A Thinking Aloud Test of the TreeTest Web App for Testing Information Hierarchies

Elias Raphael Wolfgang Doppelreiter

A Thinking Aloud Test of the TreeTest Web App for Testing Information Hierarchies

Elias Raphael Wolfgang Doppelreiter

Bachelor's Thesis

to achieve the university degree of

Bachelor of Science

Bachelor's Degree Programme: Software Engineering and Management

submitted to

Graz University of Technology

Supervisor

Ao.Univ.-Prof. Dr. Keith Andrews
Institute of Interactive Systems and Data Science (ISDS)

Graz, 11 Sep 2020

Ein Thinking Aloud Test der Web-App TreeTest zur Evaluierung von Informationshierarchien

Elias Raphael Wolfgang Doppelreiter

Bachelorarbeit

für den akademischen Grad

Bachelor of Science

Bachelorstudium: Softwareentwicklung und Management

an der

Technischen Universität Graz

Begutachter

Ao.Univ.-Prof. Dr. Keith Andrews
Institute of Interactive Systems and Data Science (ISDS)

Graz, 11 Sep 2020

Diese Arbeit ist in englischer Sprache verfasst.

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The document uploaded to TUGRAZonline is identical to the present thesis.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Dokument ist mit der vorliegenden Arbeit identisch.

Date/Datum

Signature/Unterschrift

Abstract

This thesis describes an evaluation of a web application called TreeTest [Mehic 2019], which is used for testing information hierarchies. The web application was evaluated using the thinking aloud test method.

Tree testing is a technique used to evaluate the structure of information hierarchies, ensuring that all elements in the information hierarchy are easy to find. To perform a tree test, test users are asked to find different elements in the information hierarchy. The interactions of the test users are recorded and analysed using various criteria.

Thinking aloud testing is a method for evaluating the usability of a user interface. Representative test users are given various tasks typical for the domain and are encouraged to think out loud while they work. Each test session is recorded and later analysed to create a list of potential issues and problems in the user interface. By asking users to think out loud, it is possible to gain insight into their thoughts and decision-making processes.

Kurzfassung

Diese Arbeit beschreibt die Evaluierung der Webanwendung namens TreeTest [Mehic 2019], welche verwendet wird, um Informationshierarchien zu testen. Die Webanwendung wurde mit der Thinking Aloud Testmethode evaluiert.

Tree Testing ist eine Technik, die verwendet wird, um die Struktur von Informationshierarchien zu bewerten, um sicherzustellen, dass alle Elemente der Informationshierarchie leicht zu finden sind. Um einen Tree Test durchzuführen, werden Testbenutzer gebeten, verschiedene Elemente in der Informationshierarchie zu finden. Die Interaktionen der Testbenutzer werden aufgezeichnet und anhand verschiedener Kriterien analysiert.

Thinking Aloud Testing ist eine Methode zur Bewertung der Benutzerfreundlichkeit einer Benutzeroberfläche. Repräsentative Testbenutzer erhalten verschiedene, für die Domäne typische Aufgaben gestellt und werden gebeten, während der Arbeit laut zu denken. Jede Testsitzung wird aufgezeichnet und später analysiert, um eine Liste potenzieller Probleme der Benutzeroberfläche zu erstellen. Durch die Aufforderung an die Testbenutzer, laut zu denken, ist es möglich, Einblick in ihre Gedanken und Entscheidungsprozesse zu erhalten.

Contents

List of Figures

List of Tables

Acknowledgements

I would especially like to thank my advisor Keith Andrews, who has supported me in all stages of this thesis. A great help was the fact that Keith always spontaneously took time for me. I would especially like to thank him for the many hours he invested in correcting the drafts versions of this thesis.

Special thanks to Ajdin Mehic, the developer of the web application TreeTest. He allowed me to test his web application and was always available for questions.

I would also like to thank my fellow students and friends who participated in the usability test. Without them, this thesis would not have been possible.

Last but not least, I would like to thank my family, who always supported me during my academic education.

Elias Doppelreiter
Graz, Austria, 11 Sep 2020

Credits

I would like to thank the following individuals and organisations for permission to use their material:

- The thesis was written using Keith Andrews' skeleton thesis [Andrews 2019b].
- Various materials for thinking aloud testing from Keith Andrews' HCI course were used for this thesis [Andrews 2019a].

Chapter 1

Introduction

This thesis describes a thinking aloud usability test of a web application called TreeTest [Mehic 2019], which is used for testing information hierarchies. The first part of the thesis (Chapters 2 to 4) discuss the background of information hierarchies and usability testing methods. Chapter 2 covers the application areas of information hierarchies, as well as the concept of tree testing, which is used to evaluate these structures. Chapter 3 introduces various web applications for testing information hierarchies, and in particular, the web application TreeTest developed by Ajdin Mehic [Mehic 2019]. Chapter 4 describes the different concepts used for testing user interfaces. The difference between formative and summative usability evaluation is explained, and two formative usability evaluation methods, heuristic evaluation and thinking aloud testing, are described in more detail.

The second part of this thesis (Chapters 5 and 6) describes the thinking aloud test of the TreeTest web application. Chapter 5 describes the planning and preparations for the usability study, including the test users, tasks, and the environment used. The results of the thinking aloud tests are presented and discussed in Chapter 6.

Appendix A contains the background questionnaires completed by the test users. Appendix B contains the task slips given to the different groups of test users. Appendix C contains the feedback questionnaires completed by the test users.

Chapter 2

Tree Testing

> *“ Tree testing is a powerful method in evaluating the hierarchical structure of a particular design. It ventures into quantitative research territory and can generate a large amount of data. ”*
>
> [Sam Yuan, The Things No One Tells You About Tree Testing [Yuan 2019].]

2.1 Information Hierarchies

Information hierarchies are frequently used on web sites and in applications as a central navigation element to link from the main page or screen to subpages or subscreens. Information hierarchies are structured like trees, which is why they are also called trees. Tree testing is a method for evaluating these structures. The goal is to determine how intuitive such a structure is and therefore, how easily individual elements in a tree can be found [Babich 2020].

Information hierarchies consist of nodes and their labels. Nodes can either be *inner* nodes which have further child nodes, or *leaf* nodes which do not. In a file system, folders correspond to inner nodes and files correspond to leaf nodes. For example, Figure 2.1 shows the hierarchical information structure of Graz University of Technology’s external web site [TUG 2020]. The node Studying and Teaching has one leaf and four inner nodes.

Labels must succinctly summarise the content of the nodes they are assigned to, so that users can have some intuition about what the node contains (a concept known as *information scent*) before opening it. If this is not the case, users of the information hierarchy may become frustrated and stop using the application or refrain from revisiting the web site in future [Schroeder 2018].

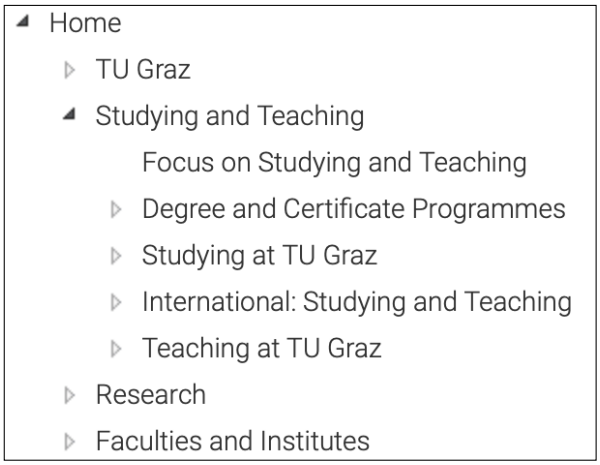

Figure 2.1: The information hierarchy of Graz University of Technology's external web site [TUG 2020] [Screenshot taken by the author of this thesis.]

2.2 Testing Information Hierarchies (Tree Testing)

Tree testing is a technique used to evaluate the structure of an information hierarchy, ensuring that the items in the hierarchy are easy to find from a user perspective. To perform a tree test, test users are asked to find different elements in the information hierarchy. The interactions of the test users are recorded and analysed [O'Brien 2020].

Tree testing is used to test how well a hierarchical information structure is set up and therefore, how well it can be used in a real scenario. This method of testing hierarchical information structures allows information structures to be evaluated before a web site or application is developed. The only things required for tree testing is the hierarchical information structure itself and corresponding tasks for the test users. Tasks specify the items that the test users should find in the tree [Whitenton 2017]. A significant advantage of tree testing is that the test users are given tasks that would also occur in a real-world scenario. This real-world approach makes it possible to obtain information that cannot be captured by other testing methods [Ratcliff 2019].

2.3 Tasks for Tree Testing

O'Brien [2020] discusses the creation and selection of tasks for tree testing. According to O'Brien [2020], choosing the right tasks for a tree test is essential. Tasks should be created to cover the following:

- *Common tasks*: Items users search for most commonly.
- *Critical tasks*: Items that are not often searched for by users, but are critical to find fast in certain situations.
- *Suspect tasks*: Items which are suspected of being difficult to find.

Various factors influence the number of tasks for a tree test. Each participant should not be given more than around 8 to 10 tasks. For a large tree structure, it may be necessary to create many more tasks but to spread them over a larger set of test users, and potentially randomise the assignment and ordering of tasks [O'Brien 2020]. When writing tasks, it is better if the names of the labels do not appear in the text of the task definition. One way to avoid this is to use scenario descriptions [Whitenton 2017].

2.4 Analysis of Results

This section is based on the material from O'Brien [2020]. The purpose of a tree test is to find out if and where the hierarchical information structure has problems in order to have the opportunity to improve the structure. Various criteria can be considered when analysing the results. The most important are discussed here, O'Brien [2020] contains a much fuller treatment.

2.4.1 Task Success

The task success rate is the most important factor for the analysis of the tree test results. This factor expresses the proportion of test users who found the correct answer for a particular task. This value is often expressed as a percentage. If the task success rate for a particular task is higher than around 65%, it can be concluded that the information hierarchy is performing well for this task, and only minor changes are required. If the task success rate for a task is less than around 50%, then the information hierarchy has not performed satisfactorily for this task and further analysis and tweaking is required.

2.4.2 First Click

The analysis of the first click is used to determine whether the top level of the information hierarchy is well structured. The top level has a significant influence on the task success rate of a particular task. It represents the starting point for each task and, therefore, the first navigation decision that users have to make.

2.4.3 Task Directness

The directness rate indicates the proportion of test users who navigated directly to the correct answer without having to backtrack. By analysing directness, it is possible to determine whether the title of a node represents its underlying content. Backtracking occurs when a user selects a node based on its title, but is disappointed with its contents after selecting it. If the directness of a task is only captured as a binary measure (deviated or not) for each user, then an average directness rate of 70% or higher is a good result. It would also be possible to count the number of deviations each user made for the task.

2.4.4 Task Completion Time

The average time taken to complete a task gives some indication of how complex the task was, but is highly dependents on how deep the correct answer lies in the tree. It can be more revealing to look for places where users slowed down, i.e. spent longer before making their next choice (click). This might indicate a lack of information scent.

2.4.5 Task Abandonment

Occasionally, test users may not be able to complete a task and decide to skip it. The abandonment rate expresses the proportion of users who decided to give up on a task. An abandonment rate of 10% or higher can indicate a problem with the information hierarchy. For more detailed insight, it can be more revealing to look for the places where users abandoned the task.

Chapter 3

Tree Testing Tools

This chapter presents three different tree testing tools: Treejack [OW 2020b], UXtweak TreeTest [UXtweak 2020a], and TreeTest [Mehic 2019]. All three tools are web-based applications. The final tool discussed, TreeTest developed by Mehic [2019], is the tool which was evaluated in a thinking aloud test for this thesis.

3.1 TreeJack

TreeJack is a well-known commercial web application for testing information hierarchies [OW 2020b], provided by Optimal Workshop [OW 2020a]. Besides TreeJack, Optimal Workshop offers a suite of other tools for user research and usability testing.

3.1.1 Creating a Study

The first step in creating a tree test is to select Treejack from a list of study types and to click on Start tree testing, as shown in Figure 3.1. Next, the study name, study link and study languages are defined, as shown in Figure 3.2. In the paid version, a password can also be assigned, and closing rules can be specified. In the following step, the tree can be defined by either creating the tree with the tool or by importing an existing tree, as shown in Figure 3.3. In the following step, tasks can be created by defining a question and selecting the answer, as can be seen in Figure 3.4. Next, the text of various messages and notifications can be configured, as shown in Figure 3.5. These are displayed to test users as various points in the test process. Finally, it is possible to define questions for pre-study and post-study questionnaires, as shown in Figure 3.6. In the paid version of Treejack, it is also possible to customise the design of the study, by specifying a logo and a colour scheme.

3.1.2 Participating in a Study

Participants in a tree test with Treejack are welcomed with the message shown in Figure 3.7. They are then given the tasks to be completed, one at a time. Figure 3.8 shows a typical task as presented to a test participant.

3.1.3 Results of a Study

Teejack provides a very clear and graphically appealing overview of the study results, as shown in Figure 3.9. In addition to summary statistics for the whole study, the results for each individual task are also provided, as shown in Figure 3.10. These include a so-called *pietree* for each task, which displays a visual overview of the paths taken by users in search of the correct answer to the task, as shown in

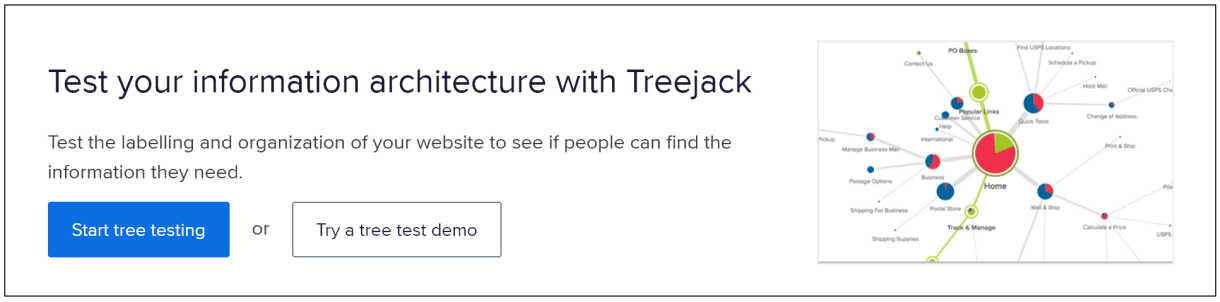

Figure 3.1: Treejack: Selecting the type of study. [Screenshot taken by the author of this thesis.]

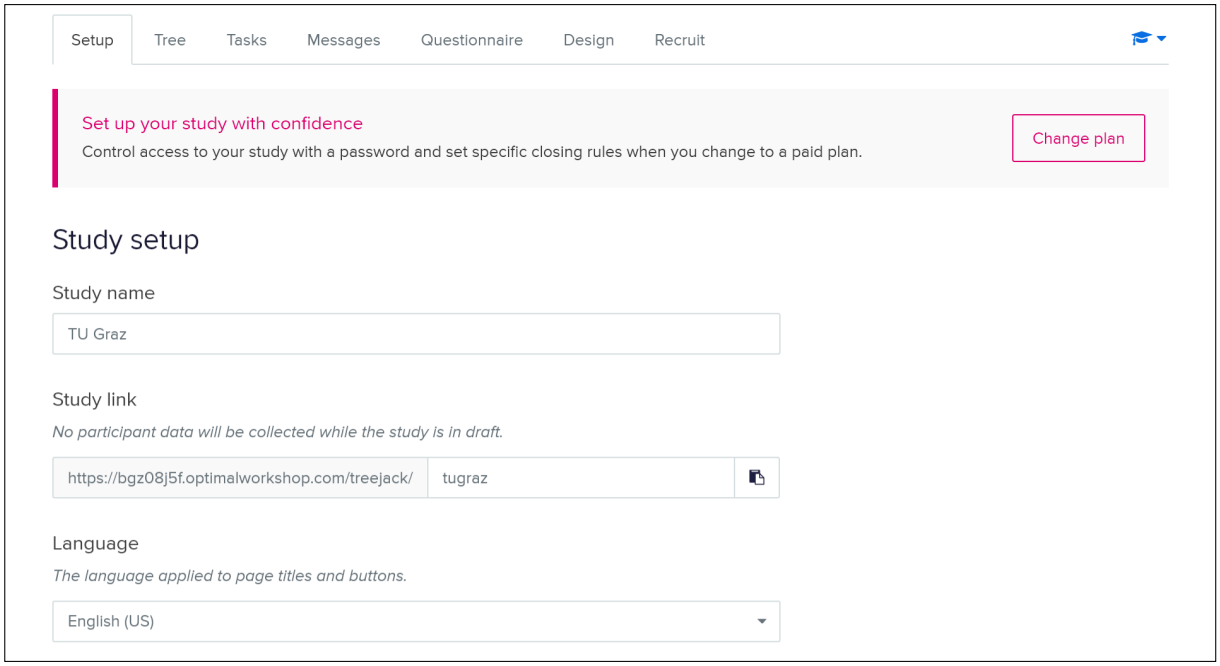

Figure 3.2: Treejack: Settings for a study specified on the Study Setup screen. [Screenshot taken by the author of this thesis.]

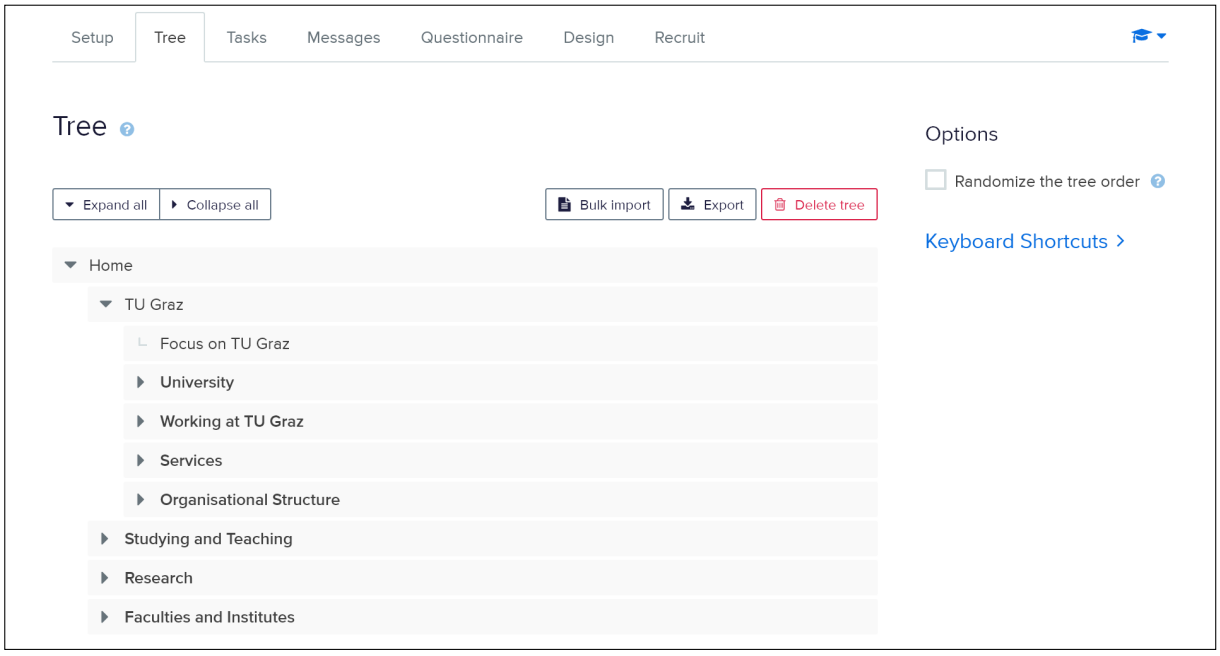

Figure 3.3: Treejack: Defining the tree for a study. [Screenshot taken by the author of this thesis.]

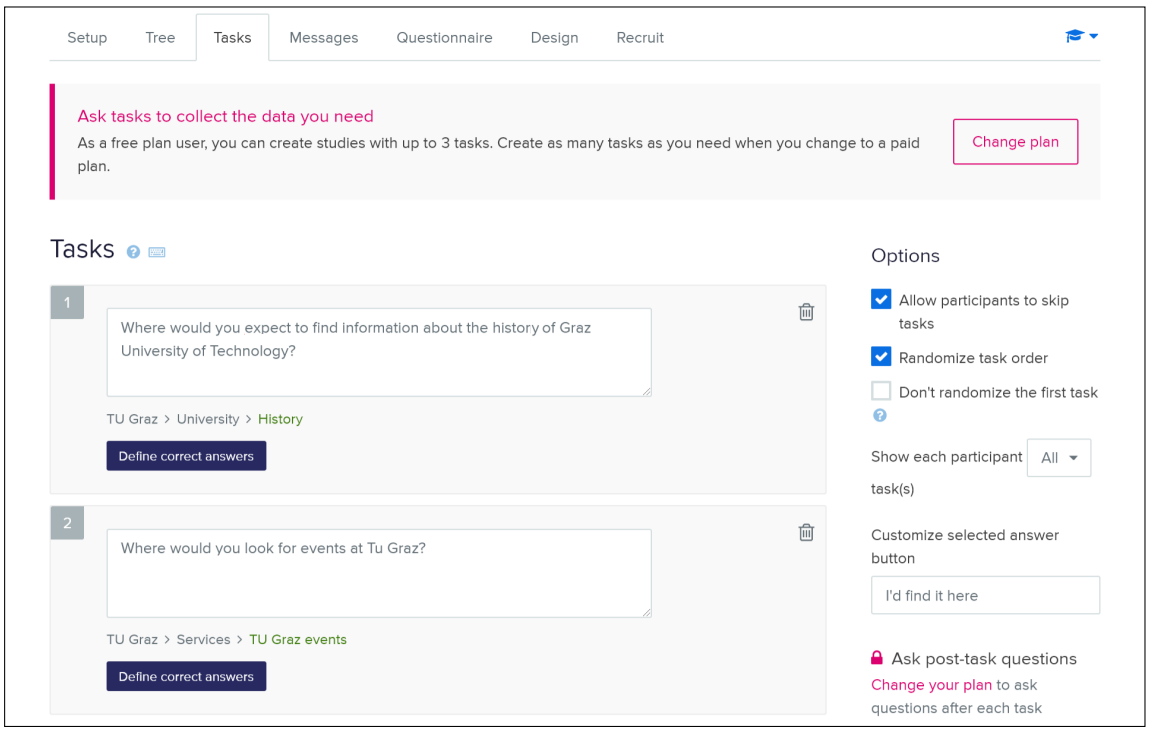

Figure 3.4: Treejack: Defining tasks for a study. [Screenshot taken by the author of this thesis.]

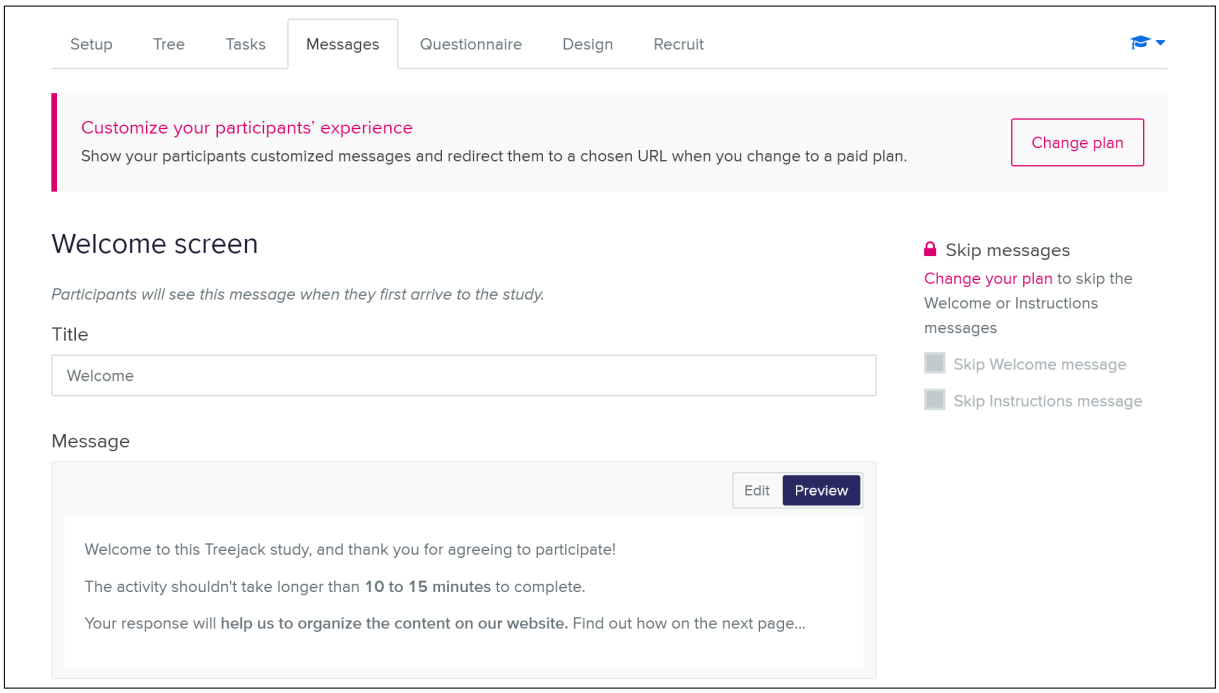

Figure 3.5: Treejack: Configuring customised messages for a study. [Screenshot taken by the author of this thesis.]

