User Interface Evaluation of the CP's Website (Proposal)

Luís Cruz[1]

[1]*MAP-i, Joint Doctoral Programme in Computer Science*

January 4, 2014

Abstract

This document proposes an usability evaluation for the website of the company *CP – Comboios de Portugal*. The company and the website are briefly described, as well as the users focused by the evaluation and the supported tasks. One analytical method and one empirical method are going to be applied in this evaluation: *Heuristic Evaluation* and the *Usability Test*, respectively, both described in this document.

1 Introduction

This project aims to evaluate the user interface of CP.pt[1] — the official website of *CP - Comboios de Portugal, E.P.E.*

CP is a public portuguese company responsible for rendering national and international passenger rail services. In the year 2012, CP had 4690 employees, transported 122 million passengers and almost 8713 thousand metric tons (CP - Comboios de Portugal, 2012). They provide 3 main kinds of rail transportation services: *urban* in the cities of Oporto and Lisbon; *National* with regional services and the fast lines of *Alfa Pendular* and *Intercidades*; and *International*.

Through the website, CP's customers can check the timetables, buy tickets, get information about the available lines and special offers and read some news related with CP services. In order to buy tickets, the website provides the *netTicket* service, which requires the customers to have an account in their *myCP* service and it is only available for the long distance trains *Intercidades* and *Alfa Pendular*.

According to the website the graphical interface was optimally designed for windows with 800×600 pixels of resolution. A view of the website is depicted in the figure 1, using a window with the same resolution.

2 Users and Context

CP customers vary according to the service provided. Many college students, workers and pensioners use the regional and urban services for small and medium distances. Long distance services are more used by college students that are away from home, tourists, and executive workers. Unfortunately, no official document stating the segmentation of the CP.pt website's users was found.

It is noticeable that CP services have a lot more passengers during school time, which means that students are an important segment of CP's customers. Besides, most of the students have good experience with the WEB, so the CP.pt website is expected to be a great tool to them. Therefore, this usability evaluation will focus in the segment of college students, which might be portuguese citizens as well as foreigners that study or want to study in Portugal and are able to speak English.

Many scenarios can apply for the use of the website by students. Some times they leave the classes earlier and need a way of quickly check if there are other alternative trains that can take them home earlier. Also sometimes there is no direct train to their destination, so they have to catch another in the the middle of the travelling. Another scenario is when the weekend is over and the student has to buy his/her ticket from home to his/her university city. Buying it from the website is more convenient since
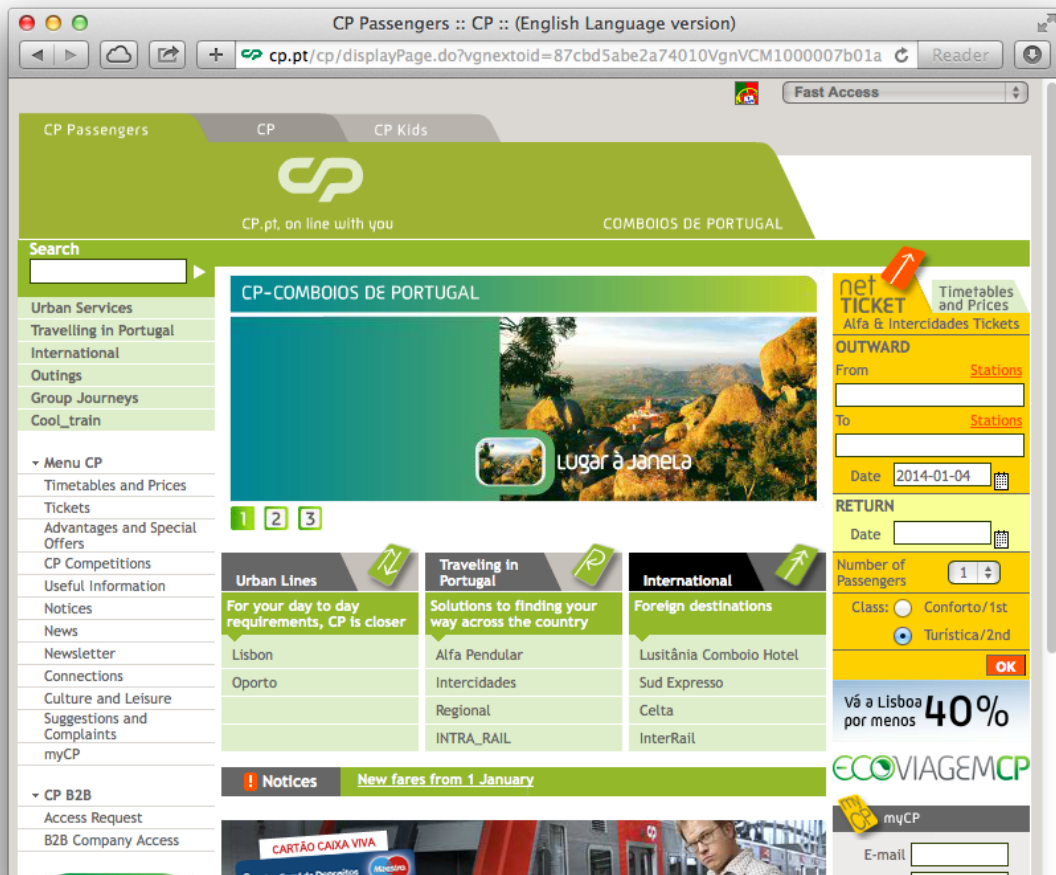
---

