HCI Report: Website Evaluation

Rakan Alassaf & Tarek Amr

Abstract

Usability is defined by ISO/IEC 9241 as the extent to which software products satisfy the users' needs in an effective and efficient manner [Rusu et al., 2011]. In this study we introduce the various sets of usability evaluation and design guidelines available today. Then, we apply a subset of those evaluation guidelines to three accommodation booking websites, and attempt to offer an alternative design that covers the deficiencies found in our evaluation.

1 Introduction

Usability is defined by ISO/IEC 9241 - the standard body concerned with ergonomics of human-computer interaction - as the extent to which software products satisfy the users' needs in an effective and efficient way [Mariage et al., 2005, Rusu et al., 2011]. [Matera et al., 2006] highlighted that usability tests were usually applied as post-production tests, to make sure that the final product meets the requirements agreed-upon with customers. However, iterative design has been proven now to be more efficient and cost-effective, and this is why software usability comes into the picture form the early phases of design; it is also checked throughout the different production processes. This comes in accordance with the user-centred design, and as pointed by [Bevan, 1997], software products (including websites) only meet the goals of their producers once they have met those of the users; hence businesses should align the two and ensure that their software products suites the key-scenarios of use.

This report is organized as follows. In section 2 we will start with a brief on websites usability guidelines, then in 2.1 we will discuss the different evaluation methods. We will then have a case-study for existing websites in section 3. Finally, we will present our alternative design in section 3.3, where we will try to cover the deficiencies found during the evaluation.

2 Website Usability Guidelines

[Nielsen and Hackos, 1993] emphasised that for a software to be usable, it has to comply with the following 5 attributes. *Learnability*, *Efficiency*, *Memorability*, *Few errors* and *Users satisfaction*. As our main goal here is to study usability of web content, we better focus on how those attributes were further explained by [Nielsen, 1995], after adapting them for web navigation. He summarized the above attributes as follows. Users should be able to learn the the basic navigation options of a web page and to find their desired content easily. They also should be able to reach such content in an efficient and quick way, and be able to remember the navigation options if they happen to return to the website after a while. Users should also be guided to follow the correct links until they reach their desired content, and if they happen to make mistakes during this process, they should be able to recover by going back to their previous location easily. Ultimately, the whole web experience should be pleasant and more satisfactory compared to any existing alternatives.

However, as noticed by [Bevan, 1997], there are opposing factors that make website navigation difficult. The content and the structure of the websites are usually built after the internal structure of the content providers, rather than the users' needs. Secondly, the content itself often needs to be adapted to the web, especially when it is originally intended to be provided in a printed form. Finally, web pages are sometimes not subject to the same quality measures for printed materials, due to the lower cost of creating the former.

To give users efficient access to the content, especially in large websites, [Matera et al., 2006] suggested complementing the hierarchical navigation with shortcrust for experience users, such as search options. Also, when it comes to in-page organization, he suggested the following:

- Core section: This is where the main message of the page is comes. (For example: at the top or in the central part of the page).

- Peripheral section: This is where auxiliary and complementary information comes. (For example: footers and side bars).

- Interconnection section: This is where links to other pages are located.

[Matera et al., 2006] warned that following the usability principles is not enough, and softwares also have to be evaluated. Also, as mentioned earlier, usability evaluation is an iterative process and should be done throughout the software production process; hence, in the next section we will assess the different evaluation guidelines.

2.1 Evaluation Guidelines

Usability evaluation can broadly be categorized into *inspection* and *testing* methods. In the former, experts study and report usability issues, while in the latter, issues are discovered by observing the users during their interaction with the interface [Rusu et al., 2011]. Four narrower evaluation categories were listed by [Nielsen, 1994]: *formal*, *informal*, *empirical* and *automatic*. We will focus on the last three here, as the first one has been abandoned in favour of the formal methods and/or grouped into other methods such as *cognitive walk-through* [Hollingsed and Novick, 2007], which is beyond the scope of this report.

