Usability Evaluation (2)
User Testing

Salvatore Fadda
944786
{name}@mail.polimi.it

Dario Miceli Pranio
928262
{name}@mail.polimi.it

Adriano Mundo
944684
{name}@mail.polimi.it

Website URL: https://wildocean.herokuapp.com

Delivery Date 6th July 2020

Table of Contents

1 Abstract

This document reports the *Usability Evaluation* of the Wild Ocean website (see the link above). *Usability* is a crucial aspect for the User Experience (UX) of an interactive application. It represents the easiness of use of an application. In this document it is evaluated with a *"user testing"* method.
The purpose of this document is to provide the reader with all the information obtained by testing our system with real users with the aim to discover issues that may arise during its usage and evaluate the effectiveness of the design choices.
The rest of this document is organised as follows:

- *Design of the study*: it defines the variables we decided to measure and how to measure them, tasks and target user
- *Execution of the study*: how, where and when the test was executed
- *Results*: describes the insights derived from the execution
- *Conclusion*: final comments from the evaluation

2 Design of the study

Usability evaluation is crucial for product development and we decided to do this at the end of the development process of our web application. The purpose is to find flaws and difficulties that user may have while interacting with the application, highlighting pain points in the design and obtain a feedback on the effectiveness of our final product.
The idea behind this kind of evaluation is to generate usability data directly from the end users in a task-fashion way. Hence, the users do things with the website as they should do during a real session.

To do what we have in mind we need to recruit a sample of what we identified as target users and assign them some tasks to perform on the website in a controlled environment. In our case, it will be the environment of a typical user session at home, office or similar location.
Among the various segment that could be final users of our website, we defined our *user profile* as a young adult whose age ranges from 20 to 35 years old who is familiar with technology. To better explain what we mean: we consider part of the target users who use websites at least one times a day for different purposes, so they are aware of the process of consulting a website to find information. We consider this profile as someone who is going to need this kind of application now or in the future.

Therefore, we recruited 5 people belonging to the user profile mentioned before and we defined some *test goals* in order to have a scope and make our test consistent. Among them, we would like to test the possibility of a user to request a service provided by the association or to take part at one of the organized events. Other goals are to test when someone want to join the association, gain information about the association objectives, directly contact the association.

Once defined the test goals and user profile we needed some tasks we want to user to perform and some variables to measure during the test; hence, we come up with a list of *tasks* to be done plus a list of *quantitative* and *qualitative variables.*

We identified the following *tasks:*

1. You heard about an event taking place in Rome during the month of July. You need to find more information about that event.
2. Sophie Turner is your favourite actress and she recently joined the association Wild Ocean. Find out what she is doing for the association.
3. Find a way to access to one of the services provided by the association.
4. Find out mission, vision, values of the association to understand more about the impact of the association for saving the ocean.
5. You are a university student or an employee who want to organize an event in collaboration with this new association you discovered. Try to get in touch with them.

During the test we need to measure the following variables for each one of the tasks we mentioned above.

Quantitative variables:

- *Effectiveness*: task success rate, where each one could be complete, partial, failure.
- *Efficiency*: as time on task
- *Errors:* wrong paths or actions during the task
- *Perceived task difficulty:* directly filled by the moderator on a scale between 1 and 5.

Qualitative variables:

- *Disorientation:* how the user felt disoriented while performing the tasks
- *Frustration:* how the user felt frustrated while performing the tasks
- *Satisfaction:* how satisfactory was completing the task in accordance with the user
- *Clarity:* if the user found the information on the website clear
- *Suggestions*: direct suggestions from the user about the website
- *Recommendability:* if the user would recommend the website to a friend
- *Unexpected behaviours:* if something unexpected happened during the navigation

3 Execution of the study

The Evaluation is divided in two parts. In the first one, moderators execute the designed test in a controlled environment while monitoring and keeping track of all the data about user behaviours. In the second phase, after all the data gathering, moderators discuss together about the collected information and derive insights on what the test has pointed out.

We prepared a form which was filled by the moderator during the tests. The form was mainly used to have an easy way to gather data. The form was useful necessary to keep the relevant information and measure the quantitative variables. A second form was directly filled by the end user for post-task data gathering. It was used to have a complete picture and gain valuable insights also for the qualitative variables.