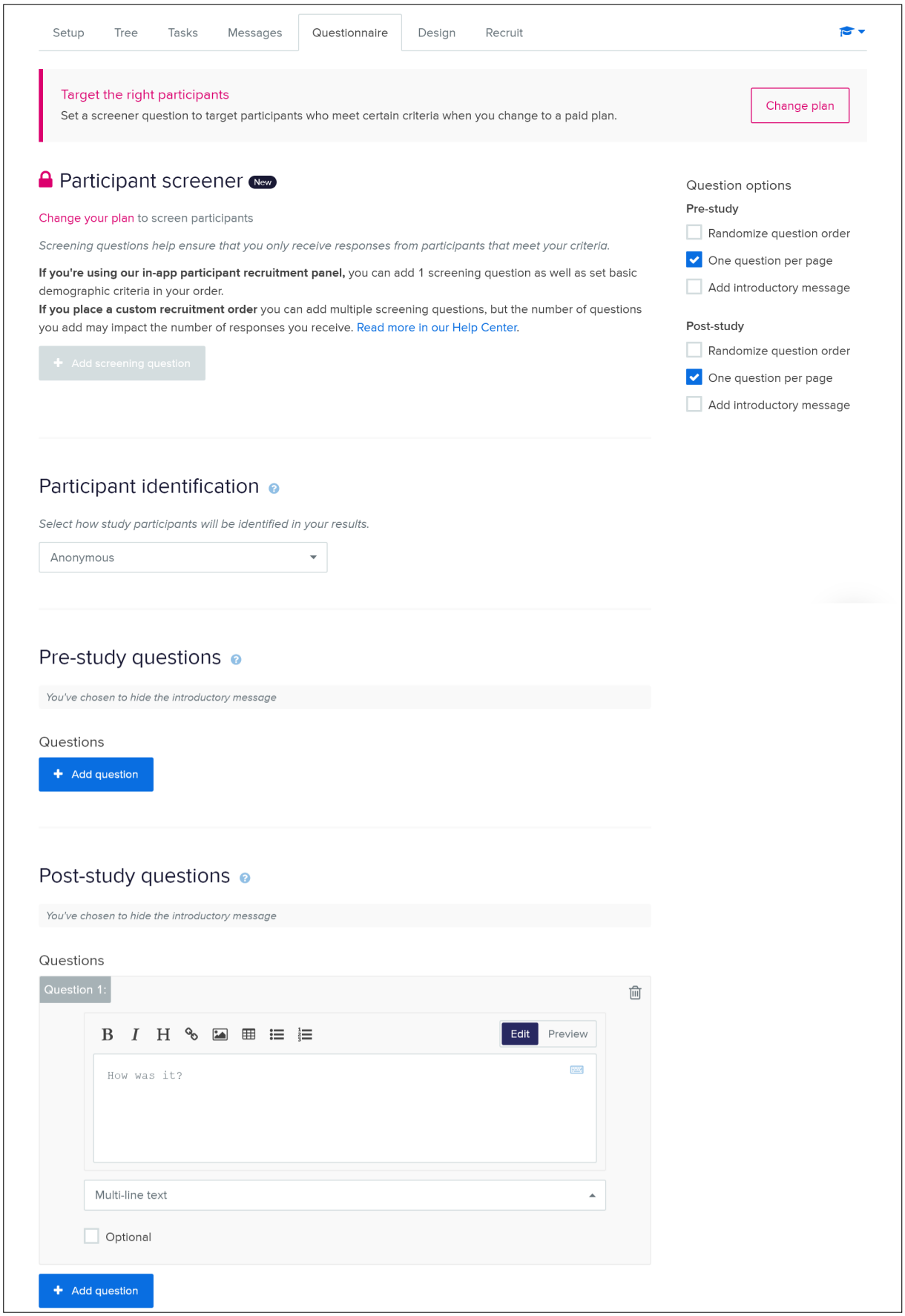

Figure 3.6: Treejack: Defining questionnaires for a study. [Screenshot taken by the author of this thesis.]

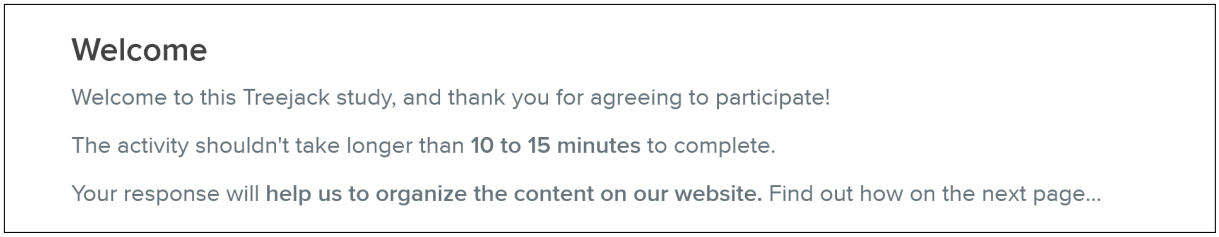

Figure 3.7: Treejack: The welcome message displayed to a participant in a study. [Screenshot taken by the author of this thesis.]

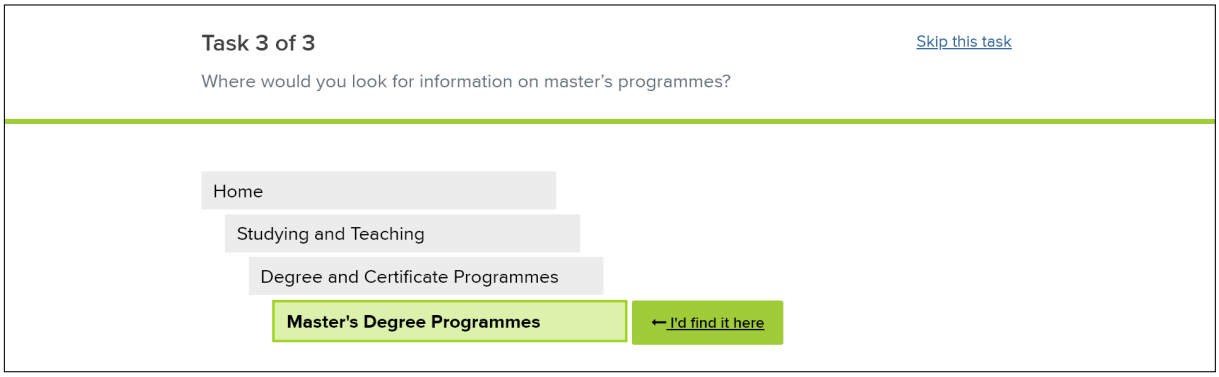

Figure 3.8: Treejack: A typical task displayed to a participant in a study. [Screenshot taken by the author of this thesis.]

Figure 3.11. Additionally, a table of destinations shows an overview of the destinations (answers) chosen by study participants for each task, as shown in Figure 3.12.

Optimal Workshop offers different versions of their usability testing tools. The free version, unfortunately, does not contain all features. For example, it is not possible to create more than three tasks or to analyse the results of more than ten participants. Paid versions start at 166$ per year [OW 2020a].

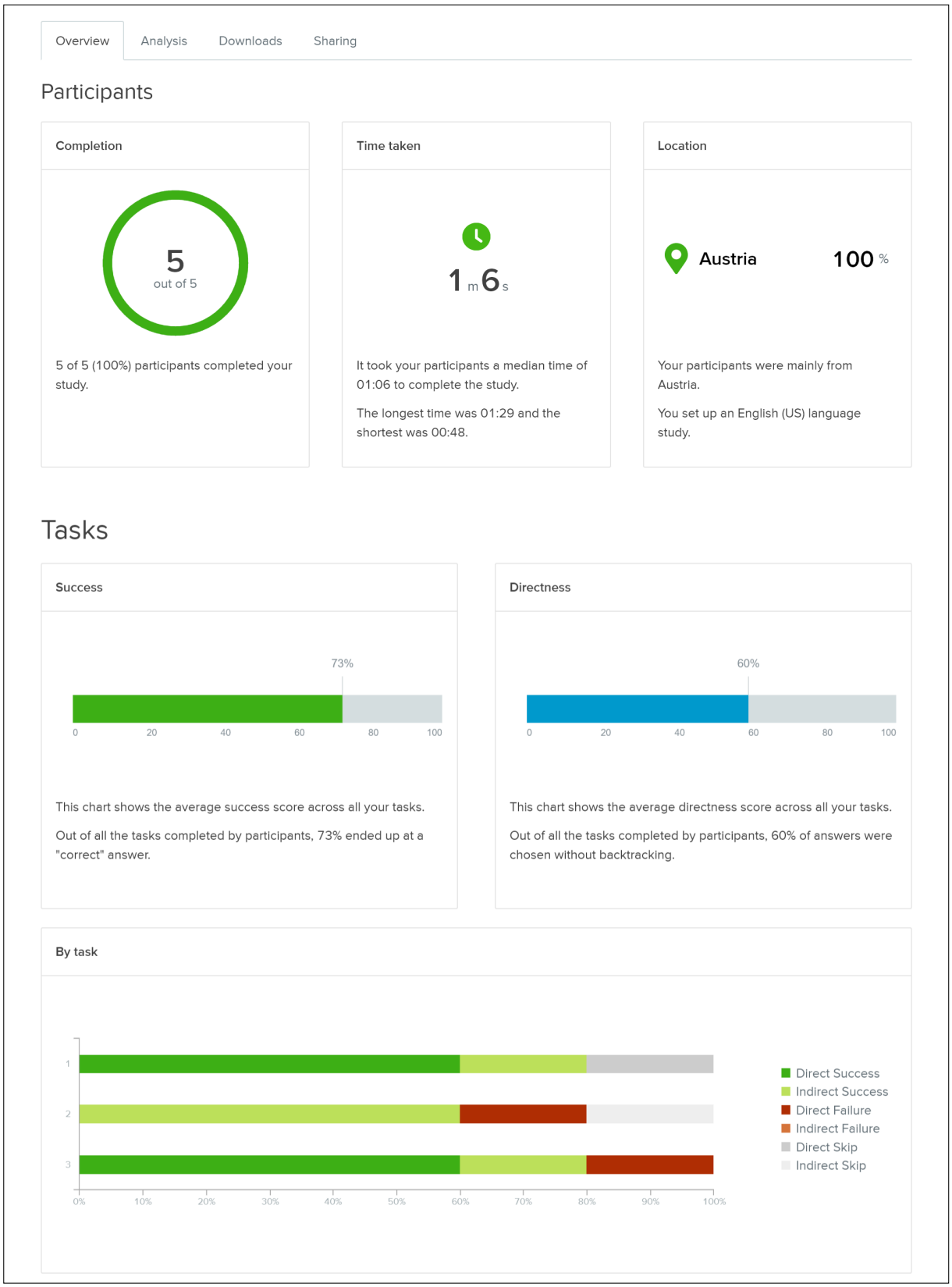

Figure 3.9: Treejack: The overview page of the study results. [Screenshot taken by the author of this thesis.]

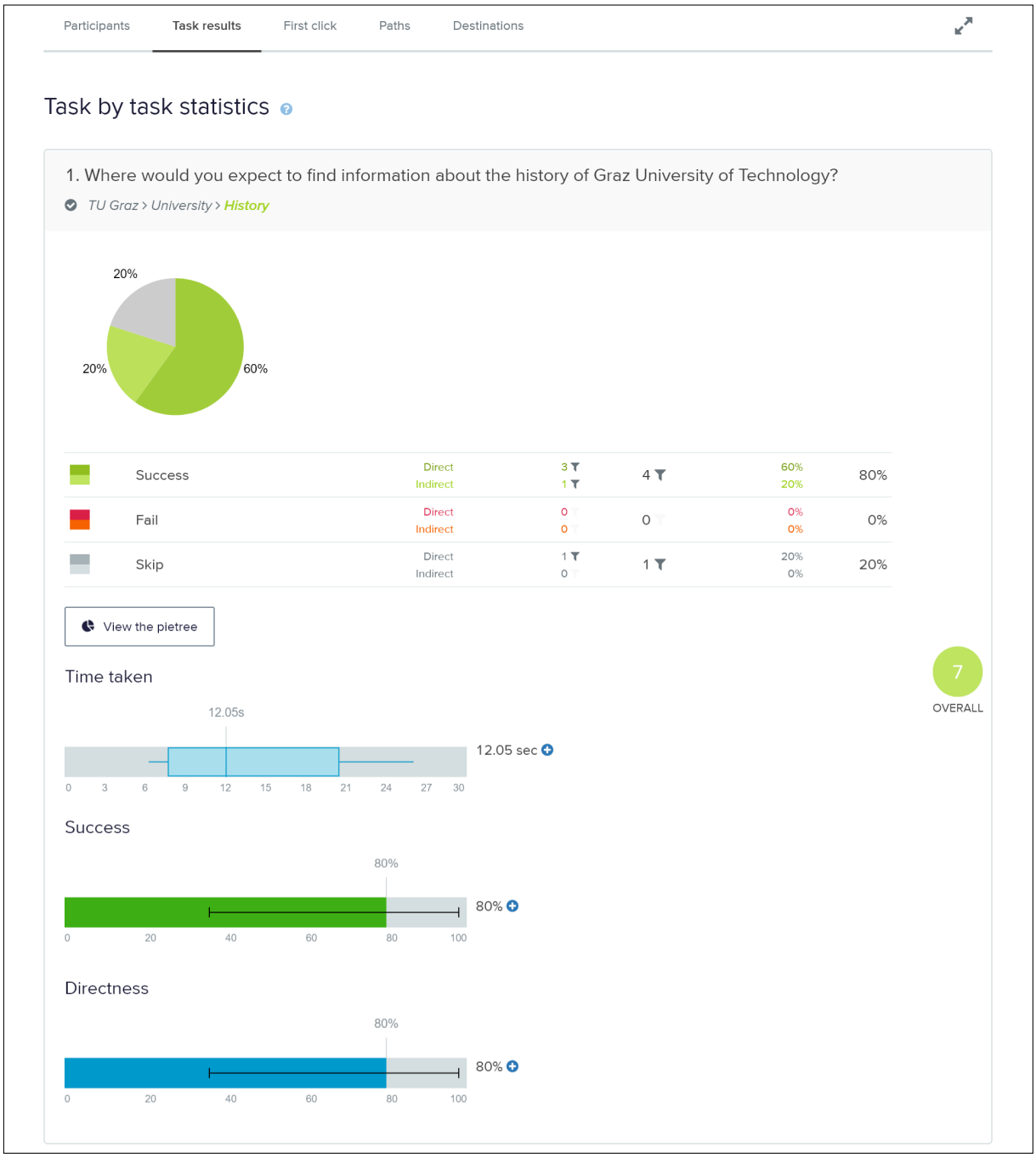

Figure 3.10: Treejack: Study results for an individual task. [Screenshot taken by the author of this thesis.]

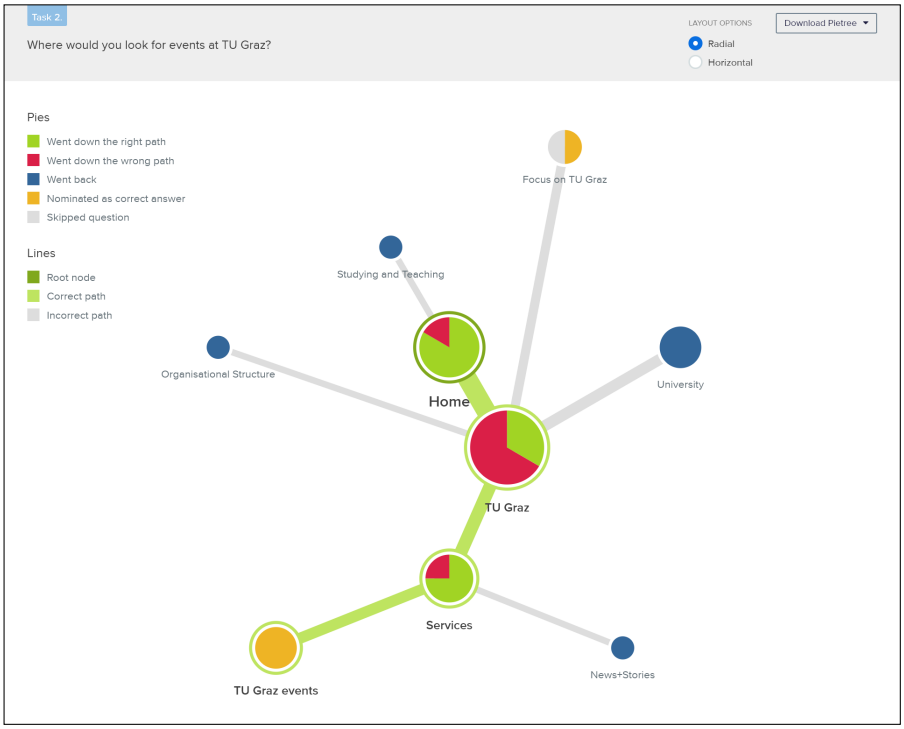

Figure 3.11: Treejack: A pietree for a task shows the proportion of users who followed particular paths through the tree while looking for an answer to the task. [Screenshot taken by the author of this thesis.]

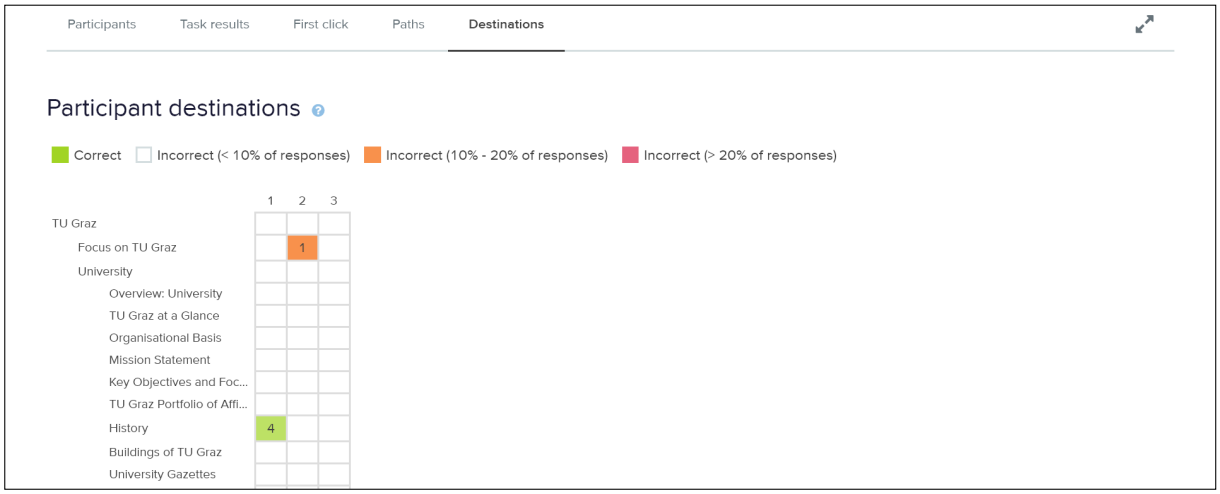

Figure 3.12: Treejack: The destinations table gives an overview of the answers (destinations) selected by study participants for each task. [Screenshot taken by the author of this thesis.]

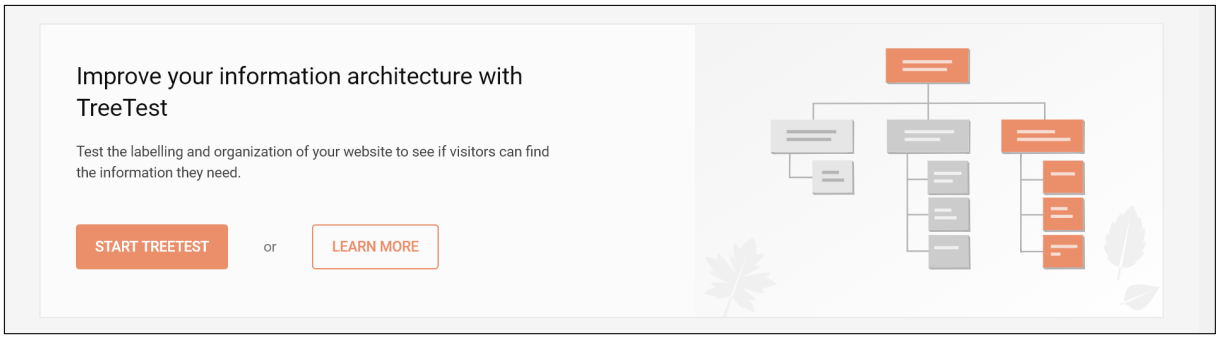

Figure 3.13: UXtweak: Selecting the type of study. [Screenshot taken by the author of this thesis.]

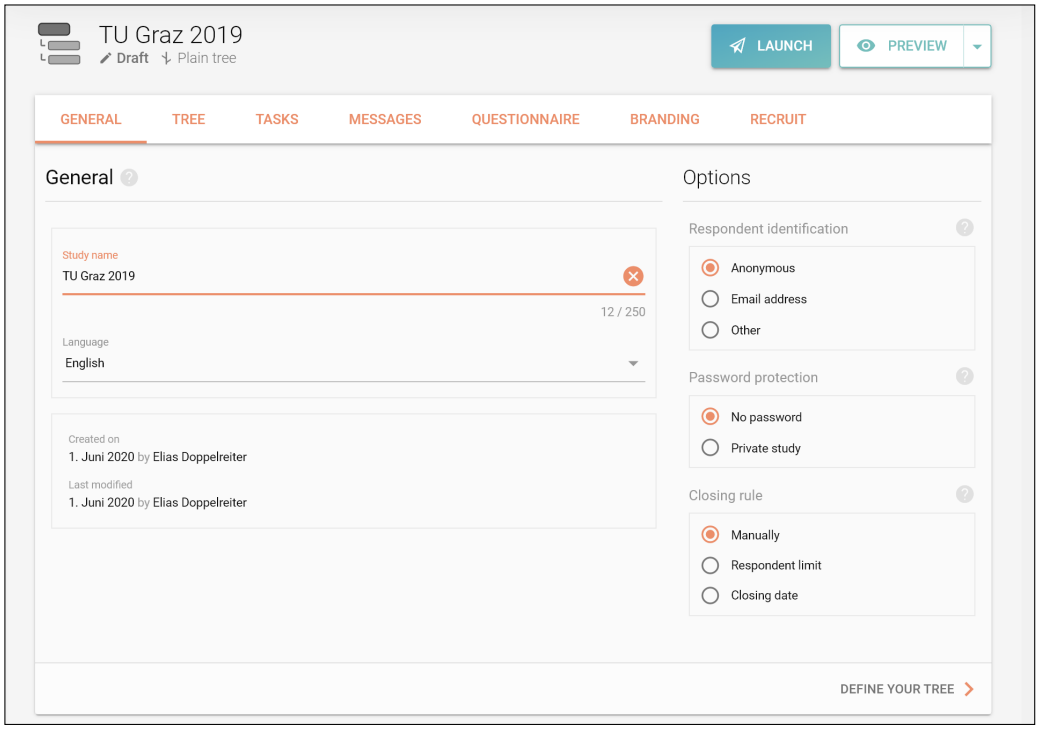

Figure 3.14: UXtweak: Configuring settings for a study. [Screenshot taken by the author of this thesis.]

3.2 UXtweak TreeTest

UXtweak Research Group is another commercial company specialised in user research and usability testing [UXtweak 2020b]. They offer various online tools, including UXtweak TreeTest for tree testing [UXtweak 2020a].

3.2.1 Creating a Study

The first step in creating a tree test is to select the study type Tree Test from a list of different study types, as shown in Figure 3.13. Then, the basic study settings can be configured, including the study name and language, as shown in Figure 3.14. The next step is to create or import the tree, as shown in Figure 3.15. Afterwards, tasks can be created, as can be seen in Figure 3.16. Messages can be configured, as shown in Figure 3.17. Questions for a post-study questionnaire can be specified, as shown in Figure 3.18. In pay-for plans, the branding of the study can be customised, as shown in Figure 3.19. Finally, it is possible to invite potential participants to the study, as shown in Figure 3.20.

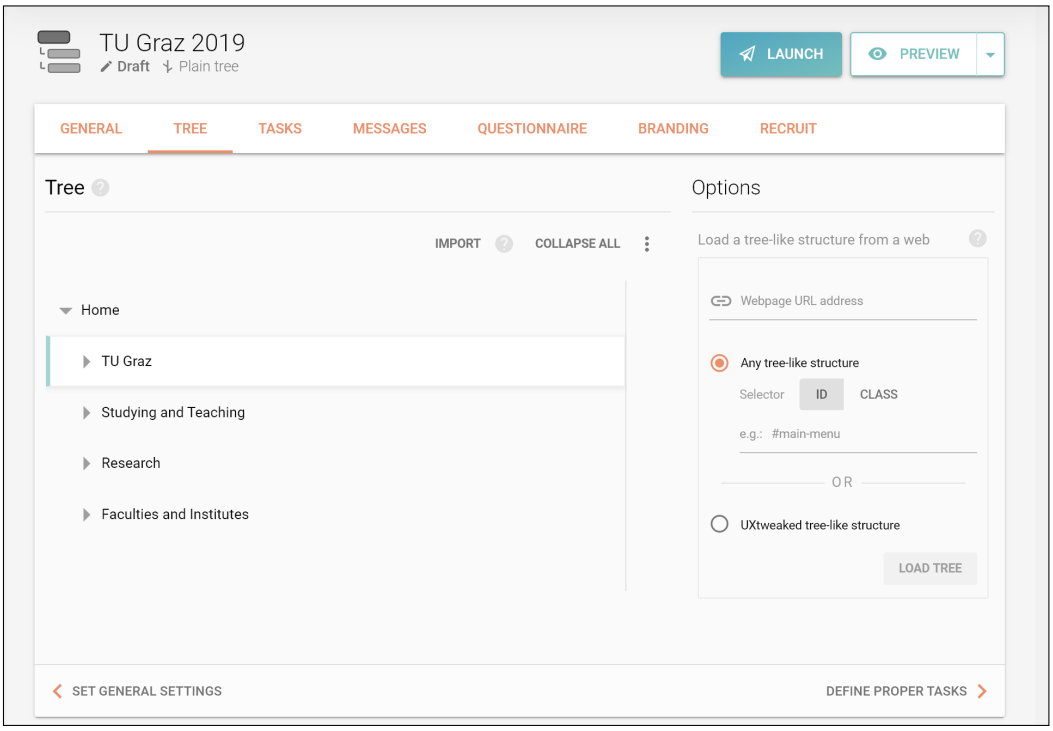

Figure 3.15: UXtweak: Defining the tree for a study. [Screenshot taken by the author of this thesis.]