2.1.1 Informal (Heuristic) Evaluation

[Hertzum and Jacobsen, 2001] elaborated that in this method a small group of experts are asked to evaluate the interface in accordance with a set of usability principles (or heuristics). They added that each of evaluator should conduct the evaluation independently of the others, and they should only be allowed to communicate after completing their evaluation. Of the widely used heuristics guidelines, are the following 10 principles presented by [Nielsen, 2005]:

1. Visibility of system status

2. Match between system and real world (Mapping)

3. User control and freedom

4. Consistency and standards

5. Error prevention

6. Recognition rather than recall

7. Flexibility and efficiency of use

8. Aesthetic and minimalist design

9. Users can recognize and recover from errors

10. Documentation and Help

[Matera et al., 2006] highlighted that evaluators normally start by visiting the system to get a general feel of its flow, they should then revisit it at least one more time to evaluate the specific components of its interface and their functionalities. They presented a sample (table 1) for how the evaluators' findings are normally presented.

In the end, [Rusu et al., 2011] commented on heuristic evaluation stating that they are easy, cheap and can help in finding many usability issues quickly; however some issues can be missed, especially the business-specific or domain-specific ones.

Problem Found	Violated heuristic	Severity	Suggested improvement
Download time is not indicated	Visibility of system status	High	Use a scrolling bar for representing the time left till the end of download

Table 1: A sample of heuristic evaluation findings, as shown by [Matera et al., 2006]

2.1.2 Empirical (Task-based) Evaluation

[Levi and Conrad, 1997] explained that empirical tests are performed by inviting users to carry out a specific task, in which the major system functionalities are involved. Observers should document the results, such as the number of tasks accomplished, time taken and/or number of pages navigated by each user during the task. The observers can also offer help to users; however they should take notice with the incidents where help was needed. They also suggested that discussions taking place between the participants afterwards is helpful in collecting their reactions and suggestions for improvements. [Matera et al., 2006] added that measurements taken during the test can vary from quantitative and objective ones to subjective ones, such as users satisfaction. Also, the observers may ask the participants to be more verbose and explain what they are thinking about at the moment, the actions they are trying to take, and why.

5 participants is a widely agreed-upon number for such kind of evaluations; [Virzi, 1992, Nielsen, 2000] justified this number after finding that they can discover about 85% of the usability issues. However, the findings of [Spool and Schroeder, 2001, Lindgaard and Chattratichart, 2007] , in which only 35% of the issues where discovered by their 5 participants, contradicted with the above suggestion. In summary, it seems that there number of participants is not the only decisive factor here. [Lindgaard and Chattratichart, 2007] concluded that the design of the tasks, their goals, the diversity of the participants and their skills play a role in the number of their usability findings. They also suggested having different sets of tasks carried out by different users for better usability findings. In the end, [Matera et al., 2006] listed the following five requirements for a good task-based usability test:

1. Setting the test goals,

2. Setting the characteristics of the participating sample,

3. Setting the scenarios and tasks to be done,

4. Setting the measurement criteria,

5. Setting the testing environment and needed materials.

2.1.3 Automatic Evaluation

This methods make use of the logs recorded by web-servers. [Levi and Conrad, 1997] noted that the value of this method is that it relies on real user data; however, it can only be applied after the web-site has been launched. [Chi et al., 2000] added that data-mining techniques can be involved here to capture users' behaviour and what information they are looking for.

3 Case Study

In the coming sections we are going evaluate the following websites:

- Venosc, http://www.venosc.com/

- Flaine, http://www.flaine.com/

- Morzine-Avoriaz, http://www.morzine-avoriaz.com/

We begin by conducting heuristic evaluation on the three websites. Then we present our findings regarding the task-based evaluation for the first two sites. Finally, we present our enhanced design for the first website.

3.1 Heuristic Evaluation

We based our study here on the principles of [Nielsen, 2005], as mentioned in section 2.1.1. We also relied on Deniese Pierotti's check-list [1] as a guideline in some cases.

Users visiting Venosc are greeted with a pop-up window in the French language. Closing it and switching the wesite's language results in getting the pop-up window once more. We found the pop-up window confusing to new visitors, and time consuming to recurring ones. In table 2, we list our findings regarding the site.