[1]Available at: http://www.cp.pt

Figure 1: View of the main page of CP.pt website in a window with resolution 800×600 pixels, using the internet browser Safari Version 7.0.1.

the student can avoid wasting time in the ticket lines and can grant a seat for his/her trip. Therefore, the following two contexts are considered the most important in terms of usability:

- Check the timetable to find any suitable train for the trip and the respective prices.

- Buy a ticket for long distance trains with reserved seats.

3   Usability Evaluation Methodology

The evaluation will be taken using two paradigms: *Analytical* and *Empirical*.

Analytical methods do not need to involve users — they are based on inspection methods. Some well known analytical methods are the *Heuristic Evaluation* (HE) proposed by Nielsen and Molich (1990), the *Cognitive Walkthrough* (Wharton et al., 1994) and its variant *Streamlined Cognitive Walkthrough* (Spencer, 2000).

Empirical methods involve the user in the evaluation process through Usability tests, involving *observation* and *query* techniques, through *controlled experiments*, in a more scientific approach, or even *questionnaires*, *focus groups*, etc.

In this evaluation, the used analytical method will be the *Heuristical Evaluation* and the empirical method will be the *Usability Test*. These methods are described in the next sections.

4 Heuristic Evaluation

The elected analytical method for this evaluation was the *Heuristic Evaluation*, because it is cheap, intuitive, easy to motivate people to do it and provides that useful results can be obtained (Nielsen and Molich, 1990).

4.1 Methodology

This method proceeds by having a small set of evaluators judging the system according to some general principles of interaction design, *usability heuristics*. It has been shown that a number of evaluators between 3 and 5 provides good results and that there is no point in having more than 10 evaluators (Nielsen and Molich, 1990). Nielsen (1995a) proposed the 10 most important usability heuristics for User Interface Design:

- Visibility of system status

- Match between system and the real world

- User control and freedom

- Consistency and standards

- Error prevention

- Recognition rather than recall

- Flexibility and efficiency of use

- Aesthetic and minimalist design

- Help users recognise, diagnose, and recover from errors

- Help and documentation.

Heuristic evaluation was originally developed for evaluators who had some knowledge in usability but who were not necessarily usability experts (Nielsen and Molich, 1990), however, it has been showed that the method is also very effective for expert evaluators (Nielsen, 1992).

In order to group the findings of this heuristic evaluation, the severity ranking proposed by Nielsen (1995b) with a scale from 0 to 4:

0 - I don't agree that this is a usability problem at all;

1 - Cosmetic problem only: need not be fixed unless extra time is available on project;

2 - Minor usability problem: fixing this should be given low priority;

3 - Major usability problem: important to fix, so should be given high priority;

4 - Usability catastrophe: imperative to fix this before product can be released.

In this evaluation, two experienced evaluators were asked to focus on the interface elements related with the following top-level user stories:

- As an unregistered user I want to check time tables and prices for travelling from a given train station to another one.

- As an unregistered user I want to buy a ticket for the fast train services *Alfa Pendular* and *Intercidades* (it may include a registration process).