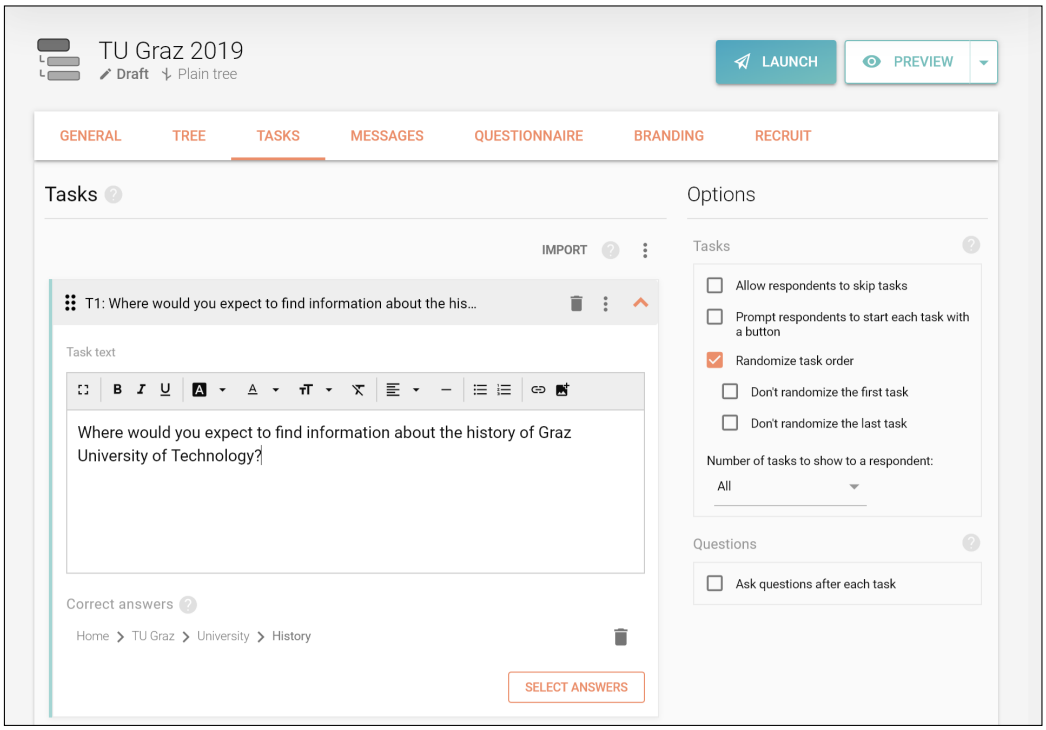

Figure 3.16: UXtweak: Defining tasks for a study. [Screenshot taken by the author of this thesis.]

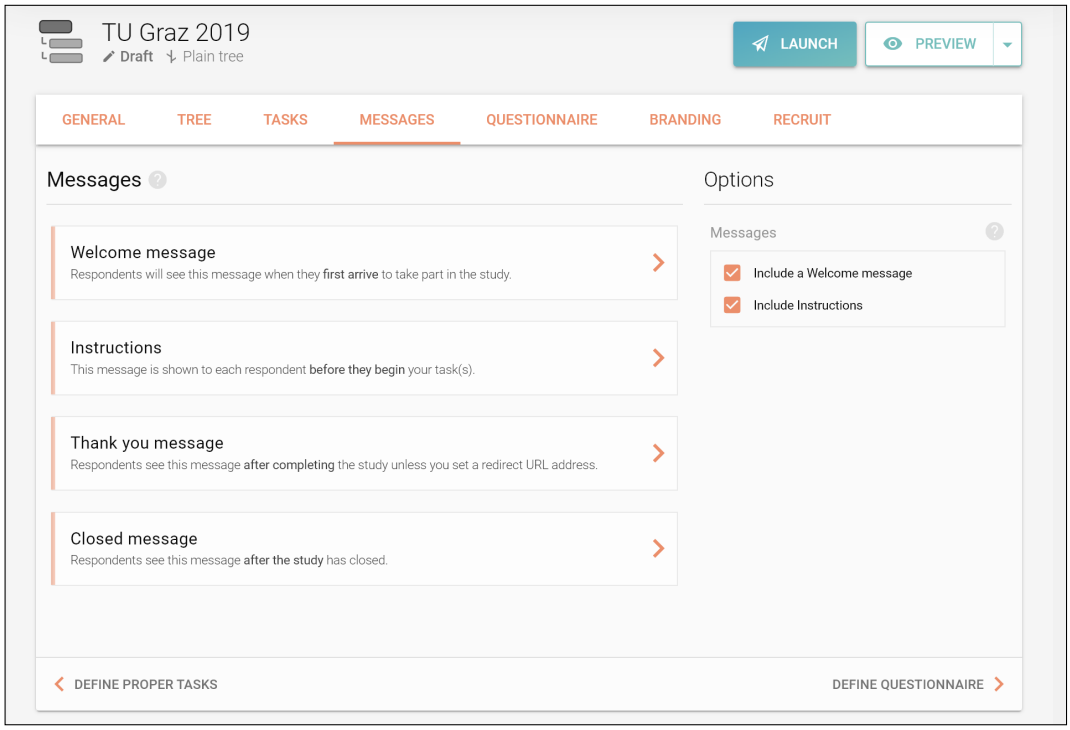

Figure 3.17: UXtweak: Configuring customised messages for a study. [Screenshot taken by the author of this thesis.]

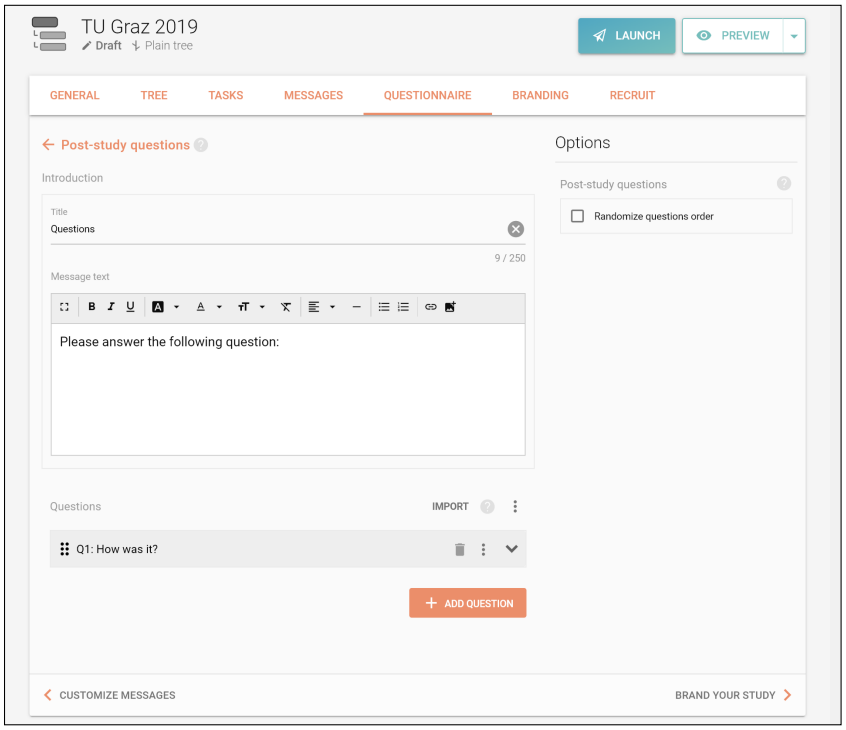

Figure 3.18: UXtweak: Defining a post-study questionnaire for a study. [Screenshot taken by the author of this thesis.]

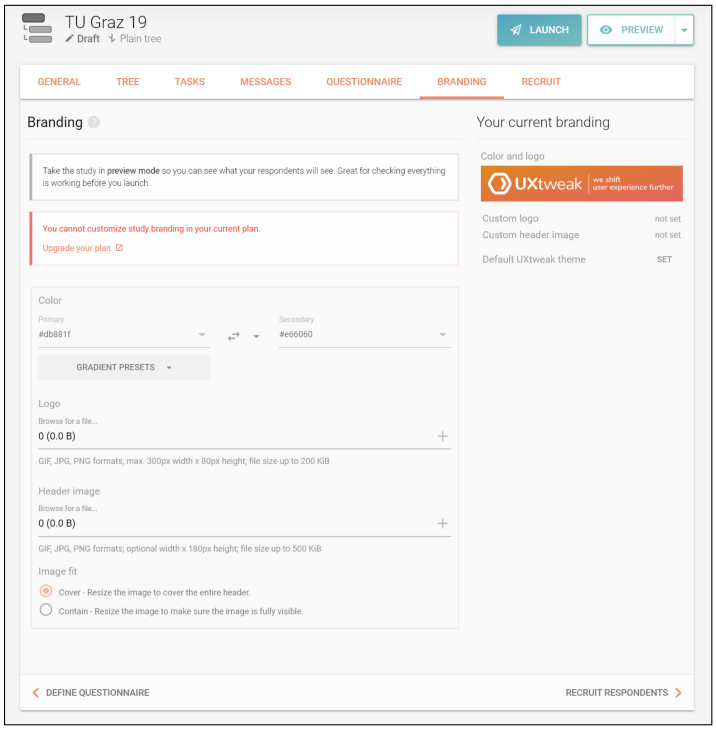

Figure 3.19: UXtweak: In pay-for plans, it is possible to customise the branding of a study. [Screenshot taken by the author of this thesis.]

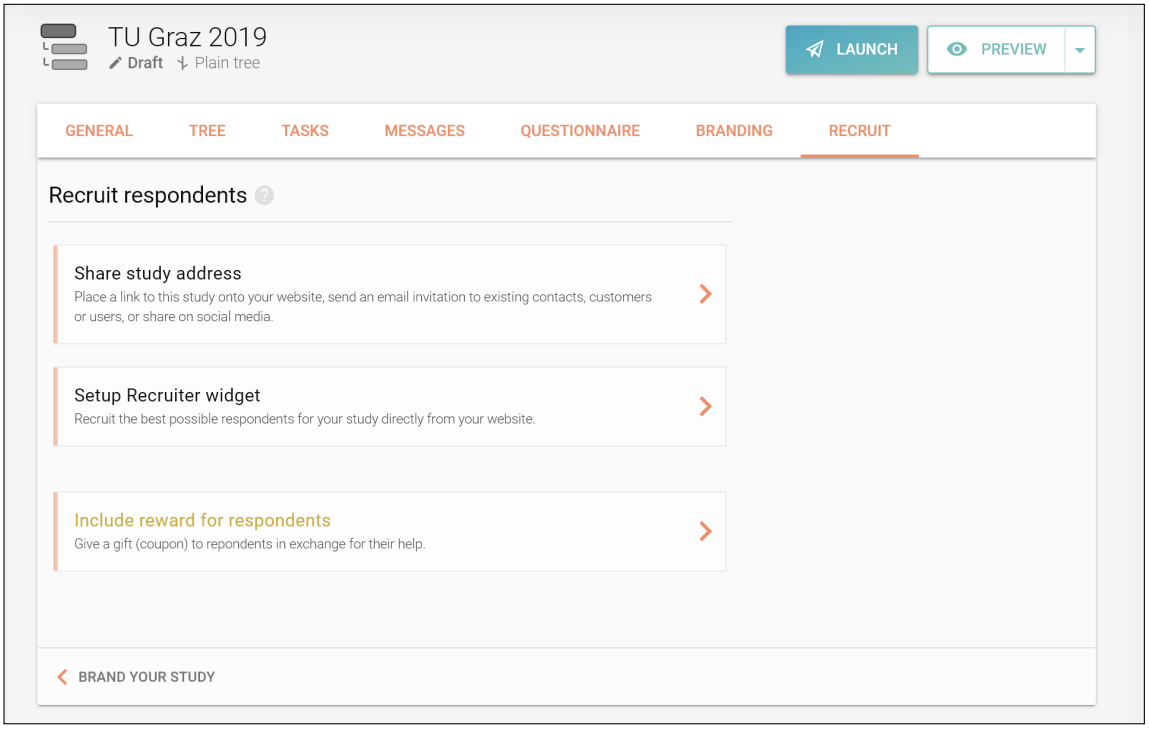

Figure 3.20: UXtweak: Inviting participants to participate in a study. [Screenshot taken by the author of this thesis.]

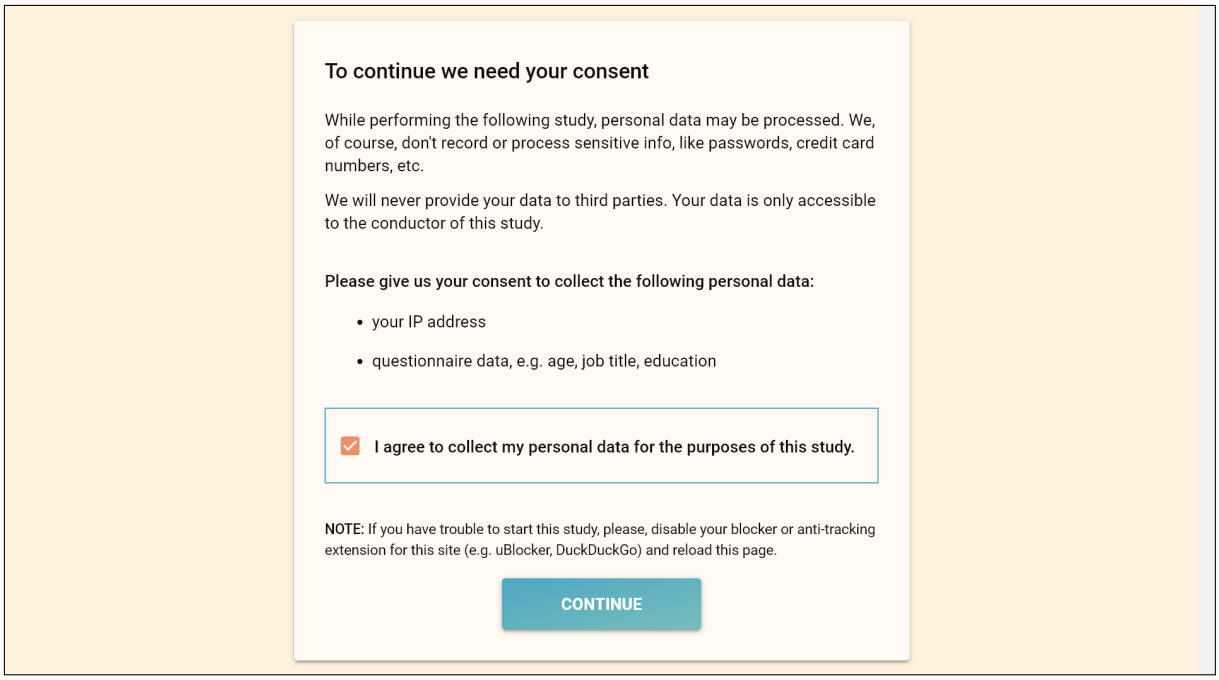

Figure 3.21: UXtweak: The consent form displayed to a study participant. [Screenshot taken by the author of this thesis.]

3.2.2 Participating in a Study

As shown in Figure 3.21, participants in a study are first shown a consent form. Next, a welcome message like the one in Figure 3.22 and a short set of instructions like in Figure 3.23 are displayed. Subsequently, tasks like the one shown in Figure 3.24 are displayed to the study participant, one after another. If a post-study questionnaire was configured, it is then displayed to the participant, as shown in Figure 3.25. At the end of a study, a thank you message like the one in Figure 3.26 is shown to the user.

3.2.3 Results of a Study

A study owner can access and manage study results from UXtweak's Active Studies page, as shown in Figure 3.27. The results of a study are collected into a page with three tabs. The first tab Overview displays overall summary statistics for the study, as can be seen in Figure 3.28. The second tab Respondents shows the results for each participant, as can be seen in Figure 3.29. It is also possible to exclude individual participants from the analysis. The third and final tab Analysis, shown in Figure 3.30, gives access to various kinds of detailed results. These include a pietree, shown in Figure 3.31, which graphically displays the paths taken by the participants in search of an answer to a particular task. The paths taken by each participant can be viewed, as in Figure 3.32. Finally, a table of destinations gives an overview of the answers (destinations) chosen by participants for a particular task, as shown in Figure 3.33.

UXtweak offers different versions for their usability testing applications. In the free version, some features are not available and there are some limitations. For example, the results of at most ten participants can be accessed [UXtweak 2020b].

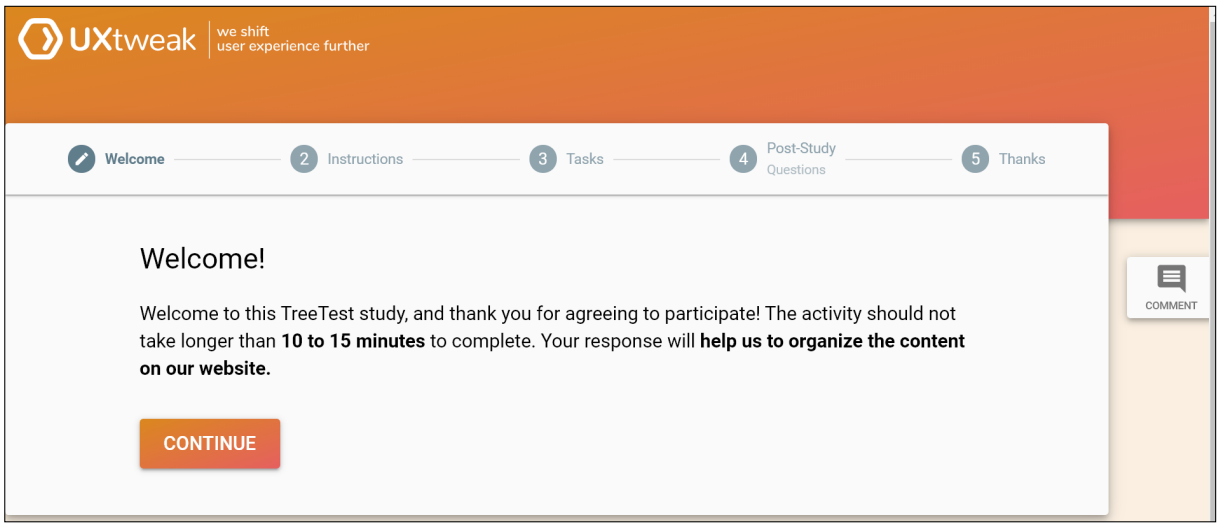

Figure 3.22: UXtweak: The welcome message displayed to a study participant. [Screenshot taken by the author of this thesis.]

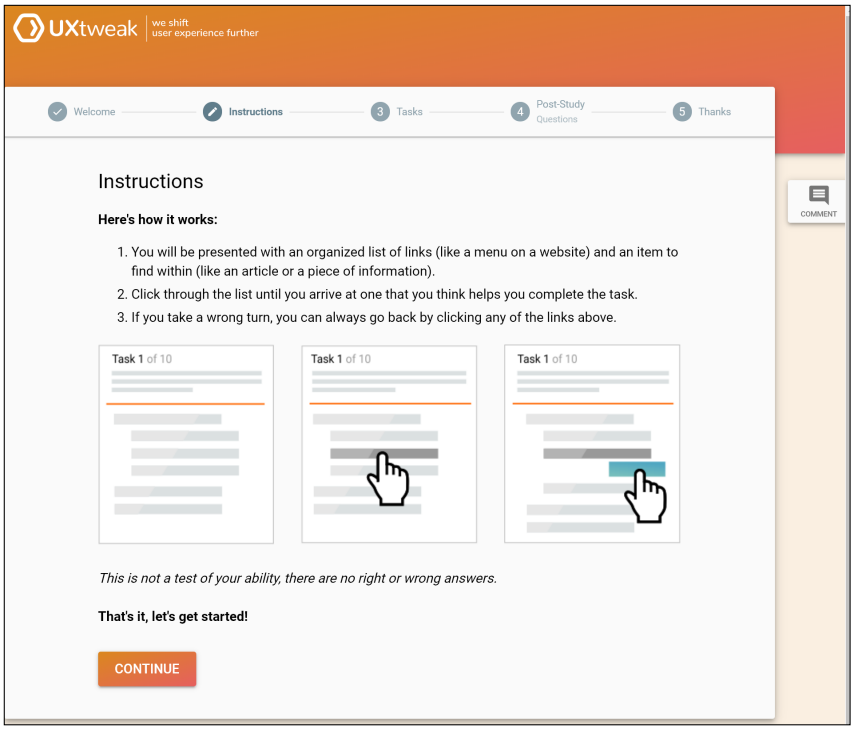

Figure 3.23: UXtweak: The instructions displayed to a study participant. [Screenshot taken by the author of this thesis.]

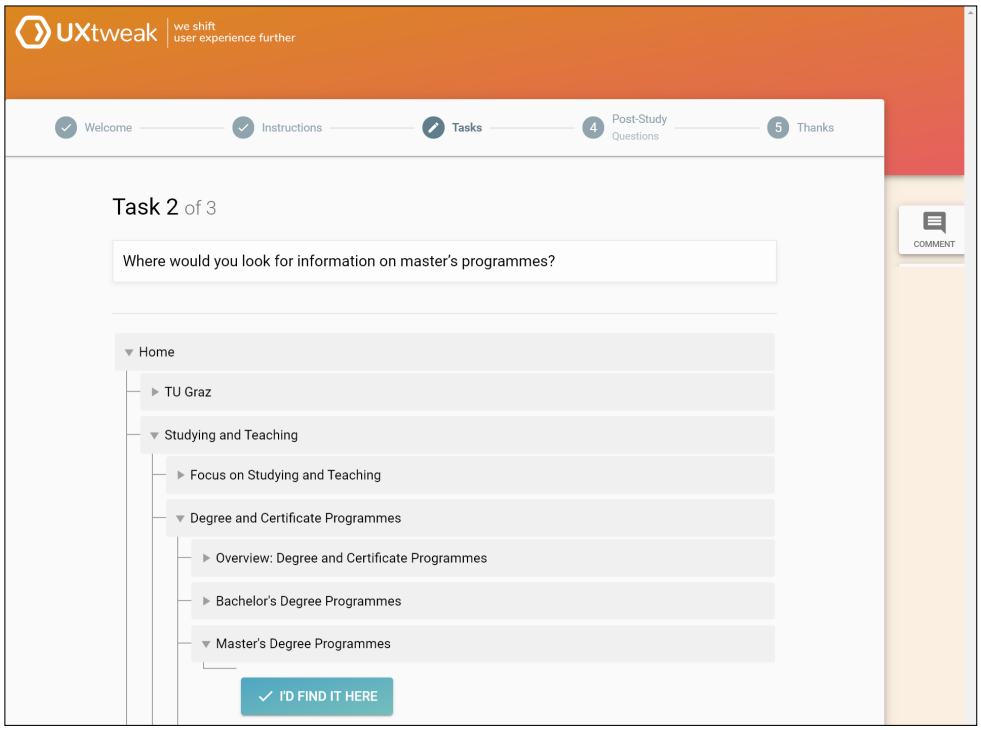

Figure 3.24: UXtweak: A typical task displayed to a study participant. [Screenshot taken by the author of this thesis.]

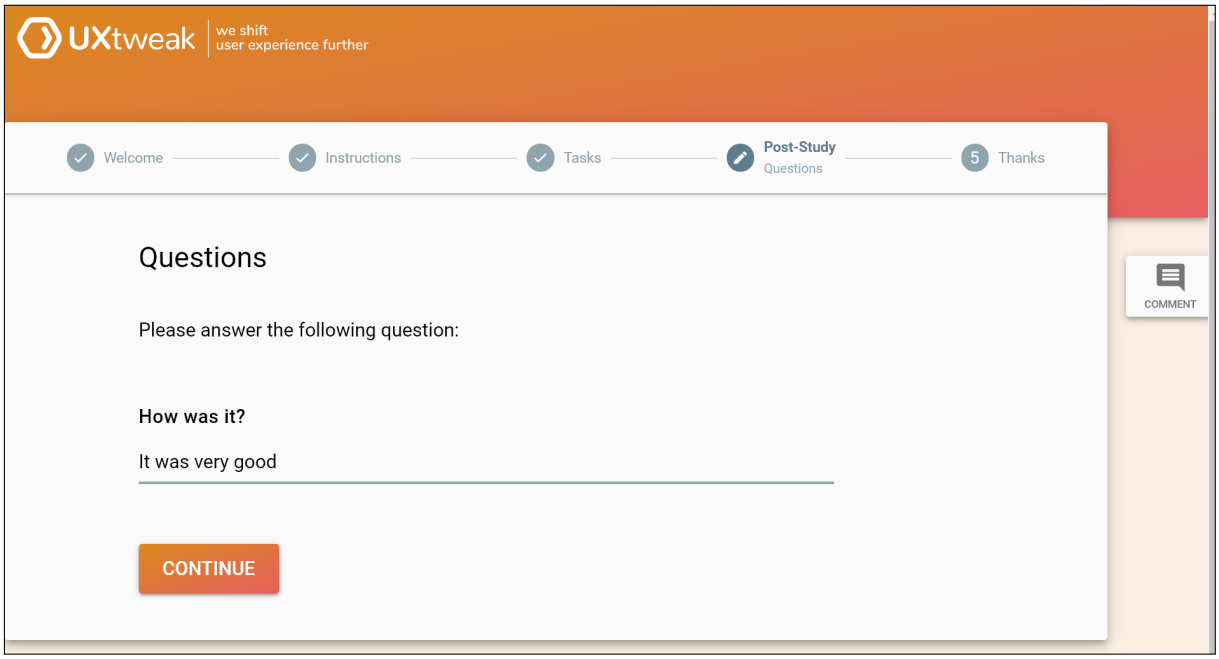

Figure 3.25: UXtweak: A question from a post-study questionnaire displayed to a participant. [Screenshot taken by the author of this thesis.]