Violated heuristics	Problem found	Suggested improvements
Home page		
Documentation / Aesthetic	Logo and circular picture on top are taking too much space. Tag-line is not telling the purpose of the site.	Combine the logo and picture, or just use the logo. Change tag-line with a more descriptive one.
Aesthetic	No clear contrast between top menu and the title bar. Users overlook it.	Increase text size and follow the conventions of menus look and feel.
Error prevention	Clicking on picture on top takes user again to homepage and showing pop-up advertisement again	Should be click-able in all pages but homepage.
Mapping / Visibility	There are two versions of the site (summer and winter), and a small button to switch between them, that is not visible. Users cannot easily tell which version they are in now.	Only one version of the site is needed.
Aesthetic	List of events taking too much space and is distracting visitors from their main goal.	Replace it with a way to search for accommodations and/or a list of recommended ones.
Convention	Button-like items in footer-menu are not click-able	Change the style of those items.
Mapping / Recognition rather than Recall	Top menu has too many types of accommodations. Users' choices are mainly based on price, size, location, etc. This confuses new users and returning ones have to remember their choices.	Two (maximum three) options in the menu, combined with a better search facility, will be more useful.
Documentation / Aesthetic	Descriptions for accommodation options is not clear	Better description here. Also remove photos in menu, since they are tiny and not helpful to foreigners.

---

[1]Heuristic Evaluation - A System Checklist, http://www.stcsig.org/usability/topics/articles/he-checklist.html

Violated heuristics	Problem found	Suggested improvements
Rental Properties: "/meubles-en"		
Convention / User control	Logo is not click-able.	Clicking on logo should take users back to home page.
Error prevention	Having a drop-box with different accommodation types again after choosing one type is confusing. Users mistakenly click on it to find themselves in places they didn't intend to go.	Replace the accommodation drop-box with a hierarchical side-menu for better error prevention and flexibility.
Consistency / Mapping	Not selecting the number of beds gets random results. Not even all results nor none.	Make the top option in drop-box be "All accommodations" and return all available results sorted by some criteria.
Mapping	No way to search by number of rooms, price or location	Add advance search options.
Consistency	The zigzag arrow-like search button is confusing.	Either update results upon changes in selection only or use a standard button.
Documentation / Aesthetic	Missing information in search results such as price and location. Some useless information there like surface area.	Better to have structured description with the following information: Price, location, number of beds and rooms, as well as a booking button.
Convention / Error Prevention	The whole area of search results is click-able. a user might click by mistake while copying the phone number for example.	Make the title only click-able or add a hyper-link with the following text: "more details".
Mapping	Some proprietary words such as "2 keys" and "Unclassified" are used in description.	Remove those words from the description.
Flexibility	No way to sort results	Make it possible to sort by price and number of rooms.
Page of one of the apartments		
Error prevention	In some pages there was a drop-down menu on top, that should help users move to other results. Whereas it confuses users making them going to places they did not intend to go	Display relative options at the bottom of the page. Core part of the page should only be dedicated to the apartment the user is checking now.
Flexibility	No online booking available. The only available options are booking by phone, email or being redirected to an external site.	Provide online booking.
Convention	Phone number displayed without country code. The word "Contact" confuses users; whether the phone is for booking, or just contacting site owners.	Change description and add country code, or preferably, just provide online booking.
Aesthetic	Information is not structured. Different fonts does not reflect differences in information importance.	Present information in more precise and ordered way.

Violated heuristics	Problem found	Suggested improvements
Recognition rather than recall	The map without landmarks on it is not useful to foreigner. They will need to compare it to an external map of the city.	Place important landmarks on the map. Provide lo-fi alternative for users with slow internet connection.

Table 2: Heuristic Evaluation for Venosc

Flaine shares many of the usability issues with Venosc. Please refer to table 2 for more details about the following issues found in the two sites: Pop-up window, winter-summer versions of the website, click-able logo in homepage, room description, city map and the lack of online booking. In addition to these issues, Flaine suffers from the following usability issues:

- Too many photos and heavy multimedia content in the homepage. (Aesthetic and Minimalist Design)

- Translation for some languages is supported via Google translate, which doesn't work with text in images and sometimes doesn't work altogether. (Error prevention)

- In the "accommodation" page, text in different colours within the description resembles hyper-links. Similarly, in one of the offered properties, a list of recommended offers is shown yet none is click-able (Consistency)

- No integration between the website and the ones users are redirected to when booking an apartment. They have to recall all the values they entered in the first website to start a new search query in the second one from scratch. (Recognition rather than recall)

- The site is using proprietary French names for the properties, without proper description. (Mapping and "Help and Documentation")

Moving to the third website. Morzine-Avoriaz is much better than the first two from a usability point of view. It has a search engine and the descriptions for the results are in more details. Nevertheless, we have found the following usability issues in it:

Violated heuristics	Problem found	Suggested improvements
Home page		
Mapping	Like the two other sites, it also has different versions for summer and winter	Combine the two, and customize the search results based on the booking date.
Mapping	In the search box, there is an option to limit search to special offers only. Normal users have no clue what are special offers.	Instead of having this option in search box, mark special offers in search results.
Mapping	A *shopping basket* is more suitable for retail sites where users can buy more than one item.	Just use a "book now" button for users to click after making their choice.