4.2 Results

After completing the heuristic evaluation 11 problems were found. They were sorted in table 1 according to the severity rating.

number of problems

#	Problem	Severity Rating	Violated Heuristic(s)
1	The "Timetable and Prices" form is very similar with the form for buying tickets. Non expert users might not notice the difference.	3	Error prevention.
2	There is an auto complete feature in the "Timetable and Prices" and NetTicket forms, which whenever the user misspells a letter of the station name, the system may complete with another station's name, and the user has to hit the backspace button one more time than usual.	3	Error prevention; Help users recognize, diagnose, and recover from errors.
3	The seat selection is made in a uncommon way. One has to first click on the previous seat and then on the desired seat.	3	Consistency and standards.
4	The services of trains are defined as acronyms which are not used by the users (eg., IC means Intercidades)	2	Consistency and standards; Match between system and the real world.
5	When an unregistered user wants to buy a ticket after choosing a specific train, the system asks him/her to register but after finishing the registration, he/she has to search again for the same train.	2	Recognition rather than recall.
6	In order to check the price in the timetable it is necessary to click in a link which when clicked replaces itself into the respective price, but only for one train at a time.	2	Flexibility and efficiency of use.
7	When searching for a travelling ticket to buy, the user selects whether is travelling in first or second class in the beginning of the search. It is defined as second class by default and might not be perceptible. Through the next steps the user cannot change it.	2	User control and freedom.
8	In the buying process there is no back button, besides the one provided by the browser that is not recommended by the system since he presents an alert warning that the user is about to leave the page.	2	User control and freedom.
9	In the timetable results, there is a clickable down arrow that does not do anything	1	Aesthetic and minimalist design; Consistency and standards.
10	There is a Time input field when in the form "Timetables and Prices", which might be unnecessary and provides no useful filtering in the results.	1	Aesthetic and minimalist design.
11	In the timetable results there is a number identifying the row that does not provide any useful information	1	Aesthetic and minimalist design.

Table 1: Problems found in the Heuristic Evaluation.

5 Usability Test

The usability test was carefully reported using the Common Industry Format (CIF) (Stanton, 2006). The

As suggested in (Mitchell, 2007) it is beneficial for the moderator to get the participant's opinion during the test session. The usability test combined both *observation* and *query*.

The participant will be asked to think aloud, describing every step he/she makes during the tasks. The moderator will be directly observing the participant and taking some notes using the Data Logging Form, provided in appendix **??**, and after each task the moderator makes a short interview focusing in the questions provided in the Usability Test Plan (see appendix **??**). The screen and the audio is also recorded for indirect observation.

After the completion of each task the participants answer a short questionnaire, as well as after the whole test they answer a post-test questionnaire. Also, they are interviewed by the moderator in an informal way, trying to answer some questions clearly stated in the Usability Test Plan document (see **??**).

5.1 Workplace

It is important to create a workplace regarding features that might affect the results.

According to (**?**), the most common resolution used in WEB is 1366×768, so in this experiment a screen with approximately the same resolution is used: 1440×990.

For accessing the application, the browser Safari 7.0.1 with default settings is used. The internet connection has an average download and upload speeds of $4Mbit/s$ and $1Mbit/s$, respectively. The screen and audio is recorded using the tool QuickTime Player 10.3. During the test, the moderator stands next to the participant which has free access to the laptop, taking in account that the participant feels comfortable.

6 Results

The reports of the usability test comply with the Common Industry Format (CIF) supporting a summative usability evaluation (Stanton, 2006).

A   Usability Test Report

The original document is available at `http://paginas.fe.up.pt/~luiscruz/cp_usability/`

<div>

Usability Test Report of CP's Website
(Proposal)

TESTED BY: LUÍS CRUZ[1]

[1]*MAP-i, Joint Doctoral Programme in Computer Science*

Tested: January 20, 2014

Common Industry Format for Usability Test Report v1.1
Date prepared: January 27, 2014
Prepared by: Luís Cruz (luiscruz@fe.up.pt)

Executive Summary

vthe identity and a description of the product
va summary of the method(s) of the test including the number of and type of participants and their tasks.
vresults expressed as mean scores or other suitable measure of central tendency
w the reason for and nature of the test
w tabular summary of performance results.

1   Introduction

1.1   Product Description

CP.pt is the official website of emphCP - Comboios de Portugal, E.P.E, the public portuguese company responsible for rendering national and international passenger rail services.

CP customers vary according to the service provided. Many college students, workers and pensioners use the regional and urban services for small and medium distances. Long distance services are more used by college students that are away from home, tourists, and executive workers. Unfortunately, no official document stating the segmentation of the CP.pt website's users was found.