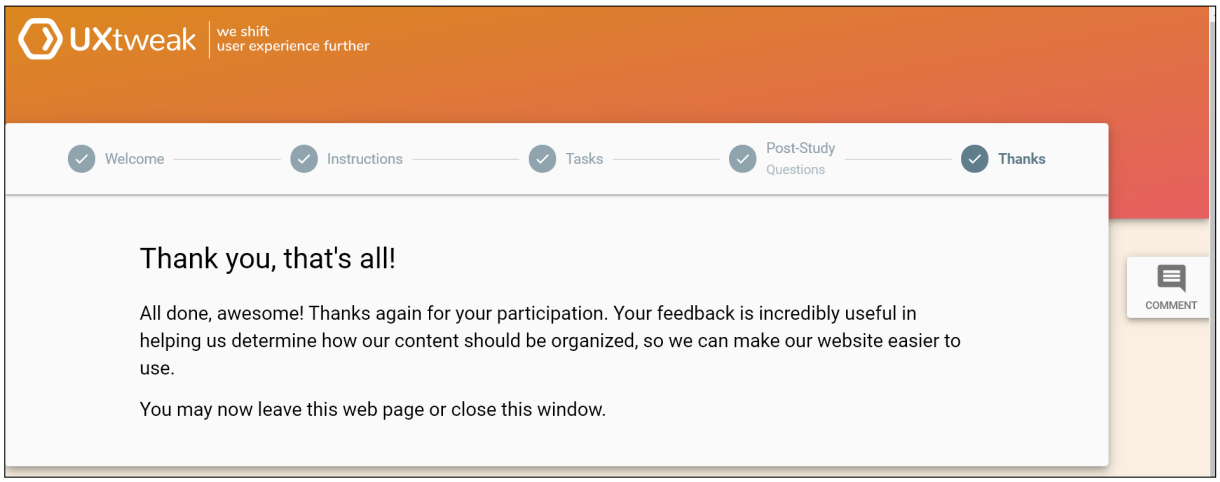

Figure 3.26: UXtweak: The thank you message displayed to a participant at the end of their session. [Screenshot taken by the author of this thesis.]

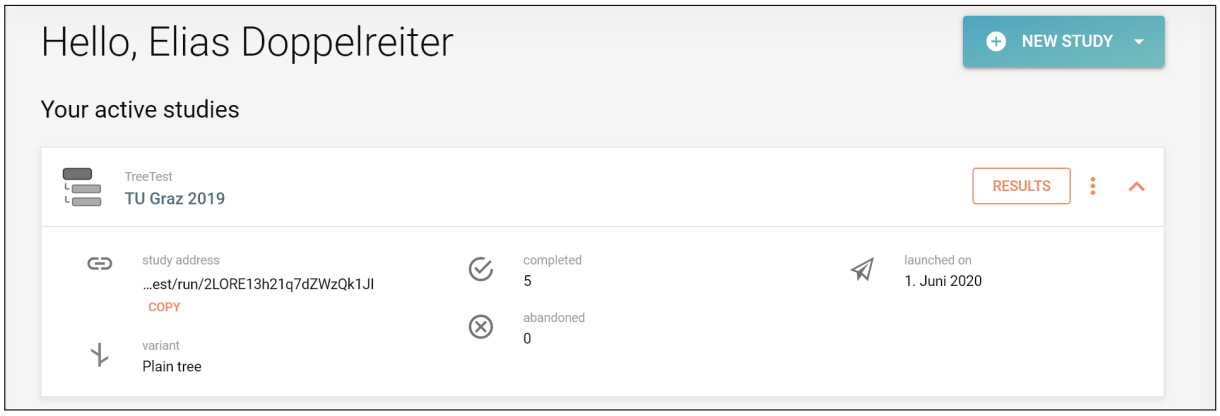

Figure 3.27: UXtweak: The Active Studies page shown to a study owner. [Screenshot taken by the author of this thesis.]

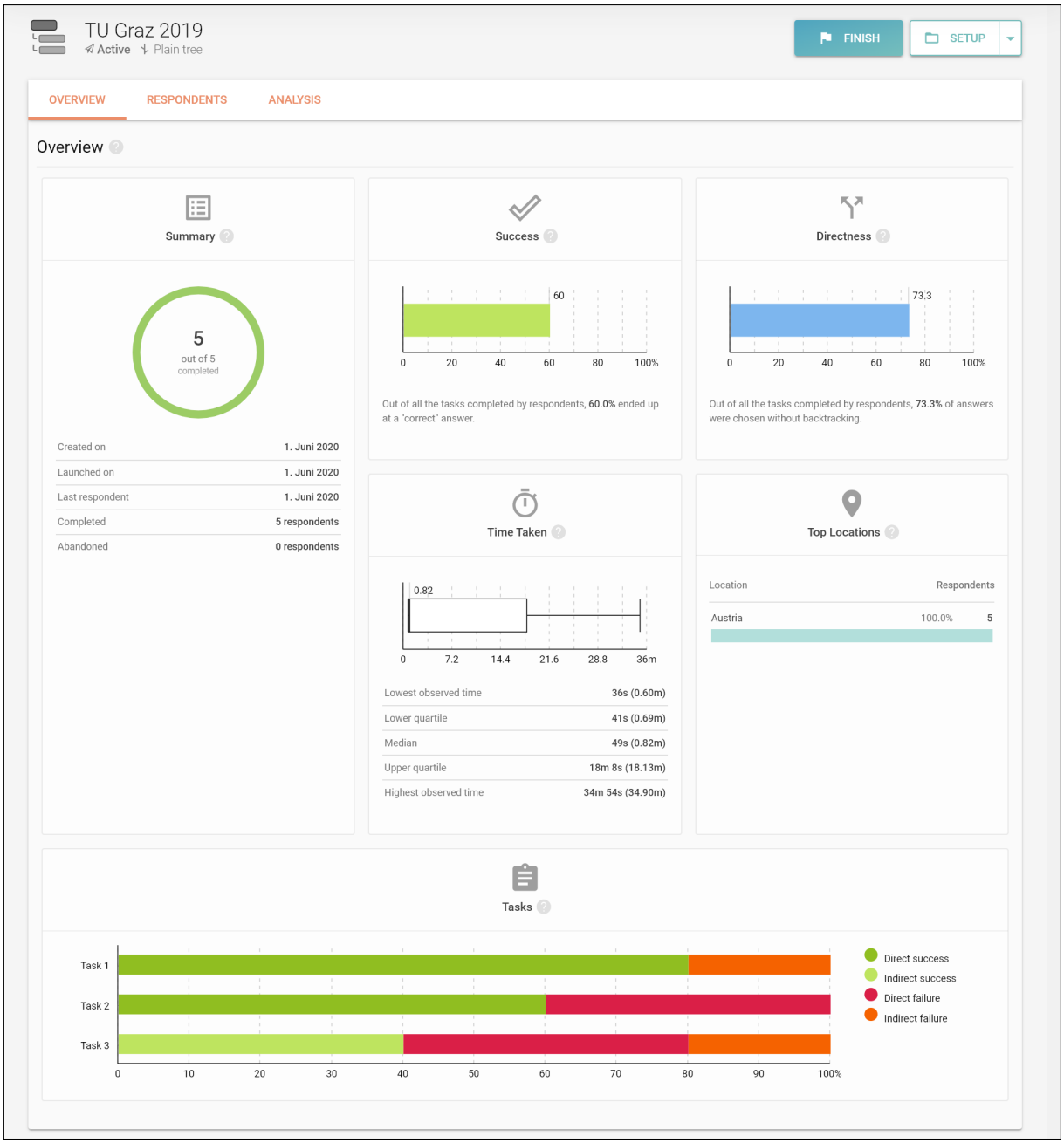

Figure 3.28: UXtweak: The Overview tab of the study results page. [Screenshot taken by the author of this thesis.]

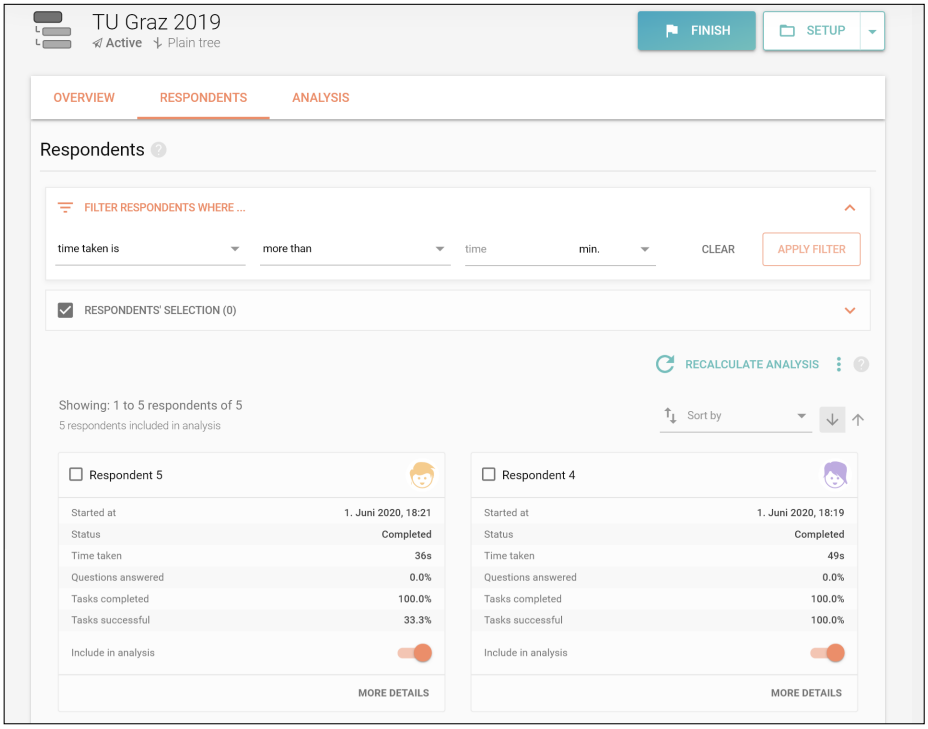

Figure 3.29: UXtweak: The Respondents tab shows the results for individual study participants. [Screenshot taken by the author of this thesis.]

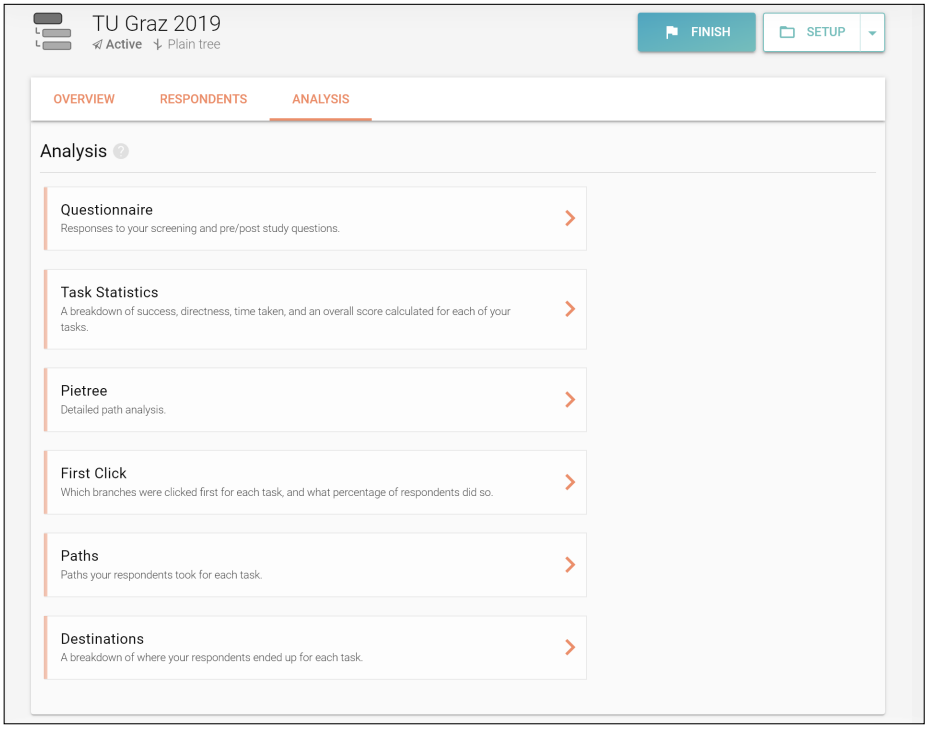

Figure 3.30: UXtweak: The Analysis tab gives access to the detailed results of the study. [Screenshot taken by the author of this thesis.]

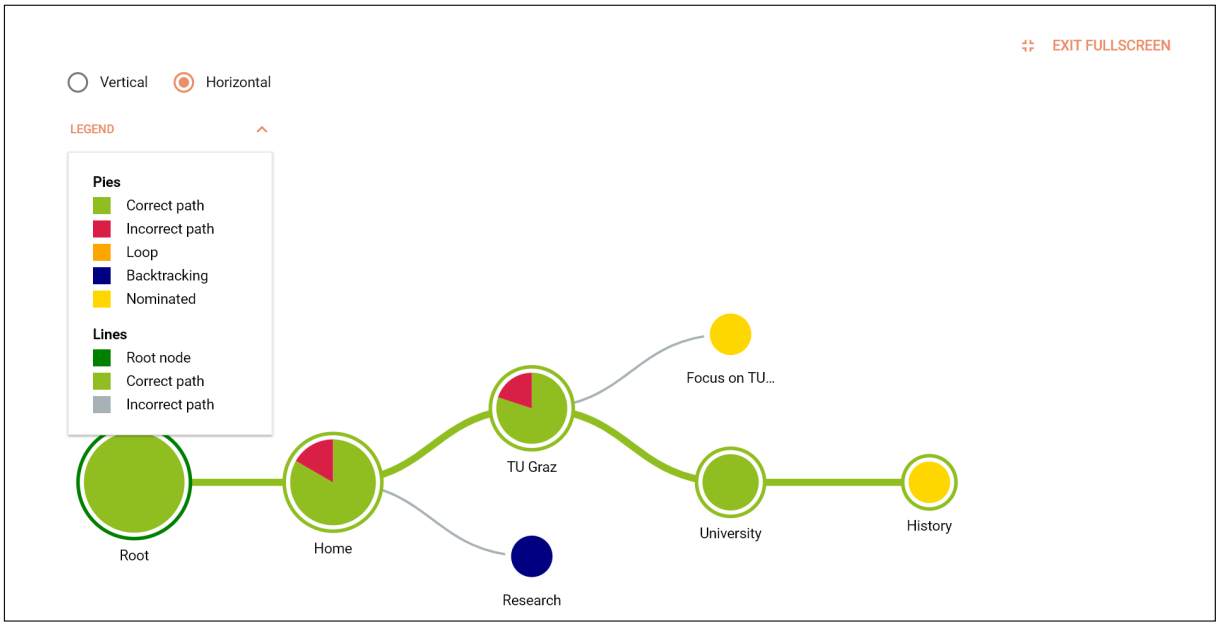

Figure 3.31: UXtweak: A pietree shows a graphical overview of the paths taken by participants for a particular task. [Screenshot taken by the author of this thesis.]

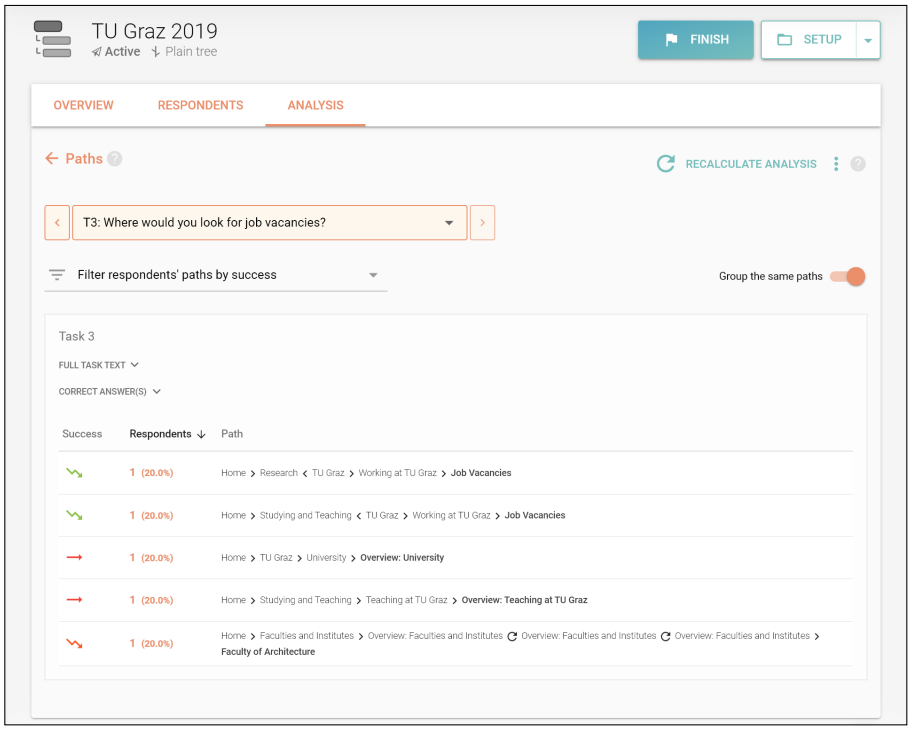

Figure 3.32: UXtweak: A textual overview of the paths taken by participants for a particular task. [Screenshot taken by the author of this thesis.]

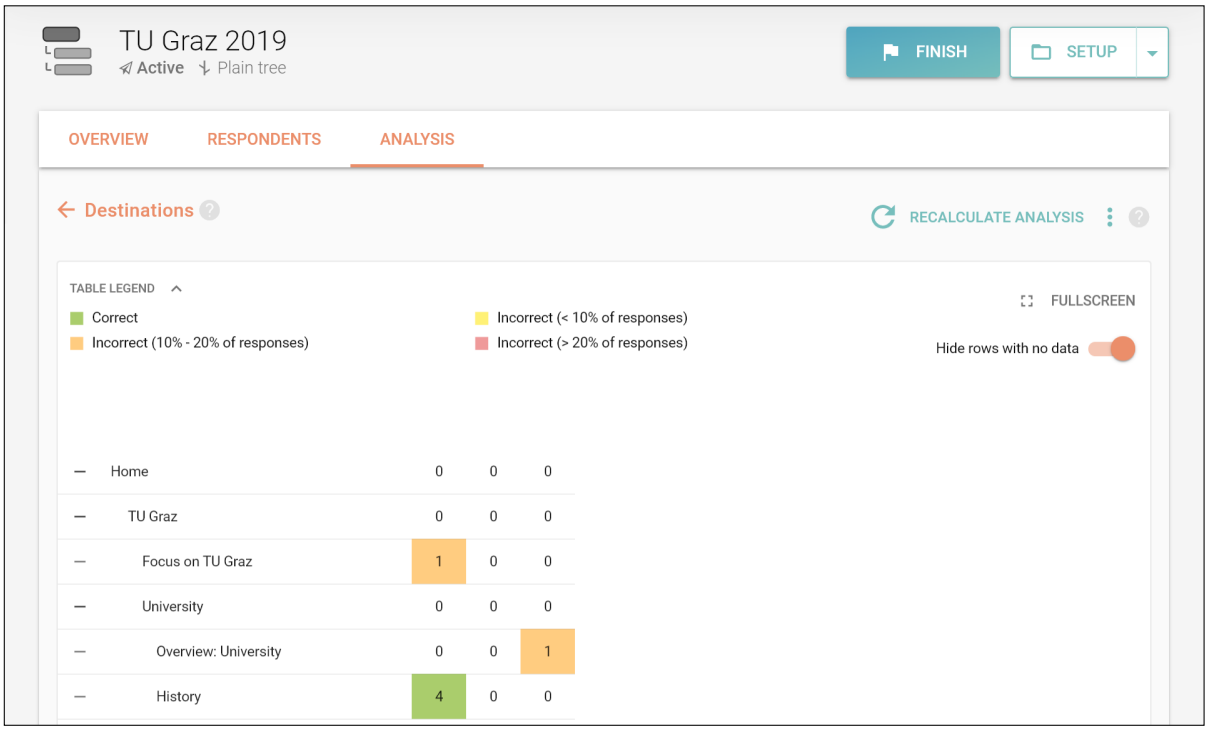

Figure 3.33: UXtweak: The destinations table gives an overview of the answers (destinations) chosen by participants for each task. [Screenshot taken by the author of this thesis.]

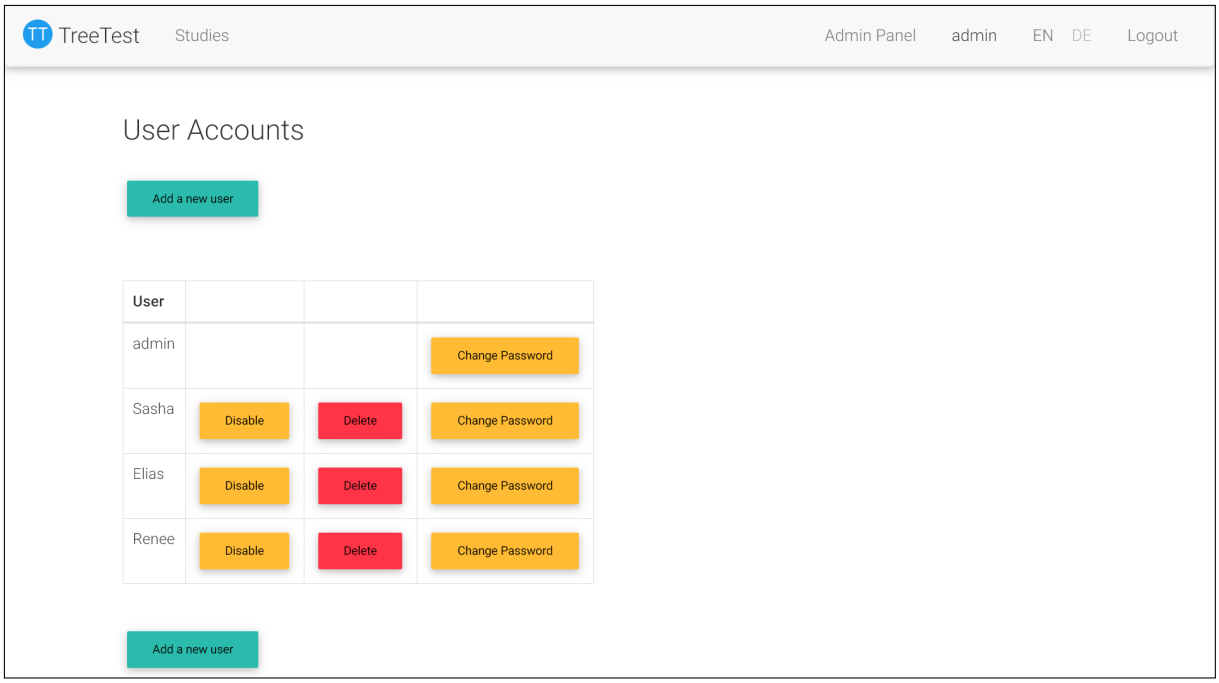

Figure 3.34: TreeTest: The Administration Panel allows administrators to manage user accounts. [Screenshot taken by the author of this thesis.]

3.3 TreeTest

Ajdin Mehic developed TreeTest for his Master's Thesis [Mehic 2019]. It is an open-source web application for tree testing built with the MEAN (MongoDB, Express, Angular, and Node) stack [MEAN 2019]. This web application was evaluated with a thinking aloud usability test in the context of this thesis.

TreeTest has three different user groups:

- Administrators: Can manage user accounts.
- Study owners: Can create studies and view the results.
- Study participants: Can participate in studies.

3.3.1 Administration Panel

Figure 3.34 shows the Administration Panel of TreeTest. Here, administrators can create user accounts and activate them as study owners. Furthermore, administrators can change the passwords of user accounts and delete user accounts.