Violated heuristics	Problem found	Suggested improvements
Aesthetic	Differences in font size and the location of horizontal bars does not reflect content importance or group related content together.	Use horizontal bars to separate unrelated content. Stick to few font sizes and colours. Only use different fonts to deliver specific message, such as content importance or the presence of a hyper-link.
Search results		
Mapping	Proprietary text in search results. For example: *(00610041)* and *(n° 690)*. Similarly 3* instead of "3 stars".	Use common language and symbols understood by all visitors.
Flexibility	Search box is set to the advanced-mode by default.	The default mode should be the one with just essential inputs.
General		
Documentation	Some parts of the website are not translated.	Make sure that all displayed content should be in the selected language only.
Convention / Documentation	Two maps are provided. One is just a satellite image while the other is a city map with landmarks on it. The former is useless to users not aware of the city, while the latter is not flexible and doesn't provide ways for zooming and panning.	Consolidate the two maps into one richer map.

Table 3: Heuristic Evaluation for Morzine

3.2   Task Based Evaluation

We asked 5 persons to visit *Venosc* and *Flaine*, and complete the task mentioned in the assignment sheet. The participants were requested to *think aloud* while doing the task [Matera et al., 2006]. We kept records with the number of clicks it took them to finish the task, and whether they succeeded or not. We did not record the task completion time, as we believe that thinking aloud may result in completing the task in more time than in normal cases.

As mentioned earlier, we noticed visitors of the two websites are greeted with a pop-up window. Changing the language in the two sites resulted in getting the pop-up once more with in the new language. We concluded that 3 wasted clicks are common in the beginning of all the tasks done.

We noticed that in all cases the users had to either send an email or visit an external website to make their booking. Since we cannot tell how reliable the email-booking would be, and since testing all other sites is beyond the scope of this report, we needed to have a third outcome other than success and failure. We will call it *uncertain success*. All of the 10 tasks, *"5 users ∗ 2 websites"*, ended with uncertain success.

3 out of 5 visitors to Flaine, needed 8 clicks before being redirected to an external website, where they were supposed to start the search again from scratch. The two other participants needed between 8 and 12 clicks to finish their task. All the 5 visitors of Venosc needed 9 or more clicks before being forced to either visit an external website or send an email with their requirements.

2 of our sample commented that the lack of online booking makes those website as good as

doing a basic web search in "Google" or "Yellow Pages". The task-based evaluation confirmed our heuristic evaluation findings regarding the pop-up windows and the lack of search options. None of the participants noticed the winter/summer versions of the websites. Only 1 out of the 5 participants noticed the "good deals" link in Flaine. 3 out of 5 visitors for Flaine did not look any further after visiting 2 choices. Our interpretation: although Vensoc was not giving detailed description for the available properties, it was slightly more helpful than the lack of description altogether in Flaine.

Many of the remaining comments made by the participants are already mentioned in the heuristic evaluation. Hence, and for the sake of brevity, we will end our task-based evaluation report here.

3.3 Design

Using the task description and the feedback provided by our users, we were able construct the following use-cases in order to build an enhanced design afterwards:

Booking a vacation - Use-case 1		
User Intention		*System Response*
	←	List supported languages
Choose website language	→	
	←	Change website language
	←	Present search box
Enter search query (number of rooms/beds, price and date ranges)	→	
	←	Display results
Order results by price, ascendingly	→	
	←	Display ordered results
Choose one of the displayed results	→	
	←	Display more details about chosen result and booking options
Make a booking	→	
	←	Display booking confirmation

For some users, having a list of recommended accommodations might be more convenient than searching. It also helps new visitors to understand the main purpose of the site. Hence, we should also have the following use-case.

Booking a vacation - Use-case 1a		
User Intention		*System Response*
	←	Display recommended apartments
Choose one of the displayed results	→	
	←	Display more details about chosen result and booking options
Make a booking	→	
	←	Display booking confirmation

Based on the use-cases, we ended with the homepage design shown in figure 1. The page should have the following two core components:

- List of recommended accommodations: Each item has a brief description, nearby landmarks, price, number of rooms/beds and a way for advanced users to view its location on a map in a new window.