It is noticeable that CP services have a lot more passengers during school time, which means that students are an important segment of CP's customers. Besides, most of the students have good experience with the WEB, so the CP.pt website is expected to be a great tool to them. Therefore, this usability evaluation will focus in the segment of college students, which might be portuguese citizens as well as foreigners that study or want to study in Portugal and are able to speak English.

Many scenarios can apply for the use of the website by students. Some times they leave the classes earlier and need a way of quickly check if there are other alternative trains that can take them home earlier. Also sometimes there is no direct train to their destination, so they have to catch another in the the middle of the travelling. Another scenario is when the weekend is over and the student has to buy his/her ticket from home to his/her university city. Buying it from the website is more convenient since the student can avoid wasting time in the ticket lines and can grant a seat for his/her trip.

From the features implemented in the website, the following were tested:

- Choose between Portuguese or English versions

- Check timetables for a trip.

- Buy fast train tickets, being the features that were tested.

1

</div>

1.2 Test Objectives

The aim of the test was to validate the usability of the main features of the CP.pt — finding the most suitable train and buying tickets. It is important that these tasks are easy to learn.

Representative users were asked to complete some tasks, measures were taken of effectiveness, efficiency and satisfaction, and some notes about the users' opinion were taken in order to have ideas for some improvements that can be made.

2 Method

2.1 Participants

This test had 2 participants. Both were college students with ages between 21 and 26 years old. They are intermediate level users, that frequently use WEB applications. They already have experience in other transportation company's websites.

Students are used to get things fast and with an attractive design. Usually they have a laptop with a 13 or 15 inches screen, and use the Eduroam network when studying in the university.

2.2 Context of Product Use in The Test

2.2.1 Tasks

The tasks that the participant has to accomplish are the following:

1. Select the English Version of the application.
2. Find the schedule for a trip from Braga to Aveiro.
3. Find a cheap trip from Braga to Aveiro.
4. Buy a ticket from Braga to Porto.

These tasks are described in more detail in the *Usability Test Plan*, available in the appendix B. For each task, all the steps were defined in order to desscribe how the task can be efficiently accomplished.

These tasks were selected for being the features that are expected to be of the most important use. Every transportation website has these features and add great value to the customers, so it is important that they meet the users needs.

All the completion and performance criterias are also described in the Usability Test Plan (see appendix B).

2.2.2 Test Facility

The test was made in a study room at the faculty. The moderator is sitting next to the participant in order to make the observation and query and give assistance. The screen and audio were recorded using the tool QuickTime Player 10.3 which is invisible to the user and does not affect the user experience.

2.2.3 Participant's Computing Environment

According to (?), the most common resolution used in WEB is 1366×768. In this experiment a 13 inches RGB screen with approximately the same resolution was used: 1440×900. For interaction with the application the participant used an Apple laptop keyboard, and an Apple Magic mouse. Some other devices were available if the user didn't feel comfortable with these devices, however, they were not used.

The browser Safari 7.0.1 with default settings was used with an internet connection which had an average download and upload speeds of $4Mbit/s$ and $1Mbit/s$, respectively.

2.2.4 Test Administrator Tools

A *Data Logging Form* (see C) was designed providing the moderator with a tool to record some notes about each task of each participant. The form has some variables . and a generic questionnaire to be asked to the user. It also provides a space to take some notes while the moderator conducts a small post-task interview.

`define variables`

All the task information is provided in the *Usability Test Plan*, available in the appendix B, defining a script about how the moderator should conduct the experiment is provided. All the steps that are necessary to finish a task are clearly described as well as some guidelines with the important questions for the post-task interview.

After the test, the participants were asked to answer a post-test questionnaire based on () provided in appendix D.

As it was already mentioned, during the experiments, the screen and voice were recorded using the tool QuickTime Player 10.3.

2.3 Experimental Design

vDescribe the logical design of the test. Define independent variables and control variables. Briefly describe the measures for which data were recorded for each set of conditions. `variables`

2.3.1 Procedure

The participants were informed that the usability of CP's website was being tested, to find out whether it met the needs of users such as themselves. They were told that it was not a test of their abilities. They were asked to sign a consent form.

Participants were given introductory instructions. The evaluator reset the state of the computer before each task, and provided instructions for the next task.

The participant could ask for assistance and make questions whenever they find necessary, in order clarify any part of the task. All assistances were logged by the moderator. Also there was no time limit for the task completions, but if the moderator feels that the participant is stuck in some part, he was allowed to give some hints if properly logged.

