</p>
<p>Provide a new flight search option on result page.</p> <p>Provide feedback to let users know where they are in the website by using breadcrumbs or changing the color of the tab.</p>	<p>Allow users to return to homepage by linking the user to the homepage by a clickable company logo.</p> <p>Provide a new flight search option on result page.</p> <p>Check-in service is frequently used. Place this hard to find tab in upper side of page with darker letters.</p>	<p>Avoid scroll up and down, especially the homepage. Remove unnecessary information from homepage and put it under a tab.</p> <p>Provide a new flight search option on result page.</p> <p>The windows are too dense to find target information. Arrange lists and links in a more structured way to decrease the density. Use shorter pages for homepage and navigation pages, and pages that need to be quickly browsed and/or read online</p>
<p>Improve contrast font size, style, and color: The color should be darker and more apparent. Contrast should be increased. Use dark letters on light background. The user should be provided to return previous pages by navigational options. Thus avoid full-screen for the links opened in new pages.</p> <p>Help users understand which parts of a website they have visited by use of text link colors (default: blue as an unvisited location/link and purple as a visited location/link).</p> <p>Avoid the use of a hazard symbol (warning sign) to direct attention of the customers to a bargain price.</p> <p>When users are expected to rapidly read and understand prose text, use black text on a plain, high-contrast, non-patterned background.</p> <p>Partition long data items such as credit card numbers into four parts to aid users in detecting entry errors, and reduce erroneous entries.</p> <p>Do not use words that typical users may not understand. Do not use unfamiliar or undefined acronyms or abbreviations</p>	<p>Improve contrast font size, style, and color: The color should be darker and more apparent. Contrast should be increased. Use dark letters on light background. The user should be provided to return previous pages by navigational options. Thus avoid full-screen for the links opened in new pages.</p> <p>Help users understand which parts of a website they have visited by use of text link colors (default: blue as an unvisited location/link and purple as a visited location/link).</p> <p>Prefer 'pointing-and-clicking,' rather than mousing over, when selecting menu items from a cascading menu structure.</p> <p>When users are expected to rapidly read and understand prose text, use black text on a plain, high-contrast, non-patterned background.</p> <p>Partition long data items such as credit card numbers into four parts to aid users in detecting entry errors, and reduce erroneous entries.</p> <p>Do not use words that typical users may not understand. Do not use unfamiliar or undefined acronyms or abbreviations</p>	<p>Improve contrast font size, style, and color: The color should be darker and more apparent. Contrast should be increased. Use black letters on light background. The user should be provided to return previous pages by navigational options. Thus avoid full-screen for the links opened in new pages.</p> <p>Help users understand which parts of a website they have visited by use of text link colors (default: blue as an unvisited location/link and purple as a visited location/link).</p> <p>Provide sufficient cues to clearly indicate to users that an item is clickable.</p> <p>When users are expected to rapidly read and understand prose text, use black text on a plain, high-contrast, non-patterned background.</p> <p>Partition long data items such as credit card numbers into four parts to aid users in detecting entry errors, and reduce erroneous entries.</p> <p>Do not use words that typical users may not understand. Do not use unfamiliar or undefined acronyms or abbreviations</p>

6 Conclusions

A usability evaluation of three airline companies in Turkey were performed by both heuristic and user testing methods. The results indicated a number of usability issues regarding the design of all three airline websites by both methods. The heuristic evaluation indicated a number of basic design mistakes with all three airline websites. User testing results indicated poor usability performance results, especially for the Turkish Airlines website. The success rate with Turkish Airlines website was a major problem. However, the success rates with the websites of the Atlasjet and Pegasus were also not too high, only about 60%. There were also huge differences among the task completion times between Pegasus and the other two airlines, indicating significant room for improvements. Post-Test Questionnaire results also supported these results: The task completion was difficult with the Turkish Airlines website; however, it was not also so easy on the other two websites. Overall, the participants had negative experience with the Turkish Airlines website. The usability experience of the participants with the other two privately-owned company websites were neutral or slightly better than neutral. Overall, the Pegasus Airlines website seems to have better design than the other two websites. These results were independent from gender, internet use frequency and education level. Recommendations were made to improve the user experience related design mistakes of the websites for all three companies.

In conclusion, all three company websites require major redesign efforts that utilize the user experience principles in order to improve customer experience, thus resulting in more business.

References

1. Nielsen, J.: Top 10 Mistakes in Web Design, <http://www.nngroup.com/articles/top-10-mistakes-web-design/>
2. Jordan, P.: An introduction to usability. Taylor and Francis, Abington (1998)
3. U.S. Dept. of Health and Human Services: The Research-Based Web Design & Usability Guidelines, Enlarged/Expanded edition. U.S. Government Printing Office, Washington (2006)
4. Nielsen, J.: Usability Engineering. Academic Press, Boston (1993)
5. Bainbridge, A.: Hotel booking process: Design & Usability (2003), http://www.tourcms.com/company/research/pdf/hotel_booking_process_february2003.pdf
6. Shneiderman, B.: Designing the User Interface: Strategies for Effective Human-Computer Interaction. Addison-Wesley Publ. Co., Reading (1987)