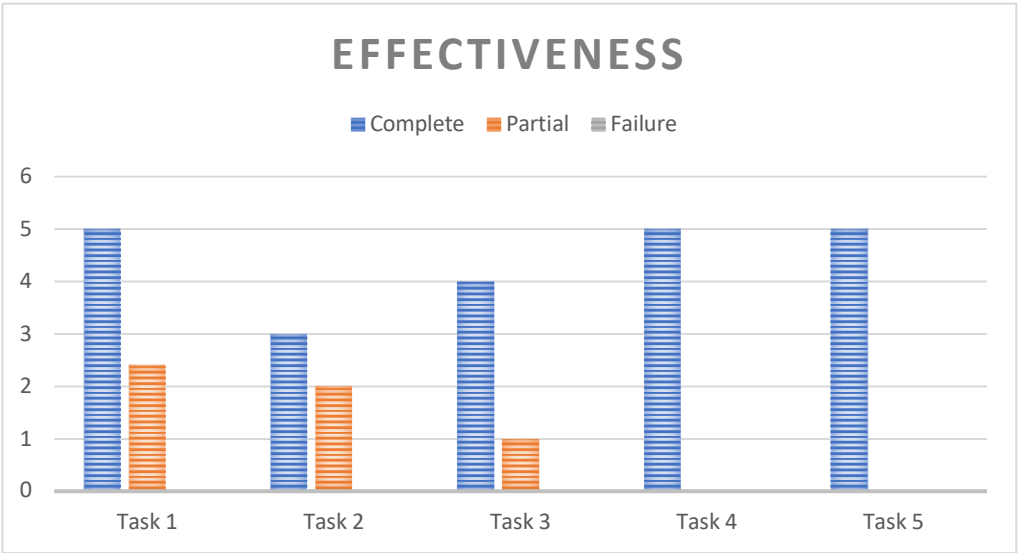
The hardware setting was composed by a PC or smartphone without specifying any OS and/or browser. We recorded the video and audio of every tester to make a check of what we tracked during the test and see again the session to discover insights in the user's behaviours we didn't notice during the live session.

During the test a lot of emphasis was pointed out to explain the objective of the experiment to the test taker, to be clear about task and environment setting while underlining that it can leave the test whenever he or she want.

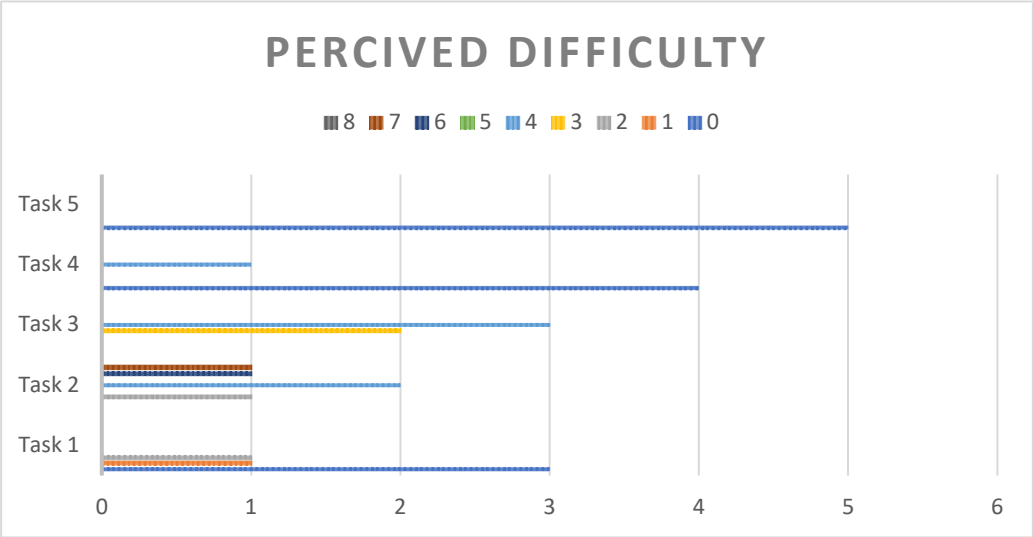
4 Results

The main problems encountered during the test, which lead to errors by test takers were in the second and third task.
In particular, for the second task test takers does not immediately find the group link to switch between different pages of the crew or they tried to reach the person page by staring from an event; while for the third task reading the paragraph text of practical info which explain to contact a member of the crew was not straightforward.