After each task the moderator conducted a small interview trying to answer some crucial questions provided in the Usability Test Plan (see appendix B).

The participants were non remunerated voluntaries and during each test session only the moderator and a participant were present.

2.3.2 Participant General Instructions

The instructions were given personally by the moderator to each participant. The test session proceeds by having only one user in the room with a moderator. Whenever the user needed help he/she could simply ask for help. The participant was asked to use the *think-aloud* technique, describing every step he/she makes during the tasks. The moderator will be directly observing the participant and taking some notes using the Data Logging Form, provided in appendix

2.3.3 Participant Task Instructions

Before starting each task the moderator explained what it was expected to accomplish in the following task. The task instructions are very simple, being described in a short sentence, as stated in section 2.2.1.

2.4 Usability Metrics

vExplain what measures have been used for each category of usability metrics: effectiveness, efficiency and satisfaction. Conceptual descriptions and examples of the metrics are given below.

2.4.1 Effectiveness

For measuring effectiveness the following measures were considered:

Completion Rate the percentage of participants that correctly finished each task.

Unassisted Completion Rate the percentage of participants that correctly finished each task without assistance.

Number of Assistances The average of assurances given in each task.

Back Button hits The average of times the user hit the Back Button in the browser.

Errors The number of times a user had to repeat parts of the task.

2.4.2 Efficiency

Efficiency was accessed by measuring the following parameters:

Task time The average time the users took to correctly complete each task.

Completion rate efficiency mean completion rate/mean task time.

2.4.3 Satisfaction

Satisfaction is a subjective measure that correlates with the user's motivation to use a product. The standardized instrument *System Usability Scale* (SUS) provides a 10 item questionnaire with five-scale responses that can be converted into a score. After the test the users answered this questionnaire. `reference SUS`

Also a post-test questionnaire based on was given to the participants (available in appendix ??). This questionnaire intended to provide some insights about the way the design should be changed to make it more suitable to the user. `http://www.`

3 Results

3.1 Data Analysis

3.1.1 Data Scoring

The method by which the data collected were scored should be described in sufficient detail to allow replication of the data scoring methods by another organization if the test is repeated. Particular items that should be addressed include the exclusion of outliers, categorization of error data, and criteria for scoring assisted or unassisted completion.

3.1.2 Data Reduction

The method by which the data were reduced should be described in sufficient detail to allow replication of the data reduction methods by another organization if the test is repeated. Particular items that should be addressed include how data were collapsed across tasks or task categories.

3.1.3 Statistical Analysis

3.2 Presentation of the Results

Effectiveness, Efficiency and Satisfaction results must always be reported

3.2.1 Performance Results

TABLE

A Consent and Recording Release Form

`reference`

CONSENT AND RECORDING RELEASE FORM

I agree to participate in the study conducted and recorded by the MAP-i student Luís Cruz.

I understand and consent to the use and release of the recording by Luís Cruz. I understand that the information and recording is for research purposes only and that my name and image will not be used for any other purpose. I relinquish any rights to the recording and understand the recording may be copied and used by Luís Cruz without further permission.

I understand that participation in this usability study is voluntary and I agree to immediately raise any concerns or areas of discomfort during the session with the study administrator.

Please sign below to indicate that you have read and you understand the information on this form and that any questions you might have about the session have been answered.

Date: _____
Please print your name: _____
Please sign your name: _____
Thank you!
Your participation is kindly appreciated.

B Usability Test Plan

The plain original document can be accessed at `http://paginas.fe.up.pt/~luiscruz/cp_usability/`

C Data Logging Form

The moderator of the usability test observes the behavior of the participant while taking notes in the Data Logging Form. Each task needs one Data Logging Form.

Logging Form Sheet
CP.pt Usability Test
MAP-i Doctoral Programme
Responsible Moderator: Luís Cruz
22 January 2014

Participant ID: _____
Date: _____
Time: _____

Task _____	Performance	Observations and Comments
Completeness	Correct / Incorrect	
Assists	Tally	
Negative Remarks	Tally	
Back Button Hits	Tally	
# of steps made differently	Tally	
# of Errors	Tally	

Task start trigger: _____ Task end: _____ Time to complete: _____

Based on the interview guided by the questions defined in the *Usability Test Plan* document, the following notes were taken:

_____
_____
_____
_____
_____
_____
_____
_____
_____

D Post-Test Questionnaire

The moderator of the usability test observes the behavior of the participant while taking notes in the Data Logging Form. Each task needs one Data Logging Form.