3.3.2 Creating a Study

The Studies page, shown in Figure 3.35, serves as the main page for study owners. On this page, the study owner can manage studies and create new studies by clicking the Create Study button. Study creation follows a five-step process through five tabs on the New Study page. Editing an existing study uses the same five tabs. First, in the Settings tab shown in Figure 3.36, settings such as the study title and an optional password can be set. A link to the study is also displayed. The second tab, Tree, shown in Figure 3.37, allows the study owner to import or create the tree to be used in the study. Tasks are created in the third tab Tasks, shown in Figure 3.38. A question is formulated in the lefthand panel, and the corresponding answer is specified in the righthand panel. The fourth tab, Messages, shown in Figure 3.39, allows the

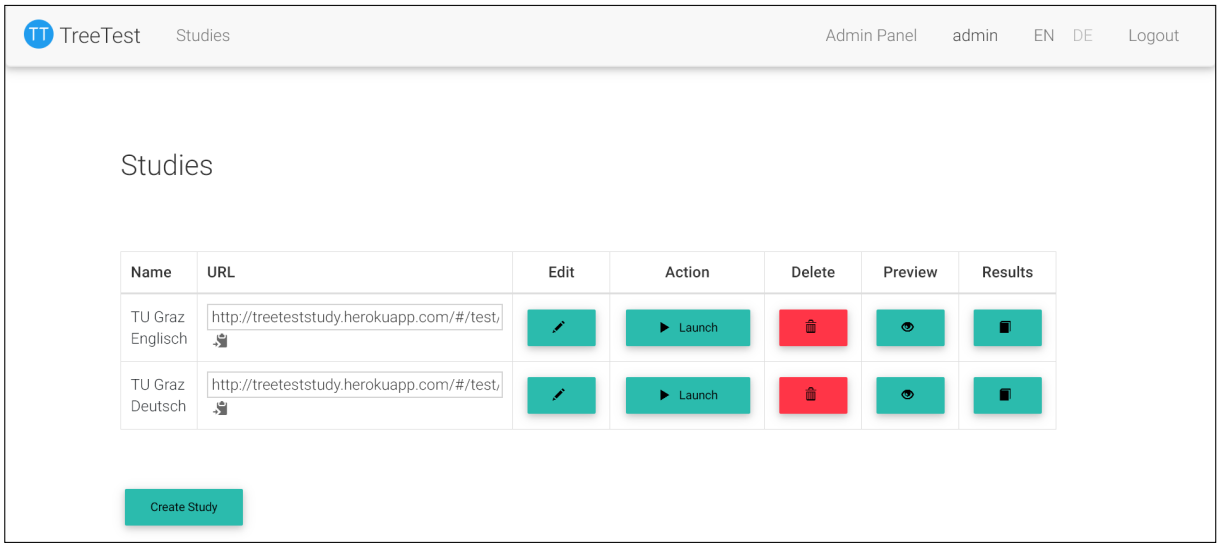

Figure 3.35: TreeTest: Studies page for study owners. [Screenshot taken by the author of this thesis.]

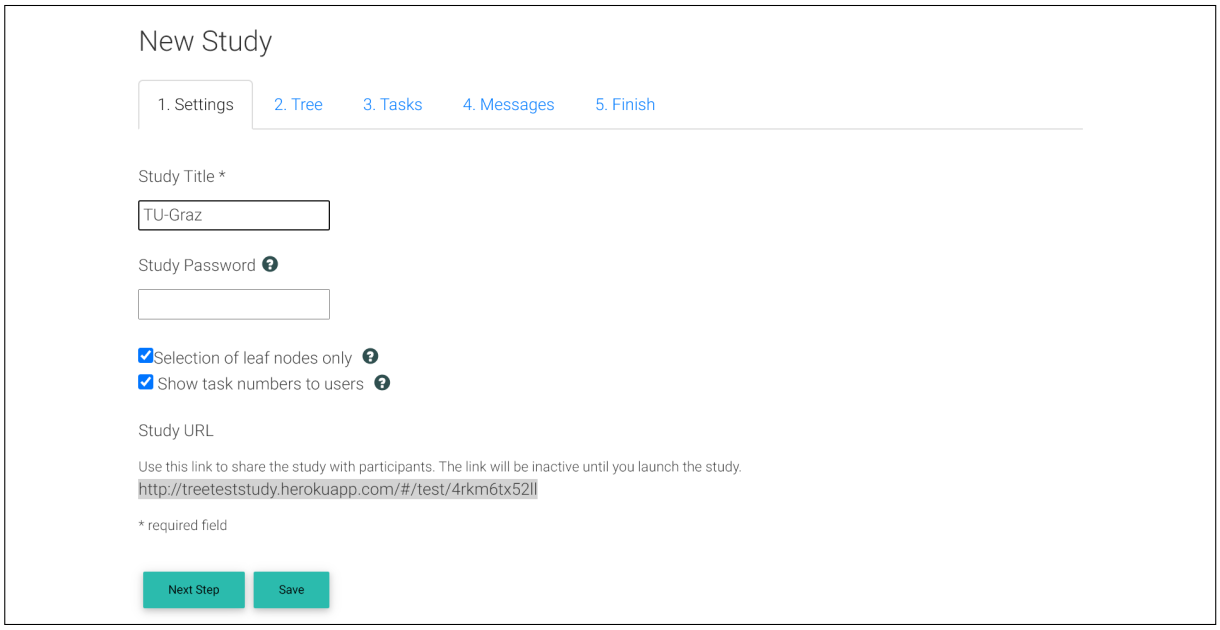

Figure 3.36: TreeTest: The Settings tab for a study allows the study owner to configure various settings. [Screenshot taken by the author of this thesis.]

study owner to customise various messages and notifications displayed to study participants during the study. The fifth and final tab, Finish, shown in Figure 3.40, confirms that the study has been set up and saved. The study owner is reminded that the study has to be launched (from the Studies page) before users can participate in the study.

3.3.3 Participating in a Study

Participant in a study are shown a short welcome message and asked to enter their name, as shown in Figure 3.41. If the study is password protected, the participant is also asked to enter the password. Next, the participant is given the tasks one by one and asked to find the answer location in the tree. A typical task displayed to a participant is shown in Figure 3.42. The participant also has the possibility to skip the current task and move to the next task. Once the participant has completed all the tasks, they have the

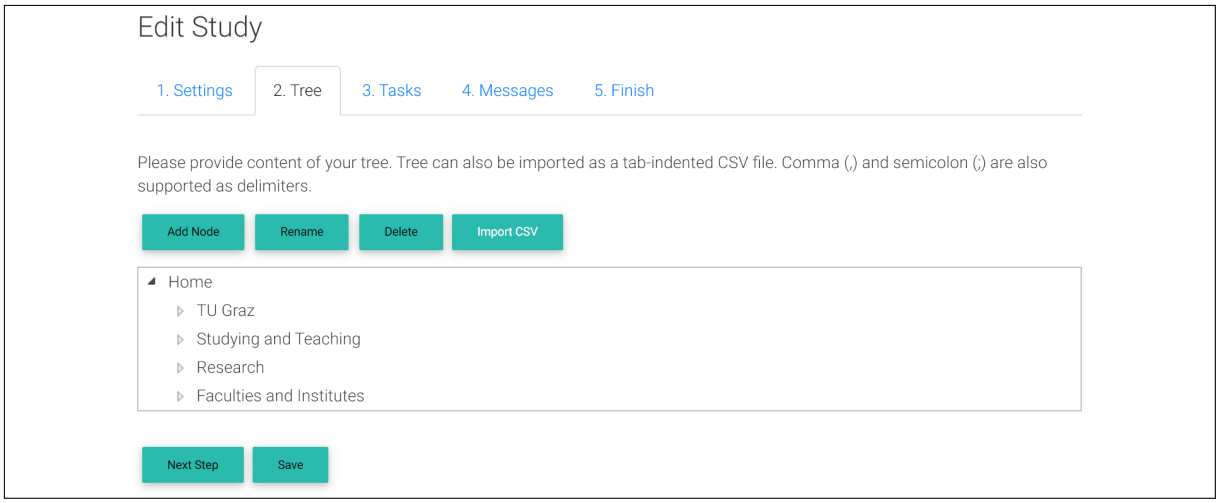

Figure 3.37: TreeTest: The Tree tab allows the study owner to create or import the tree for a study. [Screenshot taken by the author of this thesis.]

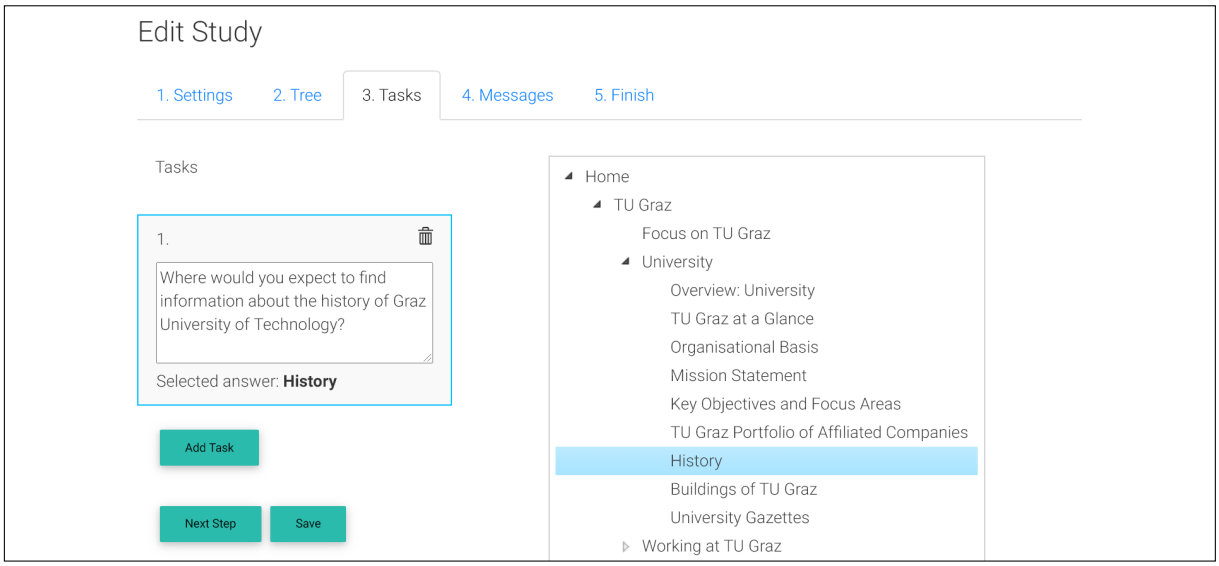

Figure 3.38: TreeTest: The Tasks tab allows the study owner to define tasks and specify the correct answer for each task. [Screenshot taken by the author of this thesis.]

opportunity to leave feedback, as shown in Figure 3.43.

3.3.4 Results of a Study

From the Studies page, a study owner can access the results of a study. The Study Results page comprises four tabs, the first of which, the Overview tab, is shown in Figure 3.44. This tab display summary statistics for the study as a whole, such as the number of participants, success rate, and directness rate. The second tab, Participants, shown in Figure 3.45, shows the results for each participant, including the duration of their session, the number of tasks completed, and the number of tasks completed correctly. The third tab, Task Analysis, displays the results per task, shown in Figure 3.46. In addition to the success and directness rates for the task, it is possible to display a path tree of the paths taken by participants during this task. An example of a path tree is shown in Figure 3.47. The fourth and final tab, Destinations, displays a table of answers (destinations) chosen by users for each task, as can be seen in Figure 3.48.

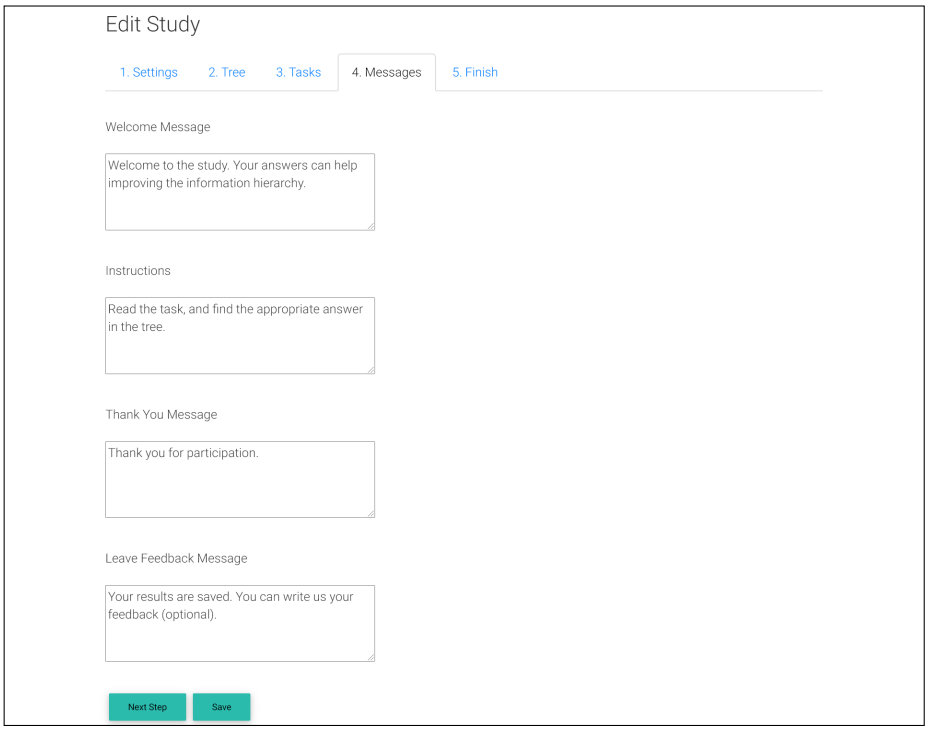

Figure 3.39: TreeTest: The Messages tab allows the study owner to configure customised messages for a study. [Screenshot taken by the author of this thesis.]

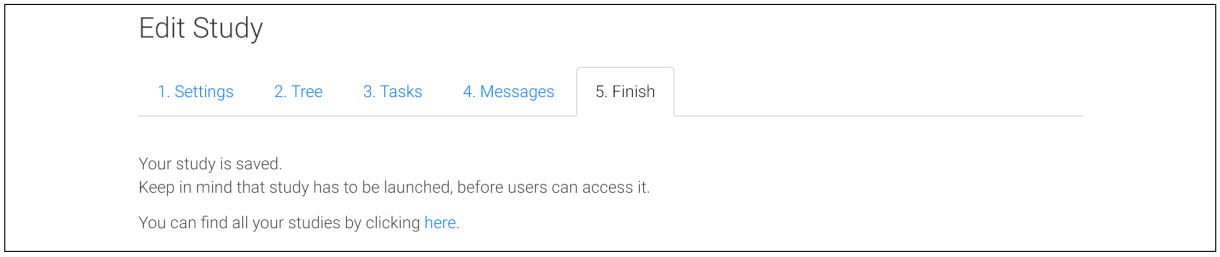

Figure 3.40: TreeTest: The Finish tab confirms that the study has been set up and saved. [Screenshot taken by the author of this thesis.]

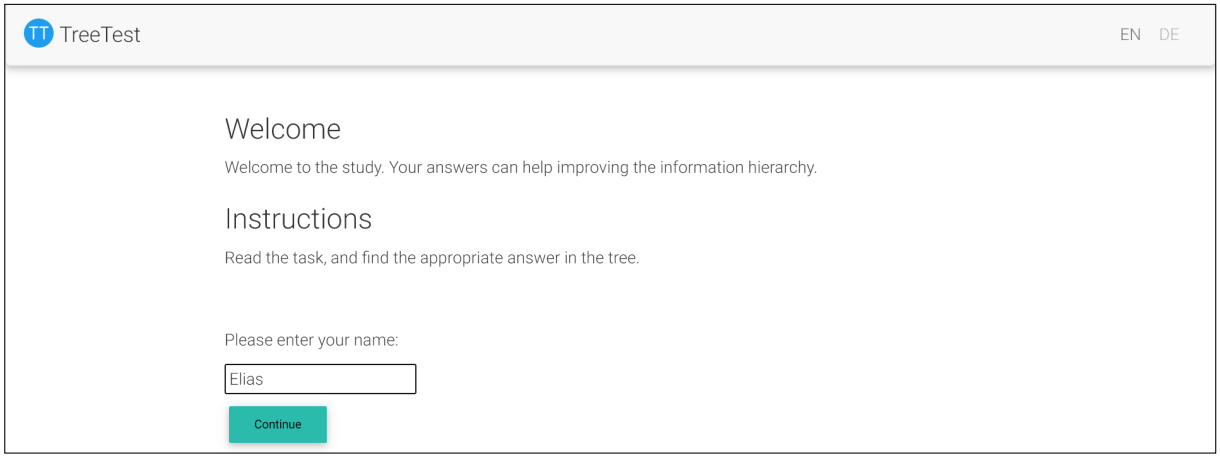

Figure 3.41: TreeTest: The welcome screen displayed to a study participant. [Screenshot taken by the author of this thesis.]

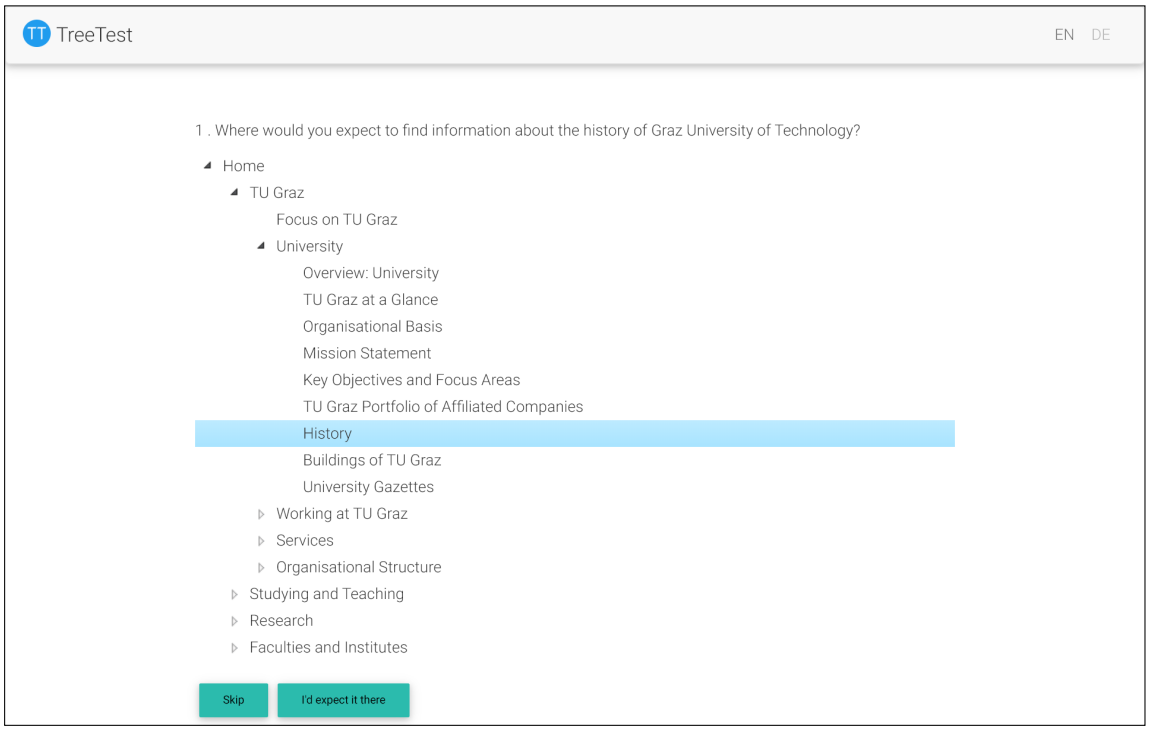

Figure 3.42: TreeTest: A typical task displayed to a participant in a study. [Screenshot taken by the author of this thesis.]

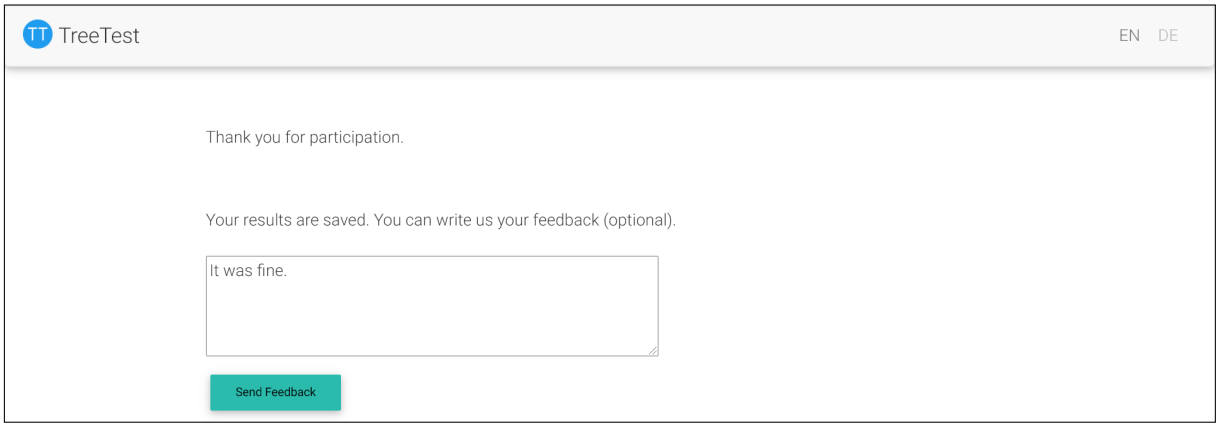

Figure 3.43: TreeTest: At the end of a test session, the participant is thanked and given the opportunity to leave feedback. [Screenshot taken by the author of this thesis.]

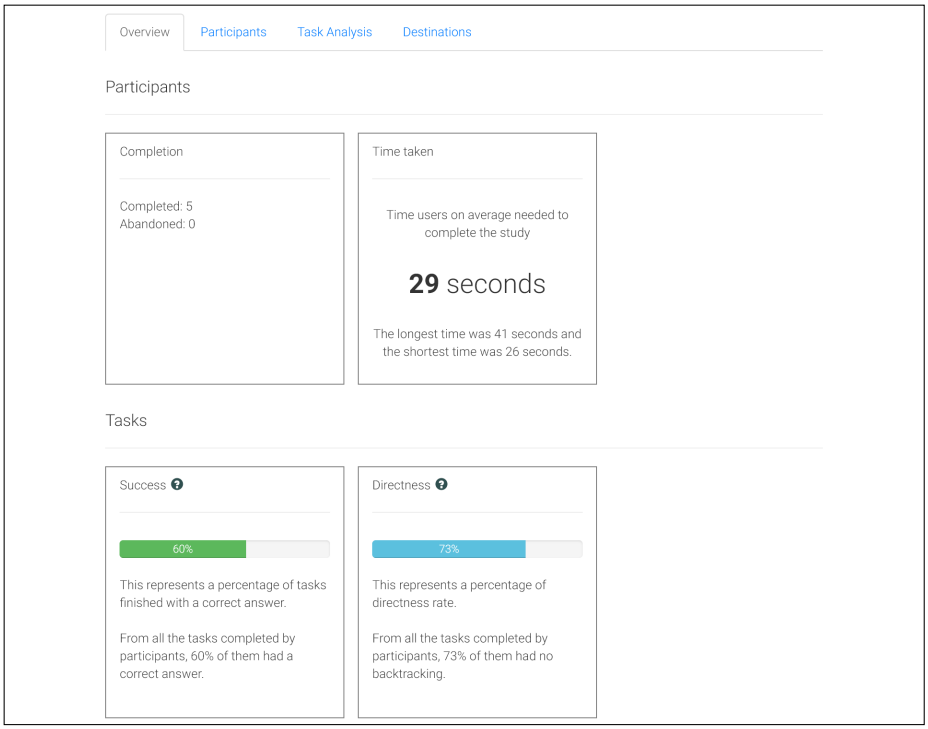

Figure 3.44: TreeTest: The Overview tab of the Study Results page display summary statistics for the whole study. [Screenshot taken by the author of this thesis.]

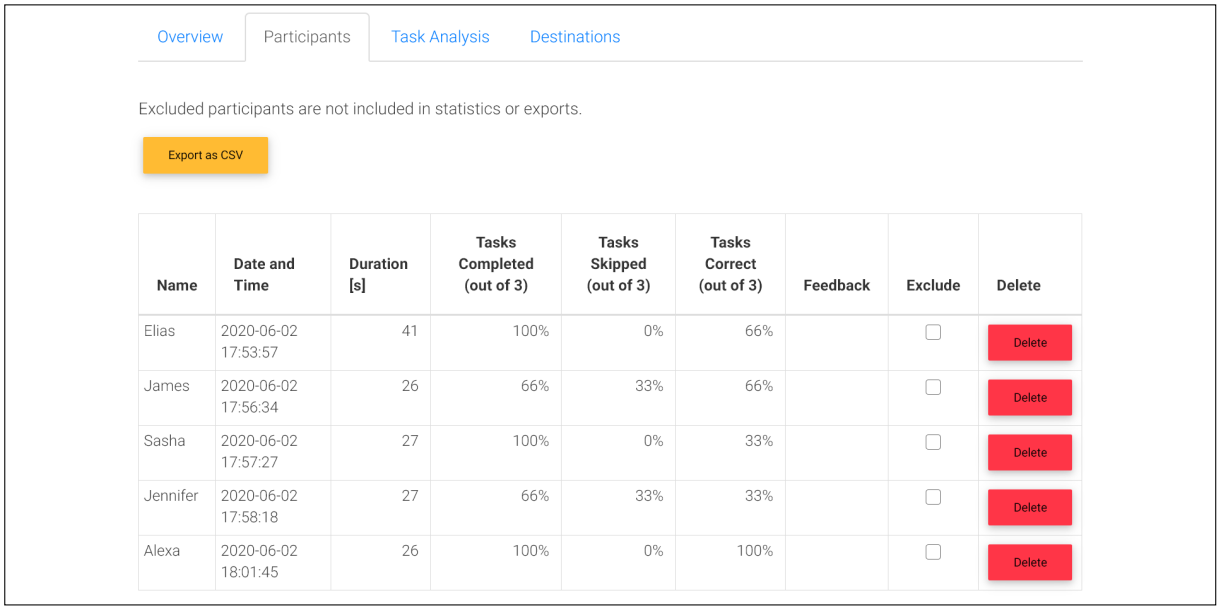

Figure 3.45: TreeTest: The study results of each individual participant. [Screenshot taken by the author of this thesis.]