  - The map should have the important landmarks in the city marked on it. Users should also be able to zoom and pan the map. We recommend using a customized 3_rd party map with the important landmarks in the city on it; however the actual implementation is beyond the scope of this report.

- Search box: Visitors can search for accommodations based on number of travellers and arrival/departure dates.

  - Advanced search options: This should include options such as price range, and whether a hotel room or a self-catering accommodation is required.

  - We have noticed that the conceptual model for a traveller tends to be aligned more with number of persons (adults and children) rather than number of rooms. Hence, we found it easier to ask for number of persons, while keeping the users aware of the maximum number of persons per room. Additionally, when it comes to the search results, we should display the two pieces of information.

  - The original site had two versions, one for summer holidays and one for winter ones. We decided to ignore this confusing option and base our results on the users' arrival and departure dates only.
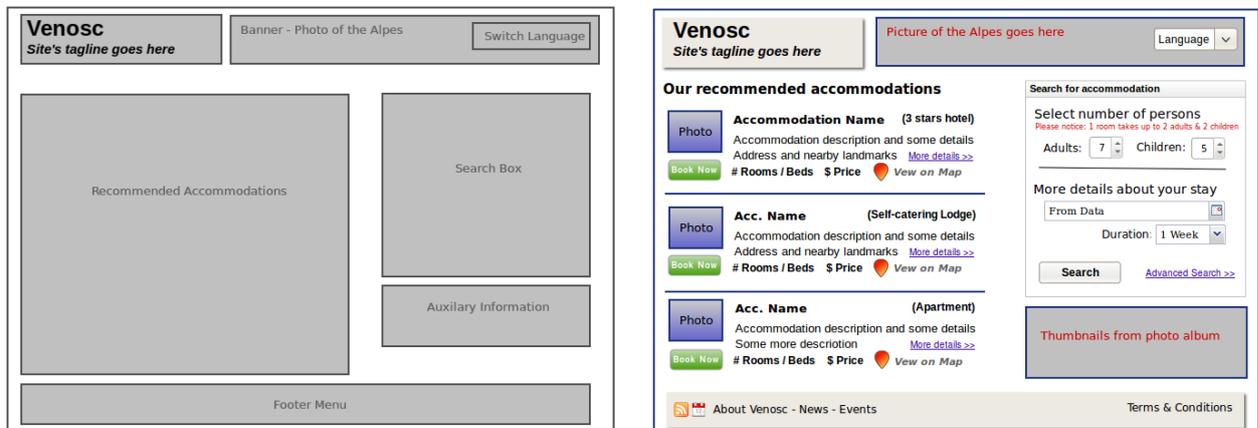

Figure1: Venosc Homepage - Low and Medium-Fi Design

We believe it will be more convenient to have the search results returned in a similar page to the homepage. The results can be displayed in a similar fashion to how the "recommended accommodations" are listed. Alternatively, the content of the homepage can be auto populated with the new results upon submitting the search criteria with no need to redirect users to a separate page. The only difference we suggest for the new results, is to have a drop-down box on top of the results for users to be able to sort results by price or size.

We asked one of our task-based sample to additionally test the website with search options[2]. We then noticed that when no results are returned, some users do not think of refining their search

---

[2]Morzine-Avoriaz

query. Therefore, when no results are found for users' queries, we suggest showing a warning message on the top of the list, then return the closest results to their query afterwards.

After users click on the booking button or follow the "more details" link, they should be given a page similar to the one in figure 2.
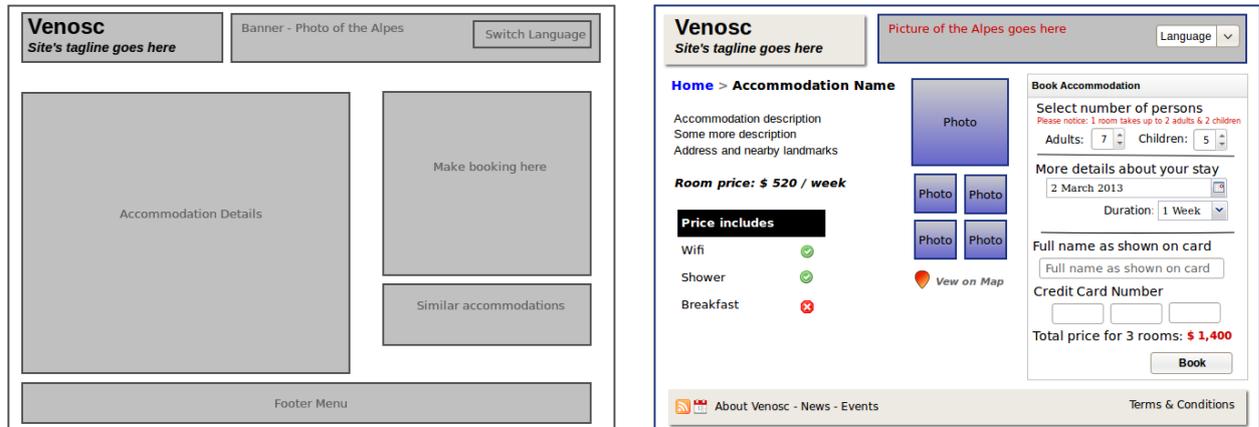