Post-Test Questionnaire[1]

CP.pt Usability Test
MAP-i Doctoral Programme
Responsible Moderator: Luís Cruz
22 January 2014

Participant ID: _____
Date: _____
Time: _____

For our last activity I'm going to give you a short questionnaire that I would like you to fill out. The questionnaire will give you a series of statements about the CP.pt Web site. I would like you to rate your agreement with each statement.

For each statement, please circle a number to indicate the level to which you agree with each statement. If you feel a statement isn't relevant to your experiences with the Web site, please feel free to skip it.

1. The homepage is attractive — Disagree ○○○○○ Agree
2. The overall site is attractive — Disagree ○○○○○ Agree
3. The site's graphics are pleasing — Disagree ○○○○○ Agree
4. The site has a good balance of graphics versus text — Disagree ○○○○○ Agree
5. The colors used throughout the site are attractive — Disagree ○○○○○ Agree
6. The typography is attractive — Disagree ○○○○○ Agree

7. The homepage's content makes me want to explore the site further — Disagree ○○○○○ Agree
8. It is easy to find one's way around the site — Disagree ○○○○○ Agree
9. You can get information quickly — Disagree ○○○○○ Agree
10. It is fun to explore the site — Disagree ○○○○○ Agree
11. It is easy to remember where to find things — Disagree ○○○○○ Agree
12. Information is layered effectively on different screens — Disagree ○○○○○ Agree

13. The homepage is attention-getting — Disagree ○○○○○ Agree
14. Information is easy to read — Disagree ○○○○○ Agree
15. Information is written in a style that suits me — Disagree ○○○○○ Agree
16. Screens have the right amount of information — Disagree ○○○○○ Agree
17. The site effectively communicates the company's image — Disagree ○○○○○ Agree
18. Information is relevant — Disagree ○○○○○ Agree

[1] Adapted from *AIR, Human Factors and Research Design* group's work.

19. The site is designed with me in mind — Disagree ○○○○○ Agree
20. The site's content interests me — Disagree ○○○○○ Agree
21. The site's content would keep me coming back — Disagree ○○○○○ Agree
22. The site has characteristics that make it especially appealing — Disagree ○○○○○ Agree
23. The site reflects progressive, leading edge design — Disagree ○○○○○ Agree
24. The site is exciting — Disagree ○○○○○ Agree

25. The site is well-suited to first-time visitors — Disagree ○○○○○ Agree
26. The site is well-suited to repeat visitors — Disagree ○○○○○ Agree
27. The site has a clear purpose — Disagree ○○○○○ Agree
28. It is always clear what to do next — Disagree ○○○○○ Agree
29. It is clear how screen elements work — Disagree ○○○○○ Agree
30. Mistakes are easy to correct — Disagree ○○○○○ Agree

And that's it! Please return this questionnaire to the moderator.
Thank you for your collaboration!

References

References

CP - Comboios de Portugal (2012). Relatório e contas consolidadas do grupo CP de 2012. Available at: `http://cp.pt/cp/displayPage.do?vgnextoid=360a4ff42133f310VgnVCM100000be01a8c0RCRD`. Accessed: January 4, 2014.

Mitchell, P. P. (2007). *A step-by-step guide to usability testing*. iUniverse, Inc.

Nielsen, J. (1992). Finding usability problems through heuristic evaluation. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 373–380. ACM.

Nielsen, J. (1995a). 10 usability heuristics for user interface design. *Nielsen Norman Group, available at* `http://www.nngroup.com/articles/ten-usability-heuristics/`.

Nielsen, J. (1995b). Severity ratings for usability problems. *Nielsen Norman Group, available at* `http://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/`.

Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256. ACM.

Spencer, R. (2000). The streamlined cognitive walkthrough method, working around social constraints encountered in a software development company. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 353–359. ACM.

Stanton, B. (2006). What is the cif? Industry USability Reporting, available at `http://zing.ncsl.nist.gov/iusr/documents/whatistheCIF.html`. Accessed: January 4, 2014.

Wharton, C., Rieman, J., Lewis, C., and Polson, P. (1994). The cognitive walkthrough method: A practitioner's guide. In *Usability inspection methods*, pages 105–140. John Wiley & Sons, Inc.