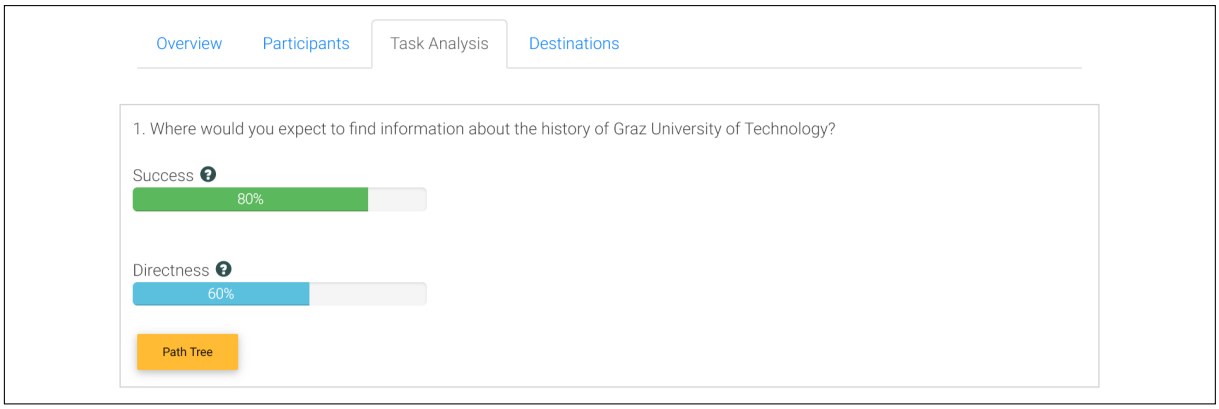

Figure 3.46: TreeTest: The study results for each individual task. [Screenshot taken by the author of this thesis.]

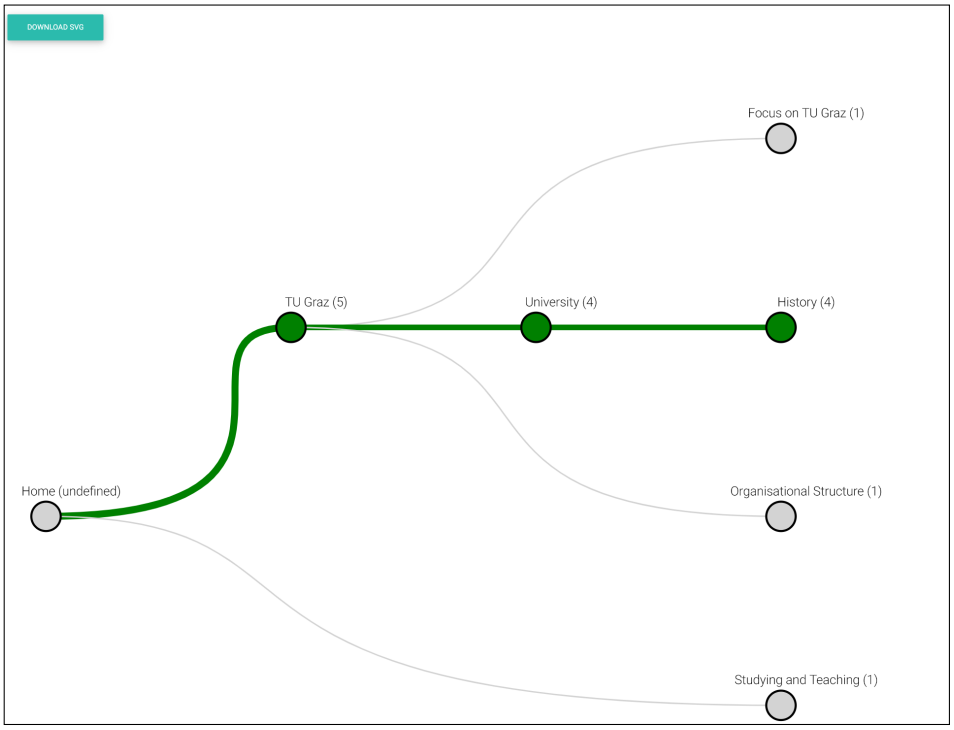

Figure 3.47: TreeTest: The path tree for a task shows a graphical overview of the paths taken by participants for that task. [Screenshot taken by the author of this thesis.]

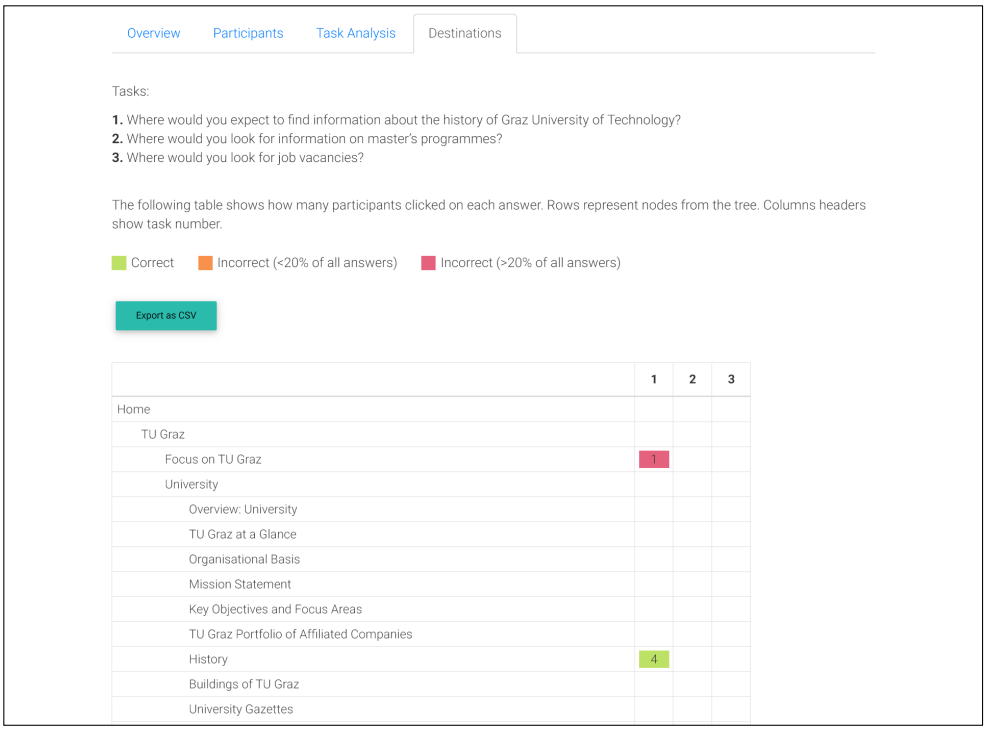

Figure 3.48: TreeTest: The destinations table gives an overview of the answers (destinations) chosen by participants for each task. [Screenshot taken by the author of this thesis.]

Chapter 4

Usability Evaluation

> *"Most people make the mistake of thinking design is what it looks like. People think it's this veneer — that the designers are handed this box and told, 'Make it look good!' That's not what we think design is. It's not just what it looks like and feels like. Design is how it works."*
>
> [Steve Jobs, interview with New York Times, 2003 [Walker 2003]]

The word usability describes a property which indicates how easy the user interface of an application is to use. If an application cannot meet the expectations of its users, they will not be satisfied and will look for alternatives [Nielsen 2012]. According to Nielsen [2012], there are five measurable criteria of usability:

- *Learnability*: How easy it is for new users to perform a task for the first time.
- *Efficiency*: The time required to complete a task.
- *Memorability*: How easy it is for casual users to remember how to use an interface.
- *Errors*: How often errors occur and how easy it is to recover from them.
- *Satisfaction*: How pleasant it is to interact with the user interface.

This chapter describes some of the methods used for evaluating the usability of applications.

4.1 Usability Evaluation Methods

Usability evaluation methods can be broadly divided into *inspection* methods and *testing methods*, according to whether the method involves test users or not [Andrews 2008; Andrews 2020, Chapter 3]. Inspection methods are carried out by experts in usability and do not involve test users [Nielsen and Mack 1994]. Examples of inspection methods include cognitive walkthrough [Wharton et al. 1994], guideline checking [Andrews 2020, page 101], and heuristic evaluation [Nielsen 1994b; Nielsen and Molich 1990b]. Testing methods, on the other hand, are managed or facilitated by usability experts, but involve representative end users performing typical tasks with the interface being evaluated [Rubin and Chisnell 2008]. Examples of testing methods include diary studies [Salazar 2016], thinking aloud tests [Krug 2009], and formal experiments [Sauro 2018].

Another way to divide up usability evaluation methods is when in the software development lifecycle they are performed (or the kind of insights they are meant to deliver). Of particular interest are *formative*

methods and *summative* methods [Andrews 2008]. Formative methods are carried out during interface development, in order to find potential usability problems, to fix them. The output of a formative evaluation is typically a list of problems with an interface. Examples of formative methods include heuristic evaluation and thinking aloud testing. Summative methods are carried out after (a particular round of) interface development has finished, in order to objectively assess the overall quality of an interface. Often, a new interface is compared with a previous version or a competitor's interface. Summative methods involve measurement and statistical analysis. Examples of summative methods include formal experiments [Sauro 2018] and A/B testing [Kohavi et al. 2020].

4.2 Heuristic Evaluation

Heuristic evaluation is a formative usability inspection method [Nielsen and Molich 1990a]. A small team of usability experts look through an interface and assess its usability against a small set of usability principles (or heuristics). It is possible to perform a heuristic evaluation very early in the development phase of the user interface, for example with paper prototypes [Nielsen 1992; Nielsen 1994c]. The problems found are classified according to the usability principles [Nielsen and Molich 1990b].

The first set of nine usability heuristics were published in 1990 by Nielsen and Molich [1990a]. In 1994, Nielsen [1994a] renamed some of the heuristics and added a tenth heuristic. The revised set of ten heuristics comprises:

1. *Visibility of system status*: The system should continuously inform the users about what is happening.
2. *Match between system and the real world*: The system should communicate in the language of the user.
3. *User control and freedom*: The system should support undo and redo.
4. *Consistency and standards*: The system should use uniform guidelines to avoid ambiguity.
5. *Error prevention*: Functions that could lead to errors should be removed, or the user should be informed that problems may occur.
6. *Recognition rather than recall*: Users should not have to remember things. Instructions to the users should be formulated understandably.
7. *Flexibility and efficiency of use*: The system should be designed in such a way that both experienced and inexperienced users can use the system.
8. *Aesthetic and minimalist design*: Less is more, the interface should be clean and simple without extraneous decorations.
9. *Help users recognize, diagnose, and recover from errors*: Error messages should not contain error codes and should be written in cleartext.
10. *Help and documentation*: Documentation should be written from the user's perspective and should be easy to find.

Individual evaluators tend to find different potential problems with an interface, with only limited overlap. Hence, it is recommended that a small team of evaluators work alone initially, but that their lists of problems are then combined into an aggregate list. The increase in usability problems found rises rapidly with aggregates of between 1 and 5 evaluators, but then flattens out rapidly. Jakob Nielsen recommends a team of 3 to 5 evaluators [Nielsen and Molich 1990a]. Typically, each evaluator spends

up to 2 hours evaluating an interface. If the system is extensive, it is then recommended to divide the interface into parts which are evaluated by different groups of evaluators [Nielsen 1994c]. Once an aggregated list of all usability problems found has been created, it is possible to assign a severity rating to each problem. For this, each evaluator is asked to rate the severity of each problem in the list according independently. Afterwards, the mean value of the severity rating is determined for each problem [Nielsen 1994d]. Nielsen [1994d] recommends using a five-point scale for severity rating:

- 0: Not a usability problem
- 1: Cosmetic problem
- 2: Minor usability problem
- 3: Major usability problem
- 4: Usability problem is a catastrophe

The output of a heuristic evaluation is a list of problems found in an interface, including a severity rating. Heuristic evaluation does not always address how these problems should be solved [Nielsen and Molich 1990a; Nielsen 1994d].

4.3 Thinking Aloud Testing

This section on thinking aloud testing is based on the material by Krug [2009] and Andrews [2020, Chapter 9]. Thinking aloud testing is a formative usability testing method. The facilitator asks test users to perform different tasks with an interface, while at the same time thinking out loud. Since users provide a running commentary on their actions, insight can be gained into *why* problems occurred. The test session can be recorded for later analysis. Additionally, observers (from the client or development team) can watch the session in an adjacent room using screen-sharing software. Typically, between 3 and 5 test users are asked to participate in a single thinking aloud test.

The test sessions are analysed and a list of problems is produced. Often, a complementary list of positives is also created. The problems are then assigned severity ratings by the facilitator and possibly other team members. The output of a thinking aloud test is a list of problems sorted in descending order of average severity.

Thinking aloud tests do not require a running system. They can be performed early in the development process with sketches and paper prototypes in order to catch potential problems as early as possible. Indeed, thinking aloud testing should be integrated into the software development process as a regular event. Krug [2009, page 23] recommends conducting a thinking aloud test once a month with three test users. A regular slot ensures that the most critical problems are identified and solved in a timely manner.

4.3.1 Defining Tasks

Krug [2009] describes a two-phase process for the creation and selection of tasks. In the first phase, an internal list of possible tasks is created and the most important ones are selected for testing:

- Critical tasks: which must work for the interface to succeed.
- Suspect tasks: which are suspected to cause users problems.
- Flagged tasks: which have been reported by the support team.

The first task is often an easy introductory task, to make the test user feel at ease. Thereafter, tasks can become longer and more involved. A typical test session lasts around 45 to 60 minutes, so an individual task should not take longer than 30 minutes. An example of a task could be: “Buying a movie ticket”.

In the second phase, the selected tasks are rewritten as scenarios. The scenarios give the tasks a context and contain further information. For the example task, the scenario description could be something like: “You would like to go to the cinema next Thursday with a friend to see the movie Avengers Endgame. Buy two tickets for the screening at 8 pm”. The scenarios are given to users one at a time on separate sheets of paper.

4.3.2 Running a Thinking Aloud Test Session

Krug [2009, Chapter 8] divides the procedure for running a thinking aloud test session into eight phases:

1. **Preparation**

   At the beginning of each test day, Krug [2009] recommends checking the following to make sure everything is set up and working properly:

   - Try out the screen-recording.
   - Try out the screen-sharing.
   - Increase the size of the mouse pointer.
   - Deactivate any unnecessary notifications.
   - Set any bookmarks needed for the test.
   - Try out the interface to be tested.
   - Reset the system to a clean state.

2. **Welcome**

   At the beginning of each thinking aloud test, the facilitator should greet the user and explain what thinking aloud testing is and how it works. To make sure nothing is forgotten, it is best to use an orientation script.

3. **Background Questions**

   Each user is typically asked some background questions before the test, in order to capture various demographic information. Through this background survey, the facilitator gets to know the test user.

4. **First Task**

   The first task is often to simply look around the interface for a few minutes. It is used to capture the first impressions of the test user and see whether the user understands what can be done with the interface. Plus, it is a good way to initiate thinking aloud.

5. **Task Scenarios**

   In this phase, the test user is given the individual task scenarios one after the other. The facilitator should read the respective task out loud and then hand over the task slip to the test user. The facilitator should also make sure that the test user thinks aloud the whole time.

6. **Interview**

   After a test user has completed the last task, it is a good opportunity to interview them about their experience. Some interview questions can be prepared in advance for all test users, other questions will arise during a test session. It is also possible to incorporate a small number of questions from observers (sent remotely).

7. **Closing**

   Afterwards, the test user should be asked if they have any final questions and then thanked for participating and given any remuneration.

8. **Housekeeping**

   Once the test user has left the room, the screen capture should be turned off and the recording saved. Finally, the facilitator should make some notes about that particular session, since specifics are very hard to remember later on.

Chapter 5

Test Procedure

This chapter describes the preparations and procedure for the thinking aloud test of the web application TreeTest. The three different user groups and the nine test users are introduced. Then, the test environment and recording are described. The training and tasks given to the users are presented. Finally, the post-test interview questions and feedback questionnaire are described.

5.1 User Profiles

The web application TreeTest has three different user groups, which were briefly described in Section 3.3:

- *Administrators*: Administrators are mainly responsible for the management of user accounts. However, they can also create and run studies and participate in studies.
- Study Owners: Study owners can create a tree test study and view the study results. They can also participate in studies.
- Study Participants: Can only participate in studies. The type of participants depends on the information hierarchy to be used in the study.

Based on the three user groups, it was decided to conduct three different thinking aloud tests: for administrators, study owners, and participants. The two users who tested administrator tasks also then participated as study owners.

5.2 Test Users

Nine test users participated in total, as can be seen in Table 5.1. Test users are referred to as TP1, TP2, etc. and are given an alias. The background information is summarised from the original completed background questionnaires, which can be found in Appendix A. The final two rows indicate which test or tests each user took part in and which device they used. Users TP1 “Jennifer” and TP5 “Chris” did both administrator tasks and study owner tasks.

Various criteria were taken into consideration for the selection of test users for the usability study. For the two administrator tests, it was important that the participants already had some experience in software development. At the time of the usability study, both test users were about to graduate with a bachelor’s degree in software development and management. This decision was made to ensure that representative test users were used for these two tests. For the study owner tests and study participant tests, attention was paid to choose test users from different areas of life and age groups to cover a broad range of potential users of the web application.

Test User	**TP1**	**TP2**	**TP3**	**TP4**	**TP5**
Alias	“Jennifer”	“Emilia”	“Robert”	“Will”	“Chris”
Date of Test	18.11.19	18.11.19	18.11.19	18.11.19	18.11.19
Time of Test	09:53	11:22	12:28	14:06	15:05
Language of Test	DE	EN	EN	EN	EN
General Information					
Sex	female	female	male	male	male
Age	22	18	20	28	23
Highest Educational Level Attained	secondary school	secondary school	secondary school	Masters in Translation	secondary school
Sight Impairment					
Sight Aid	none	none	glasses	glasses	none
Colour Blindness	no	no	no	no	no
Experience in Usability Testing					
As Test Person	yes	no	no	no	no
In Test Team	yes	no	no	no	yes
Type of Test	Thinking Aloud				Thinking Aloud
Test(s) and Device					
Participated in test(s)	administrator, study owner	study participant	study owner	study owner	administrator, study owner
Device	MacBook Pro	iPad Pro	MacBook Pro	iPad Pro	MacBook Pro

Test User	**TP6**	**TP7**	**TP8**	**TP9**
Alias	“Dave”	“John”	“Dwayne”	“Emma”
Date of Test	18.11.19	18.11.19	19.11.19	19.11.19
Time of Test	17:24	19:31	09:06	10:57
Language of Test	DE	DE	DE	DE
General Information				
Sex	male	male	male	female
Age	22	42	22	58
Highest Educational Level Attained	secondary school	secondary school	secondary school	secondary school
Sight Impairment				
Sight Aid	glasses / contact lenses	glasses	glasses	none
Colour Blindness	no	no	no	no
Experience in Usability Testing				
As Test Person	no	yes	no	no
In Test Team	no	no	no	no
Type of Test		Thinking Aloud		
Test(s) and Device				
Participated in test(s)	study participant	study owner	study participant	study participant
Device	MacBook Pro	MacBook Pro	MacBook Pro	iPad Pro

Table 5.1: Overview of the test users.

Figure 5.1: The welcome area of the test room. [Photo taken by the author of this thesis.]

5.3 Test Environment

For the thinking aloud tests, the test room was separated into a welcome area, shown in Figure 5.1, and a test area, shown in Figure 5.2. The welcome area was used to welcome the test users and to inform them why they were invited to participate in the usability study. During the welcome phase, the test users were informed about thinking aloud testing and tree testing. They were then asked to fill out the background questionnaire. Finally, users were informed that the usability test would be recorded on video and audio and they were asked to sign a consent form.

Subsequently, the test users were asked to take a seat in the test area of the room for the usability test. Before the thinking aloud test began, a training session was held with each test user to prepare them in the best possible way for the thinking aloud test. This is described in more detail in Section 5.5. Following the training, the test users were given the individual tasks one after the other. After a test user had completed all tasks, a final interview was conducted with the test user. Some of the questions were defined in advance for all test users. Other questions arose during the thinking aloud test. Finally, the test user was asked to fill out the feedback questionnaire (see Section 6.7). Figure 5.3 shows a test user during the thinking aloud test. Table 5.2 shows the exact location and date where the thinking aloud tests took place.

Six of the users performed the test with a laptop (MacBook Pro) and three of the users with a tablet (iPad Pro), in order to test the interface with both kinds of device. The details of the two devices can be found in Table 5.3.

Figure 5.2: The test area of the test room. [Photo taken by the author of this thesis.]

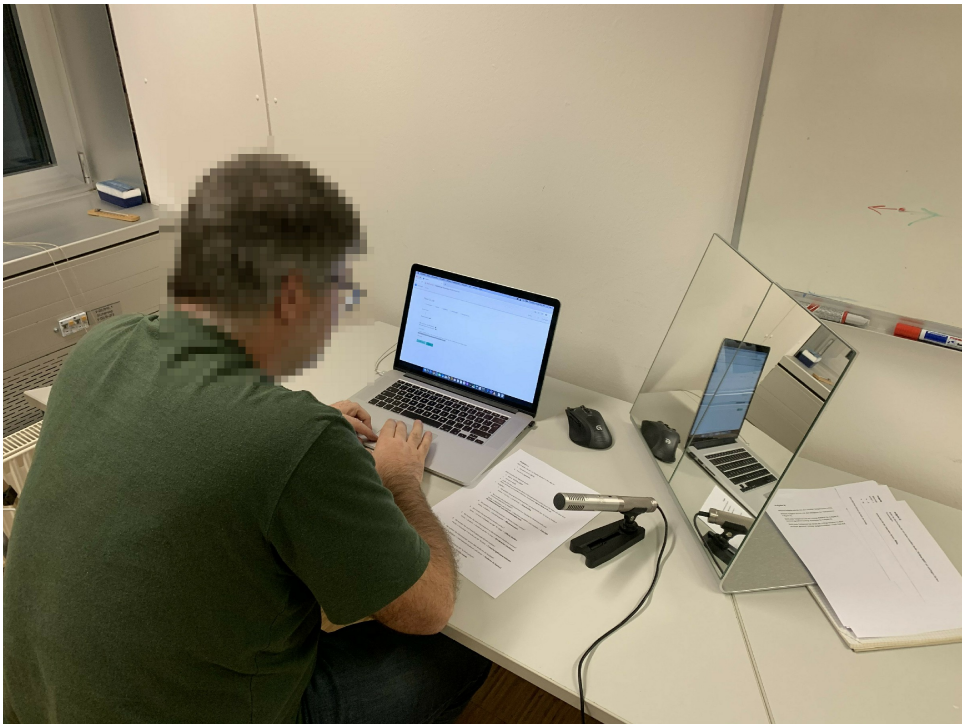

Figure 5.3: A test user while participating in the thinking aloud test. [Photo taken by the author of this thesis.]

Test Setting	
Location	Meeting Room ID01184(D2.21), Inffeldgasse 16c, Graz
Date of Pilot Test	2019-11-18
Date of Real Tests	2019-11-18 and 2019-11-19

Table 5.2: Location and Date.

Equipment		
Device	Apple MacBook Pro 15 Mid 2015	Apple iPad Pro 12.9 3.Generation
Operating System	macOS Catalina 10.15	iPadOS 13.3
Web Browser	Google Chrome 78.0.3904.97	Google Chrome 78.0.3904.84
Screen Capture Software	QuickTime Player	iPadOS 13.3 Screen Capture
Internet Connection	TU GRAZ eduroam	
Screen Resolution	2880x1800	2732x2048
Screen Size	15 IPS	12,9 IPS

Table 5.3: Equipment used.

Test Recording	
Digital Camcorder	Sanyo Xacti HD1010
Microphone	Philips SBC ME570
Tripod	Hama Profil 74

Table 5.4: External Recording Equipment.

5.4 Test Recording

The video and audio equipment used for external recording is shown in Table 5.4. The QuickTime Player application from Apple was used to record the MacBook Pro's screen. For the iPad Pro's screen recording, a function of the operating system iPadOS was used. Apple's iMovie application was used to edit all recordings, both external and internal.

5.5 Training

At the beginning of each thinking aloud test, a video of a thinking aloud test was shown. Additionally, an unrelated task was demonstrated by the facilitator whilst thinking aloud, to ensure that the test user understood the process. For the demonstration, the facilitator opened Google, searched for "cineplex.at", selected a film, and reserved two seats.

5.6 Tasks

Based on the three user groups, three different task sets were created. The various tasks are introduced in the following. For all tasks, the hierarchical information structure of the external web site of Graz University of Technology [TUG 2020] was used.

5.6.1 Tasks for Administrators

Table 5.5 shows the internal task list for administrators. The task slips given to the participants can be found in Appendix B.1.

5.6.2 Tasks for Study Owners

Table 5.6 shows the internal task list for study owners. The task slips given to the participants can be found in Appendix B.2.

Task No.	Description	Prerequisites	Completion Criteria	Possible Solution Path
1	Please log in to the application as an administrator and create a new study owner account and activate it.	Web browser opened at TreeTest's login-page.	The user has created and activated the study owner account.	Login → Admin Panel → Add a new user → Add → Enable
2	Please change the password of user "James" to "9hberf".	Web browser opened at TreeTest's Administrator Panel.	The user has changed the password.	Change Password → Change Password

Table 5.5: The tasks used for administrators.

5.6.3 Tasks for Study Participants

Table 5.7 shows the internal task list for study participants. The actual task slip given to the participants can be found in Appendix B.3.