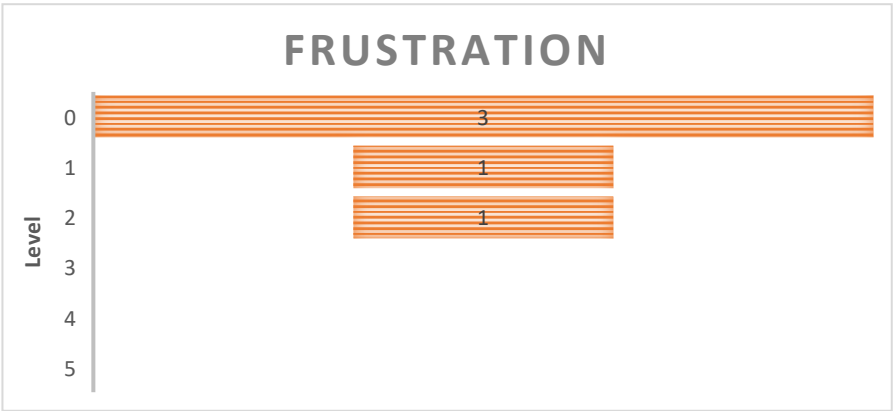
Fig.1 Andrea during the user testing of Wild Ocean

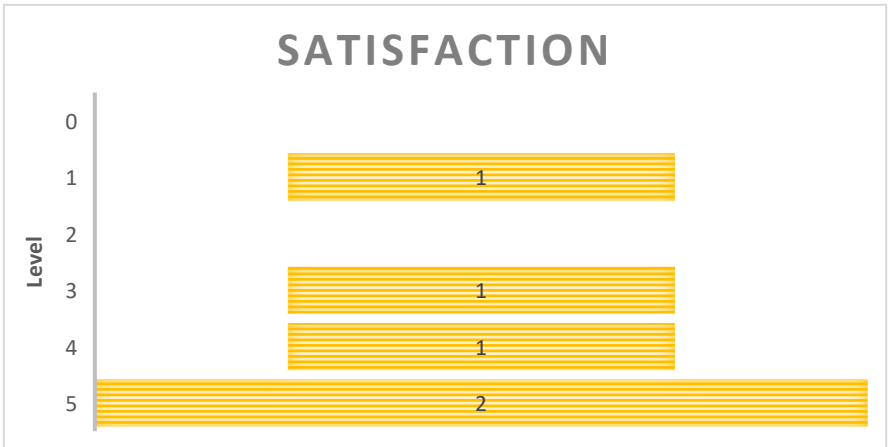


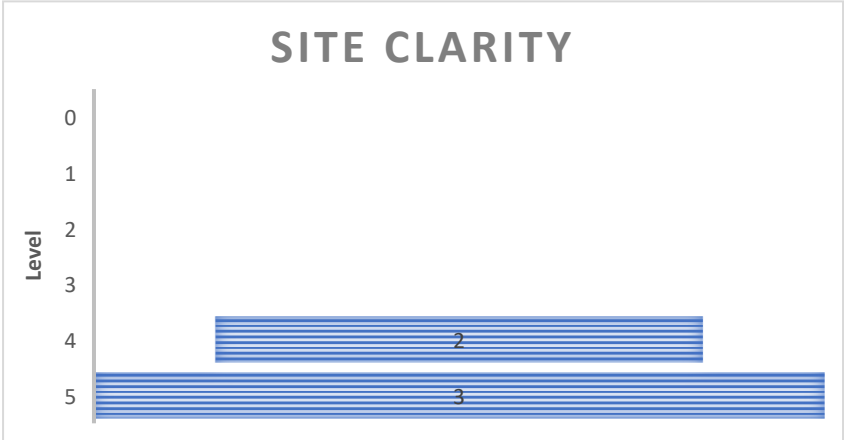
This graph represents how many testers have passed the 5 tasks completely, partially or in failure. The majority completed almost everything and only a few had difficulties completing tasks 1,2,3.



We decided to evaluate the "perceived difficulty" during the execution of tasks on a scale from 0 to 8, in which 0 represents the absence of difficulties and 8 represents great difficulties. The graph indicates for each task, how many testers have had a certain score. This representation makes us understand that tasks 5, 4 and 1 have not been complicated at all, unlike the 3 and 2 contain higher scores, especially the 2 in which appear 8,7 and 6.