Figure2: Venosc - Apartment detailed page - Low and Medium-Fi Design

For some accommodation types, user have to re-enter the number of beds when booking. Therefore, the system should memorise the values the users entered while searching and auto-fill those entries for them. System should print the total price in bold for users to be able to confirm it before paying.

In the medium-fi design, we revised our lo-fi design and decided not to include the list of similar offers shown the in lo-fi design. We though it might be confusing here.

Finally, as mentioned earlier, the presented design should be subject to iterative evaluations and improvements in the future.

References

[Bevan, 1997] Bevan, N. (1997). Usability issues in web site design. *Advances in Human Factors Ergonomics*, 21:803–806.

[Chi et al., 2000] Chi, E., Pirolli, P., and Pitkow, J. (2000). The scent of a site: A system for analyzing and predicting information scent, usage, and usability of a web site. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 161–168. ACM.

[Hertzum and Jacobsen, 2001] Hertzum, M. and Jacobsen, N. (2001). The evaluator effect: A chilling fact about usability evaluation methods. *International Journal of Human-Computer Interaction*, 13(4):421–443.

[Hollingsed and Novick, 2007] Hollingsed, T. and Novick, D. (2007). Usability inspection methods after 15 years of research and practice. In *Proceedings of the 25th annual ACM international conference on Design of communication*, pages 249–255. ACM.

[Levi and Conrad, 1997] Levi, M. and Conrad, F. (1997). Usability testing of world wide web sites. In *Conference on Human Factors in Computing Systems: CHI'97 extended abstracts on Human factors in computing systems: looking to the future*, volume 22, pages 227–227.

[Lindgaard and Chattratichart, 2007] Lindgaard, G. and Chattratichart, J. (2007). Usability testing: what have we overlooked? In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 1415–1424. ACM.

[Mariage et al., 2005] Mariage, C., Vanderdonckt, J., and Pribeanu, C. (2005). State of the art of web usability guidelines. *The Handbook of Human Factors in Web Design*, pages 688–700.

[Matera et al., 2006] Matera, M., Rizzo, F., and Carughi, G. (2006). Web usability: Principles and evaluation methods. *Web engineering*, pages 143–180.

[Nielsen, 1994] Nielsen, J. (1994). Usability inspection methods. In *Conference companion on Human factors in computing systems*, pages 413–414. ACM.

[Nielsen, 1995] Nielsen, J. (1995). *Multimedia and hypertext: the Internet and beyond*. Morgan Kaufmann.

[Nielsen, 2000] Nielsen, J. (2000). Why you only need to test with 5 users.

[Nielsen, 2005] Nielsen, J. (2005). Ten usability heuristics.

[Nielsen and Hackos, 1993] Nielsen, J. and Hackos, J. (1993). *Usability engineering*, volume 125184069. Academic press San Diego.

[Rusu et al., 2011] Rusu, C., Roncagliolo, S., Rusu, V., and Collazos, C. (2011). A methodology to establish usability heuristics. In *ACHI 2011, The Fourth International Conference on Advances in Computer-Human Interactions*, pages 59–62.

[Spool and Schroeder, 2001] Spool, J. and Schroeder, W. (2001). Testing web sites: Five users is nowhere near enough. In *CHI'01 extended abstracts on Human factors in computing systems*, pages 285–286. ACM.

[Virzi, 1992] Virzi, R. (1992). Refining the test phase of usability evaluation: How many subjects is enough? *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34(4):457–468.