Task No.	Description	Prerequisites	Completion Criteria	Possible Solution Path
1	Please log in to the application as study owner and create a study.	Web browser opened at TreeTest's login-page.	The test user has successfully created a study.	Login → Create Study → Next Step → Import CSV → Add Task (and select answer)*8 → Next Step → Next Step → here
2	Please call up the preview of the study you have created and follow the instructions.	Web browser opened at TreeTest's studies-page.	The test user has completed the preview of the study.	Preview → Continue → Start Task (and select something or skip)*8 → Send Feedback
3	Please change the fourth question of the study you created. Additionally, launch the study.	Web browser opened at TreeTest's studies-page.	The test user has successfully edited the study.	Edit → 3. Tasks → (select question and select answer) → Save → 5. Finish → here → Launch
4	Please call up the results of the study "TU Graz" and analyse the results.	Web browser opened at TreeTest's studies-page.	The test user has analysed the results.	Results → (The test user can now move freely around the web application)
5	Questions about the study results.	Web browser opened at TreeTest's results-page.	The test user has answered all questions right or wrong.	Question 1: Task Analysis, Question 2: Task Analysis, Question 3: Destinations, Question 4: Task Analysis → Path Tree

Table 5.6: The tasks used for study owners.

Task No.	Description	Prerequisites	Completion Criteria	Possible Solution Path
1	Please participate in the study and follow the instructions.	Web browser opened at TreeTest's participation in a study-page.	The test user participated in the study.	Continue → Start Task (select something or skip)*8 → Send Feedback

Table 5.7: The tasks used for study participants.

5.7 Interview Questions

After each test user completed all of the tasks, they were interviewed and asked some final questions. The opening question is designed to elicit the immediate reactions of the test user:

- "How was it?"

Then, four pre-planned questions were asked of every test user:

- "Did anything strike you as particularly bad?"
- "Did anything strike you as particularly good?"
- "How easy was it to create a study?"
- "Would you change anything on the webapp?"

Finally, any specific questions which arose during the test were asked. Some examples included:

- "Would you change anything on the test procedure?"
- "Would you change how the questions are presented?"
- "You had problems creating a task for your study. What would you do to make this step easier?"
- "You had problems coming back to the main page, what would you change?"
- "At the end you had problems understanding the results, what would you change to make it more understandable?"

5.8 Feedback Questionnaire

After the interview, the test users were asked to fill out a feedback questionnaire. The completed feedback questionnaires of the test users can be found in Appendix C. The summary of all feedback questionnaire can be found in Section 6.7.

Chapter 6

Test Results

In this chapter, the results of the thinking aloud tests of the web application TreeTest are presented. The task completion rate, positive findings, and negative findings (problems) resulting from the usability study are discussed. Furthermore, this chapter contains a summary of the feedback questionnaires.

6.1 Discussion and Analysis

In the beginning, it should be said that all test users were able to complete all tasks successfully. However, some of the test users needed some assistance. From this, it can be concluded that the web application is user-friendly and only minor changes are necessary. However, some problems were identified by the test users.

Unfortunately, some study owner test users had problems creating their first study task. Most of them first tried to select a node in the tree and then tried to define the task. This is unfortunately not supported by the web application. Furthermore, the button for creating a task is very small and not well-positioned. The test users would probably not have had this problem if the tree is only displayed after the task has been defined. Alternatively, if the button to add a task was better positioned. Another problem that several test users had was understanding the study results. Unfortunately, the explanatory texts only helped the test users to a limited extent. Furthermore, the question marks icons for the explanatory texts are not clickable, which was criticised by several test users. It is necessary to keep the mouse pointer over the question mark icon until the information is displayed. None of the test users had this patience. Explanatory texts should be formulated more clearly, and the question mark icons should react to mouse clicks.

Regarding the administrator test users, one test user complained that passwords are displayed in plain text. Apart from that, no other problems with the administrator tasks were found.

None of the study participant test users had problems when participating in studies. However, the test users noted that the design was not appealing, and the process of participation was too monotonous.

In summary, it can be said that the web application works very well except for some minor problems. If the problems are solved, and the design is updated, this application can be seen as an excellent alternative to existing applications.

6.2 Task Completion

Table 6.1 administrator, Table 6.2 study owner and Table 6.3 study participant show how many test users completed each task. Here, 1 means that the test user has completed the respective task successfully. A * means that minor assistance was given. 0 means task could not be completed successfully.

Administrator	Task 1	Task 2
TP1	1	1
TP5	1	1
Total	2	2
Percent	100	100

Table 6.1: Task completion rates for administrator test users. An asterisk (*) indicates that assistance was given.

Study Owner	Task 1	Task 2	Task 3	Task 4	Task 5
TP1	1	1	1	1	1*
TP3	1*	1	1	1	1
TP4	1*	1	1	1	1*
TP5	1*	1	1	1	1*
TP7	1*	1	1	1	1*
Total	5	5	5	5	5
Percent	100	100	100	100	100

Table 6.2: Task completion rates for study owner test users. An asterisk (*) indicates that assistance was given.

Study Participant	Task 1
TP2	1
TP6	1
TP8	1
TP9	1
Total	4
Percent	100

Table 6.3: Task completion rates for study participant test users. An asterisk (*) indicates that assistance was given.

Positivity	Meaning
4	Extremely Positive
3	Major Positive
2	Minor Positive
1	Cosmetic Positive
0	Not a Positive

Table 6.4: The positivity rating scale.

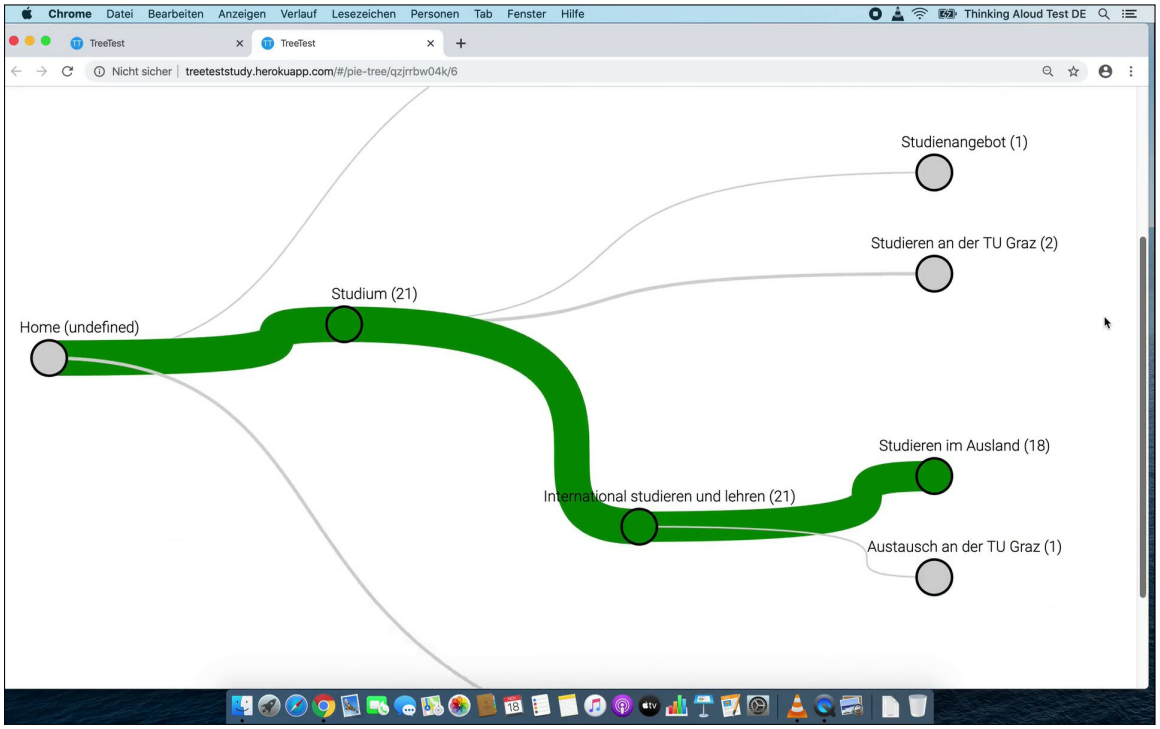

Figure 6.1: Positive 1: The path tree showing a graphical overview of paths taken for a particular task.

6.3 Top Three Positive Findings

All positive findings were assigned a positivity rating according to the scale shown in Table 6.4. The positive findings were then sorted in descending order of positivity. A list of all positive findings can be found in Section 6.4. The three most positive findings, according to the positivity rating, are presented in more detail in the following.

6.3.1 Positive 1: Path Tree

Figure 6.1 shows a screenshot of the path tree, which was described by TP1 as very useful. The path tree shows the paths taken by the study participants.

6.3.2 Positive 2: Path Tree Downloadable

TP1 and TP5 liked the ability to download the path tree, see Figure 6.2. It is possible to download the path tree as SVG files.

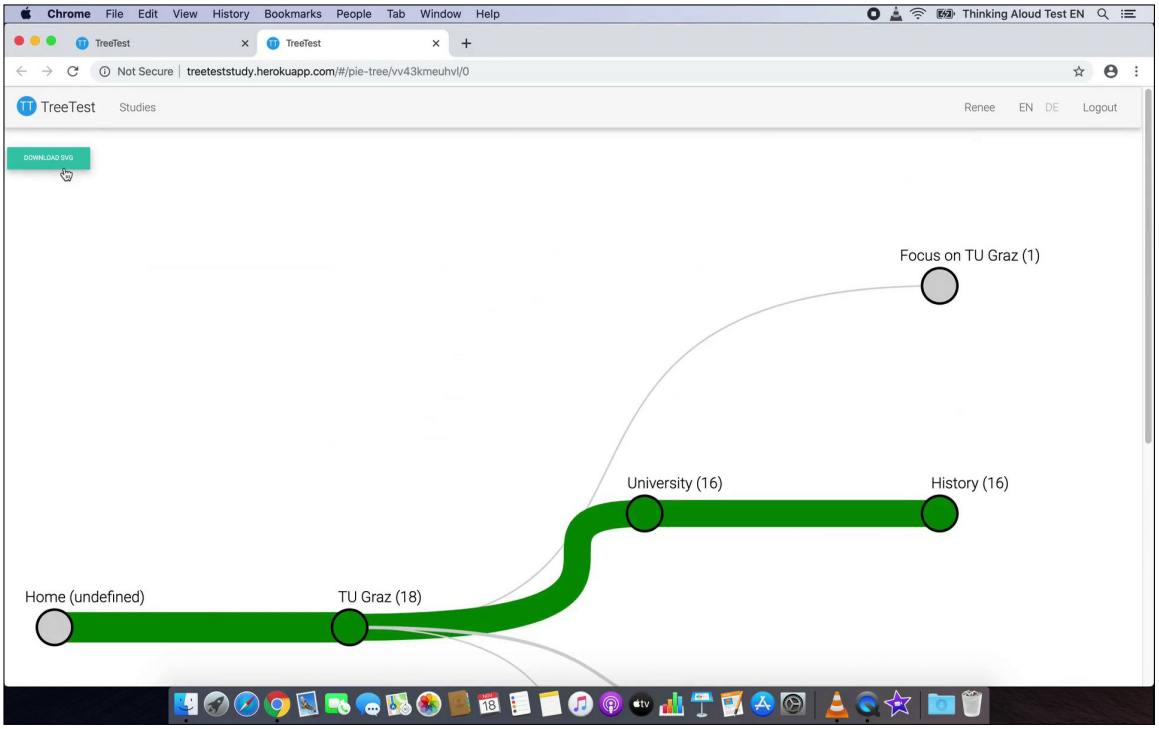

Figure 6.2: Positive 2: It is possible to download the path tree as SVG.

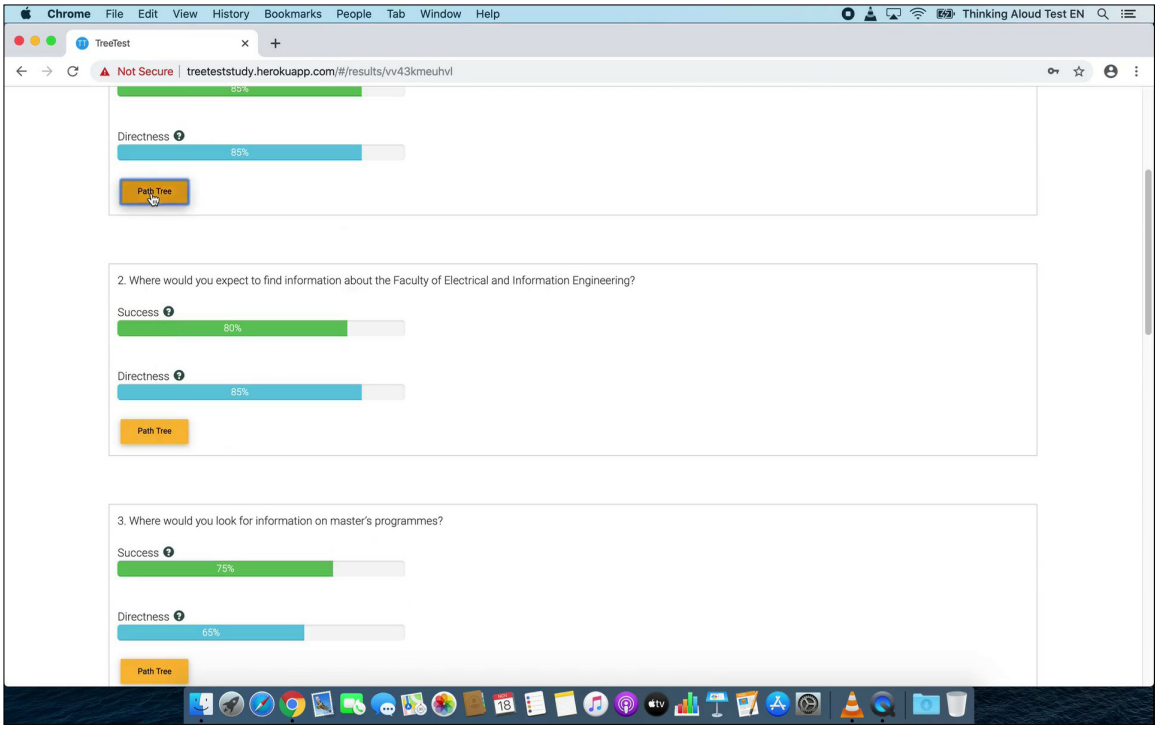

Figure 6.3: Positive 3: The path tree opens in a new browser tab.

No.	Title	Description	Found by	Location (how reproducible?)	Positivity
1	Path Tree	The path tree represents the results of the study.	TP1	Home → Studies → Results → Task Analysis → Path Tree	4
2	Path Tree Downloadable	It is possible to download the path tree as an SVG file.	TP1, TP5	Home → Studies → Results → Task Analysis → Path Tree	3
3	Path Tree Opens In A New Tab	Path tree opens in a new tab.	TP3	Home → Studies → Results → Task Analysis → Path Tree	2
4	Tree Resets After Each Question	After each question, the tree is reset while participating in a study.	TP1	Participating in a study	1

Table 6.5: List of all positive findings in descending order of positivity (most positive first).

Severity	Meaning
4	Catastophic problem
3	Serious problem
2	Minor problem
1	Cosmetic problem
0	Not a problem

Table 6.6: The severity rating scale.

6.3.3 Positive 3: Path Tree Opens in a New Browser Tab

TP3 likes that the path tree is opened in a new browser tab, see Figure 6.3.

6.4 List of All Positive Findings

Table 6.5 shows all four positive findings in descending order of positivity.

6.5 Top Five Problems

All problems were assigned a severity rating according to the scale shown in Table 6.6. The problems were then sorted in descending order of severity. A list of all problems can be found in Section 6.6. The five most significant problems, according to the severity rating, are explained in the following.

6.5.1 Problem 1: User Interface Selection Error

TP1 noticed a user interface error, as shown in Figure 6.4. Sometimes the user interface for creating a task is not selected as expected. This leads to difficulties in selecting the answer in the hierarchical structure. It is necessary to click a second time to select the correct user interface.

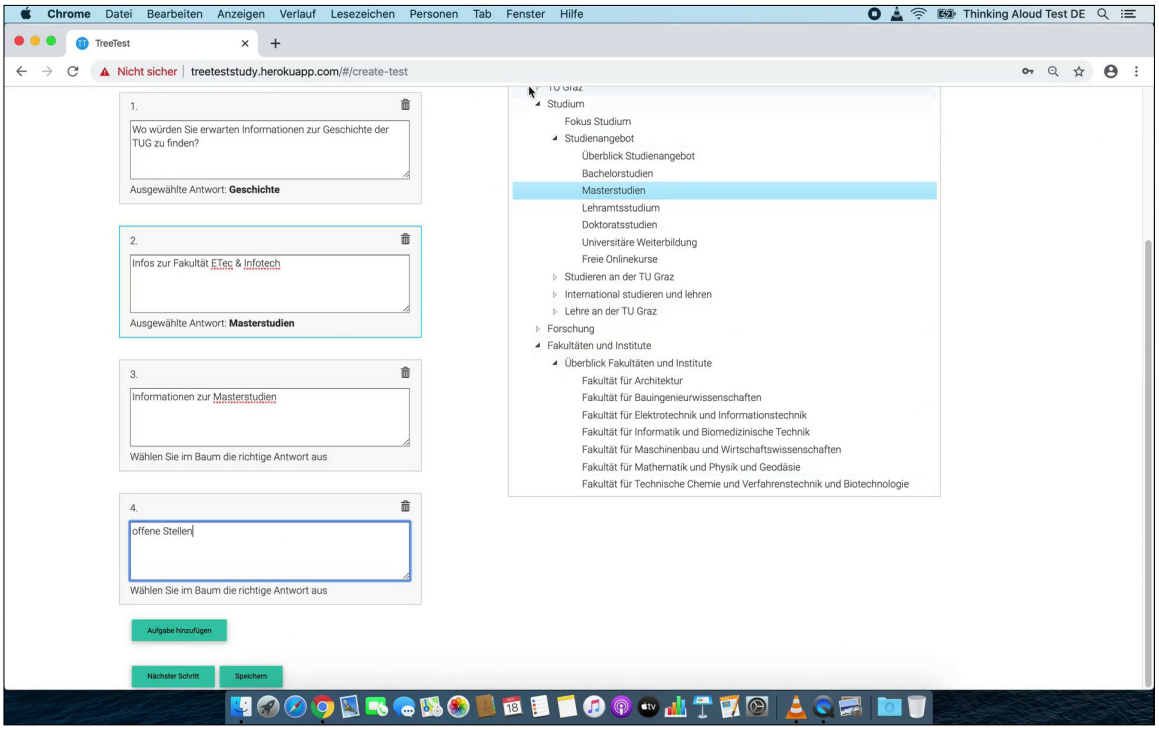

Figure 6.4: Problem 1: Sometimes the user interface for creating a task is not selected.

6.5.2 Problem 2: Question Mark Icons

TP1, TP4, TP5 and TP6 had problems with the question mark icons, as shown in Figure 6.5. The question mark icons react very slowly, and it is not possible to click on them. Furthermore, the information is only available in English.

6.5.3 Problem 3: Save Button Without Feedback

TP3, PT4 and PT5 have noticed that no confirmation is displayed when saving a study. This is shown in Figure 6.6. The save button by creating a study does not inform the user if saving the study was successful.

6.5.4 Problem 4: Unintentional Tree Modification

TP4 noticed that scrolling on an iPad causes the tree to be modified, as shown in Figure 6.7. When creating tasks, the hierarchical information structure can be changed by scrolling.

6.5.5 Problem 5: Web Application Changes Language

TP1 discovered that the web application changes its language when a user signs out, as shown in Figure Figure 6.8. Reloading or logging out of the application changes the language to English.

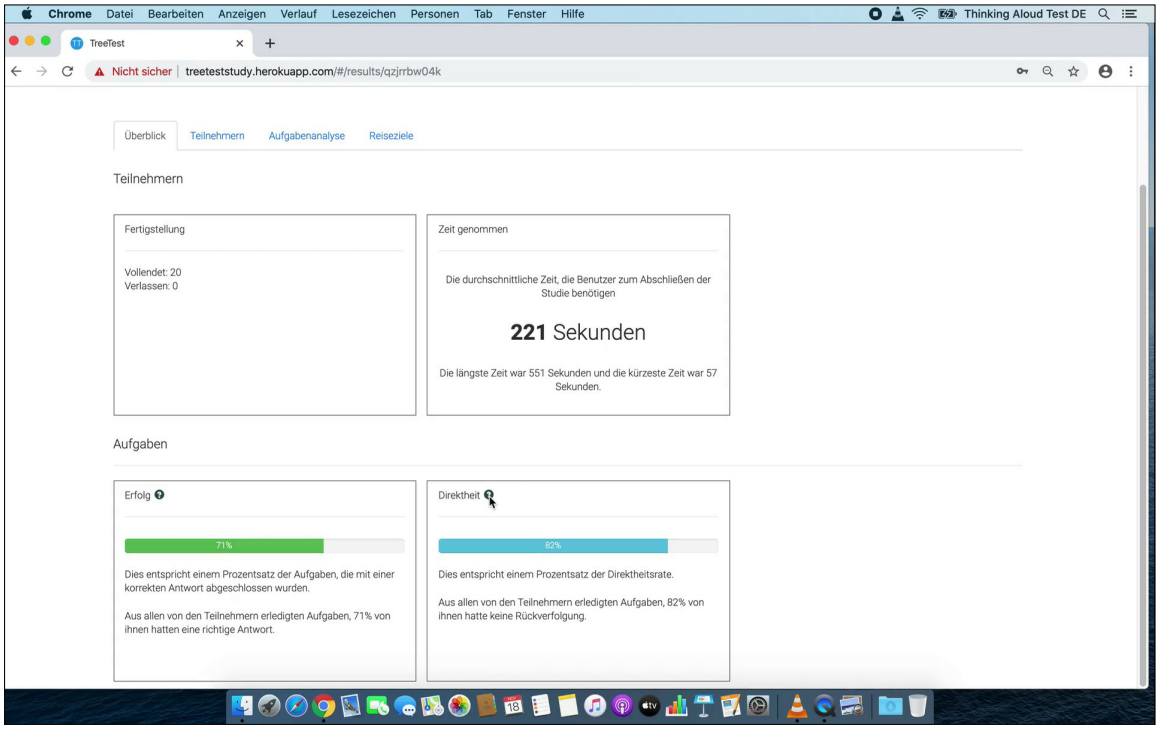

Figure 6.5: Problem 2: The question mark icons do not work as expected.

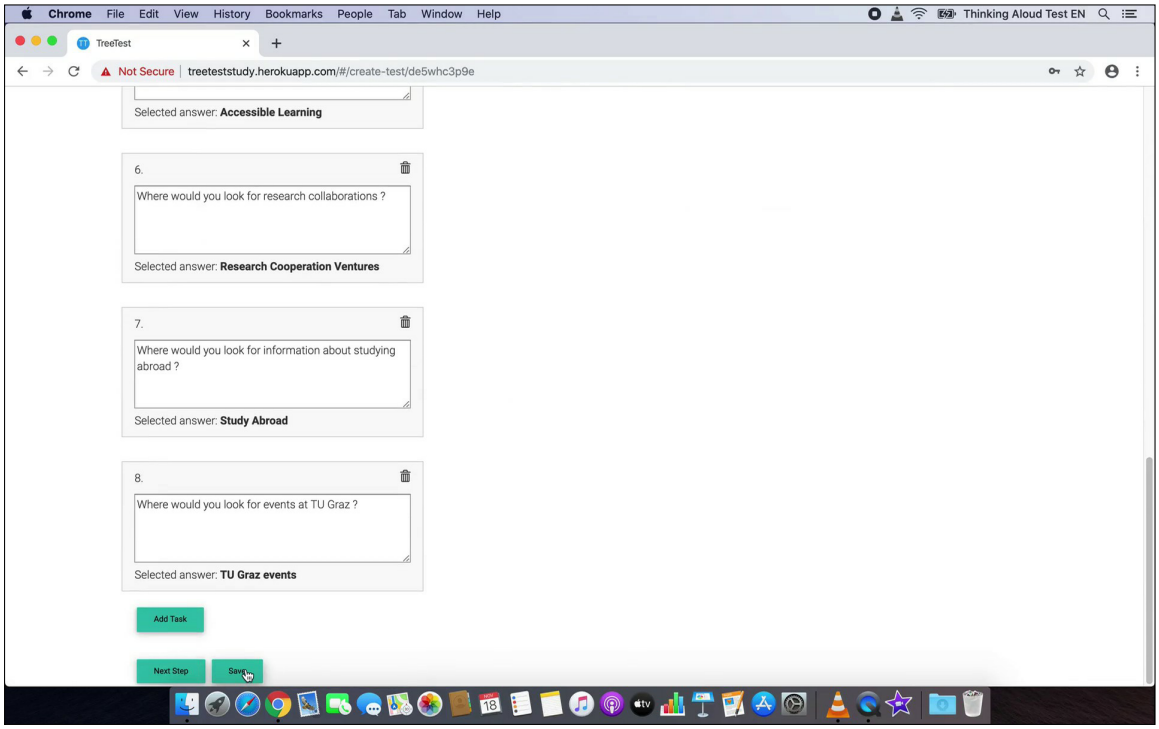

Figure 6.6: Problem 3: The Save button does not provide any confirmation or feedback.