We decided to evaluate the sensations perceived by our testers on an emotional scale, expressed numerically by the values 0,1,2,3,4,5 in which the value 0 indicates the absence of this sensation and 5 a notable perception of this during the test. The graph represents how many testers gave a certain score. It can be seen that almost nobody perceived the feeling of frustration, while everyone was satisfied with the completion of the test and happy with the clarity of the site.

	AVG TIME	AVG ERRORS	AVG DIFFICULTY	AVG EFFECTIVENESS
T1	18.8 sec	0	0.6	100%
T2	70.2 sec	1.6	4.6	60%
T3	57 sec	0.8	4.4	80%
T4	10 sec	0.2	0.8	100%
T5	14.6 sec	0	0	100%
TOTAL AVG	34.12 sec	0.52	2.08	88%

	TIME VARIANCE	ERRORS VARIANCE	DIFFICULTY VARIANCE	EFFECTIVENESS VARIANCE
T1	234.7 sec	0.2704	2.1904	1.44%
T2	1301.77 sec	1.1664	6.3504	4%
T3	523.49 sec	0.0784	5.3824	0.0064%
T4	581.77 sec	0.1024	1.6284	1.4%
T5	381.03 sec	0.2704	4.3264	1.4%

Reccomendability: 100%

The site behaves correctly without unexpected behaviours: 100%

Analysing the results we noticed that, in general, effectiveness is high. Task 2 has the lowest value and this makes us understand that the choice of using group links in "Crew" page, maybe is not the optimal thing to do. Furthermore, the number of errors per task is very low and clarity is high, and these tell us that the entire site is clear and intuitive for all the users.

5 Conclusions

To wrap-up we can conclude that most of the times people who use Wild Ocean website are comfortable at doing simple tasks and navigate the website to find all the information they need without encounter too much difficulties.
Most of the times tasks were done fasts and people find the website clear.

We noticed from errors and direct feedbacks that there's space for improvements in the usability of the website.
We can better design the group link of the "Crew" page to better highlight the possibility to switch between different page or we can have all the Crew members in one single page, but this can be an issue if the number of members increase.
Practical information section in the "Service" page can be put more in evidence to help people understand they need to contact one of the members in order to access to the service.
In All Services page title and image could also be a link to the kind of topic page and sometimes the fade-in effect during scroll has disoriented people to understand that there are more services.

At the end, we can conclude that this kind of test is very useful to find issues and flaws in our website and we noticed that even with a small number of users (5 in our case) is possible to find the most evident issues. This was another different kind of experience for our team that learnt a lot about how to test an application from the usability perspective and in this second case with a direct user testing. We appreciate this opportunity and we are sure this will be very useful in different situations during our future working life.

Annexes

Partecipant: Sara *Evaluator:* Adriano Mundo

TASK	EFFECTIVENESS	TIME ON TASK	ERRORS	DIFFICULTY
1	Complete	16s	0	1
2	Complete	35s	1	3
3	Partial	55s	1	5
4	Complete	6s	0	0
5	Complete	9s	0	0

Partecipant: Andrea *Evaluator:* Adriano Mundo

TASK	EFFECTIVENESS	TIME ON TASK	ERRORS	DIFFICULTY
1	Complete	18s	0	1
2	Partial	81s	1	4
3	Complete	47s	1	4
4	Complete	8s	0	0
5	Complete	19s	0	0

Partecipant: Pietro *Evaluator*: Salvatore Fadda

TASK	EFFECTIVENESS	TIME ON TASK	ERRORS	DIFFICULTY
1	Complete	20s	0	0
2	Partial	65s	1	7
3	Complete	25s	0	4
4	Complete	17s	1	4
5	Complete	15s	0	0

Partecipant: Francesco *Evaluator:* Salvatore Fadda

TASK	EFFECTIVENESS	TIME ON TASK	ERRORS	DIFFICULTY
1	Complete	20s	1	2
2	Complete	35s	1	6
3	Complete	30s	0	4
4	Complete	10s	0	0
5	Complete	18s	0	0

Partecipant: Chiara *Evaluator:* Dario Miceli

TASK	EFFECTIVENESS	TIME ON TASK	ERRORS	DIFFICULTY
1	Complete	20s	0	0
2	Complete	135s	4	7
3	Complete	128s	2	4
4	Complete	9s	0	0
5	Complete	12s	0	0