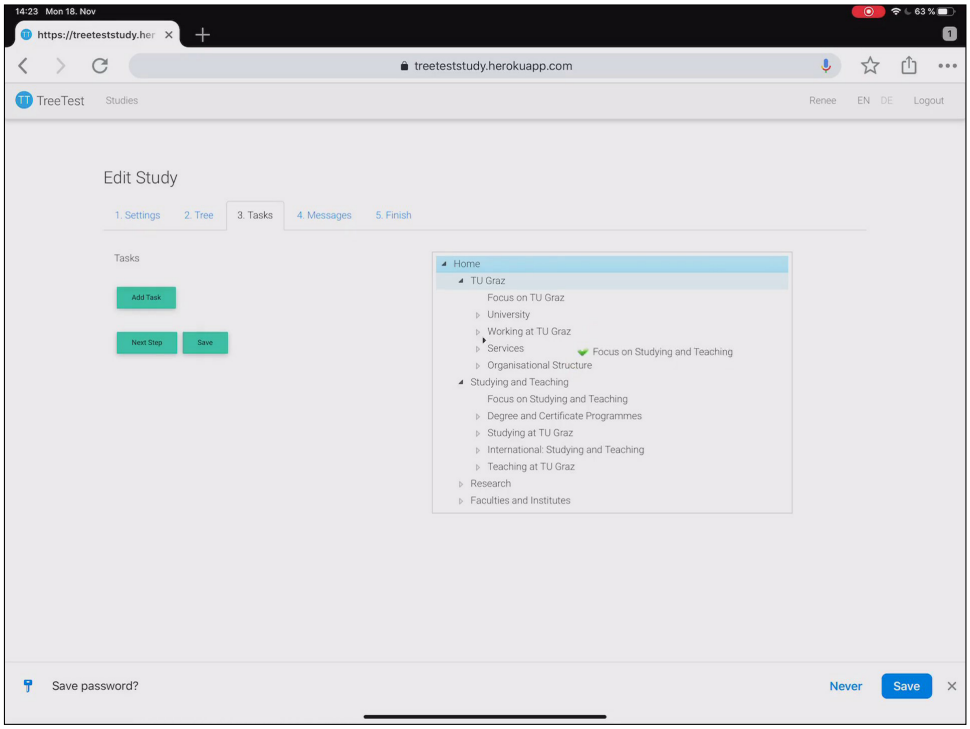

Figure 6.7: Problem 4: Scrolling on an iPad causes the tree to be modified unintentionally.

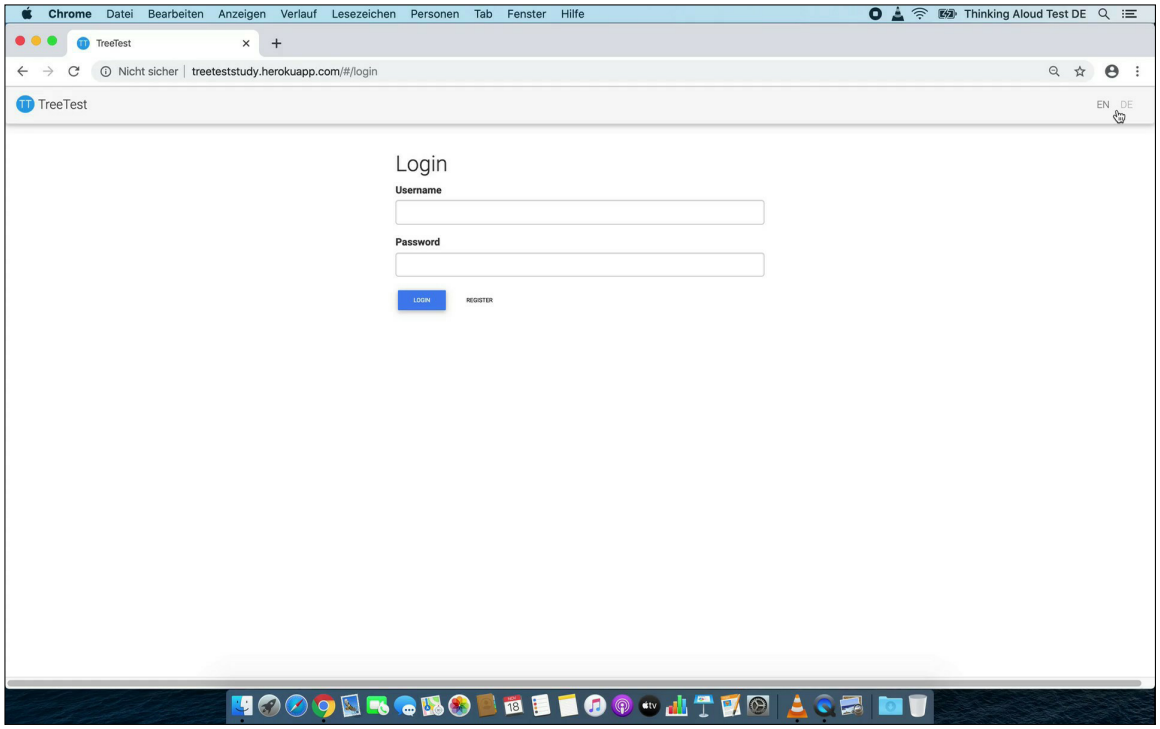

Figure 6.8: Problem 5: The web application changes its language when a user signs out.

6.6 List of All Problems Found

Table 6.7 shows all fifteen problems found in descending order of severity.

No.	Title	Description	Found by	Location (how reproducible?)	Severity
1	User Interface Selection Error	Sometimes the user interface for creating a task is not selected as expected. This leads to difficulties in selecting the answer in the hierarchical structure. It is necessary to click a second time to select the correct user interface.	TP1	Studies → Create Study → 3. Tasks	3
2	Question Mark Icon Does Not Work As Expected	The question mark icons react very slowly, and it is not possible to click on them. Furthermore, the information is only available in English.	TP1, TP4, TP5, TP6	Any question mark, for example, the question marks on the results page.	3
3	Save Button Without Feedback	The save button by creating a study does not inform the user if saving the study was successful.	TP3, TP4, TP5	Studies → Create Study	3
4	Unintentional Tree Modification	Scrolling on an iPad causes the tree to be modified. When creating tasks, the hierarchical information structure can be changed by scrolling.	TP4	Studies → Create Study → 3. Tasks	3
5	Web Application Changes Language	Reloading or logging out of the web application changes the language to English.	TP1	Anywhere in the web app	2
6	TreeTest Icon Not Clickable	Usually, the icon of a web site brings the user back to the main page when clicked. Unfortunately, this application does not support this.	TP3, TP7	Anywhere in the web app	2
7	Destination Table Very Long	Unfortunately, the destination tables are very long.	TP4, TP5	Studies → Results → Destinations	2
8	Password In Clear Text	Passwords are stored/displayed in plain text.	TP1, TP5	Administrator Panel	2
9	Tree Move	When many study tasks are created, the page becomes very long. Unfortunately, the tree has a fixed position. Therefore it is necessary to scroll frequently.	TP5	Studies → Create Study → 3. Tasks	2
10	Answer Before Question	Some users first looked for the answer in the path tree and then wanted to create the question. Unfortunately, the application does not support this.	TP3, TP4, TP5, TP7	Studies → Create Study → 3. Tasks	1

11	English Text On German Page	A pop-up message has English text on a German page. The path tree is only available in English.	TP1	Administrator Panel and Studies → Results → Task Analysis	1
12	Spelling Mistake	The German word “Fertig machen” is misspelled “Vertig machen”.	TP1	Studies → Create Study	1
13	Skipped Tasks Do Not Come Again	When a question is skipped during participation in a study, the question is not asked again at the end of the study.	TP1	Participating in a study	1
14	Example Text Must Be Deleted Manually	The example text for creating a task “Where would you expect to find...?” must be deleted manually, it should behave like a hint.	TP1	Studies → Create Study → 3. Tasks	0
15	Tree Does Not Reset Itself	When tasks are created, the tree is not reset for each additional task added.	TP1	Studies → Create Study → 3. Tasks	0

Table 6.7: List of all problems found in descending order of severity (most severe first).

6.7 Feedback Questionnaires

After the thinking aloud tests, test users fill out a feedback questionnaire. Two different feedback questionnaires were used, depending on the type of test user. The first questionnaire was used for administrator test users and study owner test users. The second questionnaire was used for study participant test users. A neutral 7-point scale `3 2 1 0 1 2 3` was used on the feedback questionnaire for users to circle. For the analysis of the feedback questionnaires, the neutral scale was converted into points from 0 (worst) to 6 (best).

Table 6.8 shows a summary of the feedback questionnaires from administrators and study owners. Table 6.9 shows a summary of the feedback questionnaires of the study participants. The completed feedback questionnaires of the individual participants can be found in Appendix C.

No.		TP1	TP3	TP4	TP5	TP7	Mean	Std Dev
1.	The app's user interface is clear and logical?	6	6	4	5	4	5.00	0.89
2.	Text is easy to read?	1	6	6	6	4	4.60	1.96
3.	Appearance of app, including colours and graphics?	5	6	6	6	1	4.80	1.93
4.	Getting to the right part of the app is easy?	5	6	5	5	3	4.80	0.97
5.	Consistency of app?	6	6	5	5	2	4.80	1.47
6.	Response speed of app?	5	6	6	6	6	5.80	0.40
7.	How well does the app guide you through the process of conducting a study?	6	5	4	4	3	4.40	1.02
8.	How easy is it to create a study?	6	6	4	5	5	5.20	0.75
9.	How easy is it to create/upload a tree?	6	6	4	6	5	5.40	0.80
10.	How easy is it to create tasks?	4	6	3	5	5	4.60	1.02
11.	How easy is it to navigate through the tree to select the correct answer?	4	5	2	5	6	4.40	1.36
12.	How well are the results of a study presented?	5	6	3	4	3	4.20	1.17
13.	How useful is the path tree?	3	6	3	5	3	4.00	1.26
14.	This app cares about my satisfaction as a customer?	3	4	3	4	2	3.20	0.75
15.	Do you consider this app to be trustworthy?	4	3	4	6	3	4.00	1.10
16.	How relevant is the functionality of this app to you?	2	3	1	1	4	2.20	1.17
17.	Would you consider using this app in your own work?	1	5	1	1	1	1.80	1.60
18.	Overall impression of the app?	4	6	4	4	2	4.00	1.26

Table 6.8: Summary of ratings by administrators and study owners from the feedback questionnaire.

No.		TP2	TP6	TP8	TP9	Mean	Std Dev
1.	The app's user interface is clear and logical?	4	5	5	6	5.00	0.71
2.	Text is easy to read?	4	6	2	6	4.50	1.66
3.	Appearance of app, including colours and graphics?	5	0	0	6	2.75	2.77
4.	Consistency of app?	6	5	5	6	5.50	0.50
5.	Response speed of app?	6	5	6	6	5.75	0.43
6.	How well did the app guide you through the test?	6	0	4	5	3.75	2.28
7.	How easy is it to navigate through the tree?	5	4	4	6	4.75	0.83
8.	This app cares about my satisfaction as a user?	5	4	2	6	4.25	1.48
9.	Do you consider this app to be trustworthy?	4	1	1	6	3.00	2.12
10.	How relevant is the functionality of this app to you?	4	5	5	5	4.75	0.43
11.	Would you consider using this app in your own work?	4	1	3	6	3.50	1.80
12.	Overall impression of the app?	6	1	3	6	4.00	2.12

Table 6.9: Summary of ratings by study participants from the feedback questionnaire.

Appendix A

Background Questionnaires

In Section 5.2, the test users were already introduced with a shortened version of the background questionnaires. In this section, the test users are presented in a non-shortened version. The format and content of the background questionnaires were taken from the course material of the course Human-Computer Interaction by Andrews [2019a]. The domain-specific questions were added.

Datum: 18.11.19 Uhrzeit: 9:53 Test User ID: 1

Hintergrundbefragung

1. Allgemeine Informationen

Geschlecht: [ ] *männlich* [X] *weiblich*

Alter: 22

Beruf: Student (SEW)

2. Sehvermögen

1. Verwenden Sie eine Sehhilfe bei der Arbeit am Computer?

[X] *Keine* [ ] *Brille* [ ] *Kontaktlinsen* [ ] *sonstige* _____

2. Sind Sie farbenblind?

[X] *Nein* [ ] *Ja, und zwar* _____

3. Ausbildung

1. Höchste abgeschlossene Ausbildung:

[ ] *Lehre* [X] *Matura* [ ] *Studium* [ ] *Doktorat*

2. Wenn Sie studieren oder studiert haben, beschreiben Sie bitte Ihr Hauptstudiengebiet:

SEW

4. Verwendung eines Computers

1. Welche Art von Personal Computer verwenden Sie am häufigsten?

[ ] *Keinen* [X] *Microsoft Windows* [ ] *Apple Macintosh* [ ] *Unix* [X] *sonstige* Linux

2. Wie lange benutzen Sie bereits einen Personal Computer?

13 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie einen Personal Computer?

40 *Stunden*

5. Verwendung eines Tablets

1. Welche Art von Tablet verwenden Sie am häufigsten?

[ ] Keines [X] Android [ ] Apple (iOS) [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Tablet?

3 Jahre

3. Wie viele Stunden pro Woche verwenden Sie ein Tablet?

4 Stunden

6. Verwendung eines Smartphones

1. Welche Art von Smartphone verwenden Sie am häufigsten?

[ ] Keinen [X] Android [ ] Apple (iOS) [ ] Windows Phone [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Smartphone?

13 Jahre

3. Wie viele Stunden pro Woche verwenden Sie ein Smartphone?

20 Stunden

7. Verwendung des Web

1. Wie viele Stunden pro Woche benutzen Sie das World Wide Web?

50 Stunden

2. Welches Gerät verwenden Sie am häufigsten zum surfen?

[ ] Desktop PC [X] Laptop [ ] Tablet
[ ] Smartphone [ ] sonstige _____

3. Welche Art von Internet Zugang verwenden Sie normalerweise?

[ ] xDSL [ ] Kabel Modem [X] Glasfaser
[ ] 3G mobiles Internet [ ] LTE mobiles Internet [ ] sonstige _____

4. Welchen Web-Browser verwenden Sie normalerweise?

[ ] Chrome [X] Firefox [ ] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] sonstige _____

5. Haben Sie Erfahrung als Web Site Administrator?

Wenn ja, wie lange? _____ *Jahre* X

8. Domänenspezifische Fragen

1. Haben Sie schon einmal eine Informationshierarchie entworfen?
(z.B. Produktkategorien und Unterkategorien, hierarchische Menüstruktur, Website-Navigationshierarchie)

[X] Nein [ ] Ja, _____

2. Haben Sie schon einmal als Testperson an einer Studie zum Testen einer Informationshierarchie teilgenommen?

[X] Nein [ ] Ja, _____

3. Haben Sie schon einmal ein Online-Tool zur Durchführung einer Nutzerstudie benützt? Wenn ja, welche(s)?

[ ] Nein [X] Ja, _____

9. Erfahrung mit Usability Tests

1. Haben Sie schon an einer Usability Studie teilgenommen?

[X] als Testperson [X] als Mitglied des Testteams

Wenn ja, was war das für eine Studie?

[X] Thinking Aloud [ ] Formal Experiment [ ] sonstige _____

Date: 18.11.19 Time: 11:22 Test User ID: 2

Background Questionnaire

1. General Information

Sex: [ ] male [X] female

Age: 18

Occupation: Student (SEW)

2. Sight Impairment

1. Do you use a sight aid when working on the computer?

[X] none [ ] glasses [ ] contact lenses [ ] other _____

2. Do you have any form of colour blindness?

[X] no [ ] yes, _____

3. Education

1. Highest Educational Level Attained:

[ ] vocational training [X] secondary school [ ] university degree [ ] doctorate

2. If you are studying or have studied, please describe your main area of study:

SEW

4. Personal Computer Use

1. Which kind of personal computer do you use most?

[ ] None [X] Microsoft Windows [ ] Apple Macintosh [ ] Unix [ ] other _____

2. How long have you been using a personal computer?

10 years

3. How many hours per week do you use a personal computer?

35 hours

5. Tablet Use

1. Which kind of tablet do you use most?

[ ] None [ ] Android [x] Apple (iOS) [ ] other _____

2. How long have you been using a tablet?

3 years

3. How many hours per week do you use a tablet?

5 hours

6. Smartphone Use

1. Which kind of smartphone do you use most?

[ ] None [ ] Android [x] Apple (iOS) [ ] Windows Phone [ ] other _____

2. How long have you been using a smartphone?

5 years

3. How many hours per week do you use a smartphone?

30 hours

7. Web Use

1. How many hours per week do you use the World Wide Web?

63 hours

2. Which kind of device do you use most often to surf the web?

[ ] desktop computer [x] laptop [ ] tablet
[ ] smartphone [ ] other _____

3. What kind of internet connection do you normally use?

[ ] xDSL [ ] cable modem [ ] fibre optic
[ ] 3G mobile internet [x] LTE mobile internet [ ] other _____

4. Which web browser do you normally use?

[x] Chrome [ ] Firefox [ ] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] other _____

5. Do you have experience as a web site administrator?

If yes: 1 *years*

8. Domain-Specific Questions

1. Have you ever designed an information hierarchy?
(e.g. product categories and sub-categories, hierarchical menu structure, website navigation hierarchy)

[ ] no [x] yes, Web

2. Have you ever participated as a test user in a study to test an information hierarchy?

[x] no [ ] yes, _____

3. Have you ever used an online tool to carry out a user study? If so, which one(s)?

[x] no [ ] yes, _____

9. Experience with Usability Tests

1. Have you participated in any kind of usability study before?

[ ] as a test user [ ] as part of the test team

If yes, what kind of study was it?

[ ] Thinking Aloud [ ] Formal Experiment [ ] other _____

Date: 18.11.19 Time: 12:28 Test User ID: 3

Background Questionnaire

1. General Information

Sex: [x] male [ ] female

Age: 20

Occupation: Student

2. Sight Impairment

1. Do you use a sight aid when working on the computer?

[ ] none [x] glasses [ ] contact lenses [ ] other _____

2. Do you have any form of colour blindness?

[x] no [ ] yes, _____

3. Education

1. Highest Educational Level Attained:

[ ] vocational training [x] secondary school [ ] university degree [ ] doctorate

2. If you are studying or have studied, please describe your main area of study:

SEW

4. Personal Computer Use

1. Which kind of personal computer do you use most?

[ ] None [ ] Microsoft Windows [ ] Apple Macintosh [x] Unix [ ] other _____

2. How long have you been using a personal computer?

12 years

3. How many hours per week do you use a personal computer?

40 hours

5. Tablet Use

1. Which kind of tablet do you use most?

[ ] None [ ] Android [ ] Apple (iOS) [ ] other _____

2. How long have you been using a tablet?

_____ *years*

3. How many hours per week do you use a tablet?

_____ *hours*

6. Smartphone Use

1. Which kind of smartphone do you use most?

[ ] None [x] Android [ ] Apple (iOS) [ ] Windows Phone [ ] other _____

2. How long have you been using a smartphone?

10 *years*

3. How many hours per week do you use a smartphone?

20-30 *hours*

7. Web Use

1. How many hours per week do you use the World Wide Web?

40 *hours*

2. Which kind of device do you use most often to surf the web?

[ ] desktop computer [x] laptop [ ] tablet
[ ] smartphone [ ] other _____

3. What kind of internet connection do you normally use?

[ ] xDSL [x] cable modem [ ] fibre optic
[ ] 3G mobile internet [ ] LTE mobile internet [ ] other _____

4. Which web browser do you normally use?

[x] Chrome [x] Firefox [ ] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] other _____

5. Do you have experience as a web site administrator?

If yes: ____ years

8. Domain-Specific Questions

1. Have you ever designed an information hierarchy?
(e.g. product categories and sub-categories, hierarchical menu structure, website navigation hierarchy)

[x] no [ ] yes, _____

2. Have you ever participated as a test user in a study to test an information hierarchy?

[x] no [ ] yes, _____

3. Have you ever used an online tool to carry out a user study? If so, which one(s)?

[x] no [ ] yes, _____

9. Experience with Usability Tests

1. Have you participated in any kind of usability study before?

[ ] as a test user [ ] as part of the test team

If yes, what kind of study was it?

[ ] Thinking Aloud [ ] Formal Experiment [ ] other _____

Date: 18.11.19 Time: 14:06 Test User ID: 4

Background Questionnaire

1. General Information

Sex: [x] male [ ] female

Age: 28

Occupation: Student

2. Sight Impairment

1. Do you use a sight aid when working on the computer?

[ ] none [x] glasses [ ] contact lenses [ ] other _____

2. Do you have any form of colour blindness?

[x] no [ ] yes, _____

3. Education

1. Highest Educational Level Attained:

[ ] vocational training [ ] secondary school [x] university degree [ ] doctorate

2. If you are studying or have studied, please describe your main area of study:

Translate: English - Deutsch

4. Personal Computer Use

1. Which kind of personal computer do you use most?

[ ] None [x] Microsoft Windows [ ] Apple Macintosh [ ] Unix [ ] other _____

2. How long have you been using a personal computer?

20 years

3. How many hours per week do you use a personal computer?

10 hours

5. Tablet Use

1. Which kind of tablet do you use most?

[ ] None [ ] Android [x] Apple (iOS) [ ] other _____

2. How long have you been using a tablet?

6 * *years*

* asked later

3. How many hours per week do you use a tablet?

4 * *hours*

6. Smartphone Use

1. Which kind of smartphone do you use most?

[ ] None [ ] Android [x] Apple (iOS) [ ] Windows Phone [ ] other _____

2. How long have you been using a smartphone?

11 *years*

3. How many hours per week do you use a smartphone?

20 *hours*

7. Web Use

1. How many hours per week do you use the World Wide Web?

20 *hours*

2. Which kind of device do you use most often to surf the web?

[ ] desktop computer [ ] laptop [ ] tablet
[x] smartphone [ ] other _____

3. What kind of internet connection do you normally use?

[ ] xDSL [ ] cable modem [ ] fibre optic
[ ] 3G mobile internet [x] LTE mobile internet [ ] other _____

4. Which web browser do you normally use?

[ ] Chrome [ ] Firefox [x] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] other _____

5. Do you have experience as a web site administrator?

If yes: ____ years

8. Domain-Specific Questions

1. Have you ever designed an information hierarchy?
(e.g. product categories and sub-categories, hierarchical menu structure, website navigation hierarchy)

[X] no [ ] yes, _____

2. Have you ever participated as a test user in a study to test an information hierarchy?

[X] no [ ] yes, _____

3. Have you ever used an online tool to carry out a user study? If so, which one(s)?

[X] no [ ] yes, _____

9. Experience with Usability Tests

1. Have you participated in any kind of usability study before?

[ ] as a test user [ ] as part of the test team

If yes, what kind of study was it?

[ ] Thinking Aloud [ ] Formal Experiment [ ] other _____

Date: 18.11.19 Time: 15:05 Test User ID: 5

Background Questionnaire

1. General Information

Sex: [X] male [ ] female

Age: 23

Occupation: Student

2. Sight Impairment

1. Do you use a sight aid when working on the computer?

[X] none [ ] glasses [ ] contact lenses [ ] other _____

2. Do you have any form of colour blindness?

[X] no [ ] yes, _____

3. Education

1. Highest Educational Level Attained:

[ ] vocational training [X] secondary school [ ] university degree [ ] doctorate

2. If you are studying or have studied, please describe your main area of study:

SEW

4. Personal Computer Use

1. Which kind of personal computer do you use most?

[ ] None [ ] Microsoft Windows [X] Apple Macintosh [ ] Unix [ ] other _____

2. How long have you been using a personal computer?

10 years

3. How many hours per week do you use a personal computer?

40 hours

5. Tablet Use

1. Which kind of tablet do you use most?

[ ] None [ ] Android [x] Apple (iOS) [ ] other _____

2. How long have you been using a tablet?

5 *years*

3. How many hours per week do you use a tablet?

4 *hours*

6. Smartphone Use

1. Which kind of smartphone do you use most?

[ ] None [ ] Android [x] Apple (iOS) [ ] Windows Phone [ ] other _____

2. How long have you been using a smartphone?

8 *years*

3. How many hours per week do you use a smartphone?

15 *hours*

7. Web Use

1. How many hours per week do you use the World Wide Web?

40+4+15
59 *hours*

2. Which kind of device do you use most often to surf the web?

[ ] desktop computer [x] laptop [ ] tablet
[ ] smartphone [ ] other _____

3. What kind of internet connection do you normally use?

[ ] xDSL [ ] cable modem [ ] fibre optic
[ ] 3G mobile internet [x] LTE mobile internet [ ] other _____

4. Which web browser do you normally use?

[x] Chrome [ ] Firefox [ ] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] other _____

5. Do you have experience as a web site administrator?

If yes: 5 *years*

8. Domain-Specific Questions

1. Have you ever designed an information hierarchy?
(e.g. product categories and sub-categories, hierarchical menu structure, website navigation hierarchy)

[ ] no [x] yes, _____

2. Have you ever participated as a test user in a study to test an information hierarchy?

[x] no [ ] yes, _____

3. Have you ever used an online tool to carry out a user study? If so, which one(s)?

[x] no [ ] yes, _____

9. Experience with Usability Tests

1. Have you participated in any kind of usability study before?

[ ] as a test user [x] as part of the test team

If yes, what kind of study was it?

[x] Thinking Aloud [ ] Formal Experiment [ ] other _____

Datum: 18.11.19 Uhrzeit: 17:29 Test User ID: 6

Hintergrundbefragung

1. Allgemeine Informationen

Geschlecht: [x] *männlich* [ ] *weiblich*

Alter: 22

Beruf: Student

2. Sehvermögen

1. Verwenden Sie eine Sehhilfe bei der Arbeit am Computer?

[ ] Keine [x] Brille [x] Kontaktlinsen [ ] sonstige _____

2. Sind Sie farbenblind?

[x] Nein [ ] Ja, und zwar _____

3. Ausbildung

1. Höchste abgeschlossene Ausbildung:

[ ] Lehre [x] Matura [ ] Studium [ ] Doktorat

2. Wenn Sie studieren oder studiert haben, beschreiben Sie bitte Ihr Hauptstudiengebiet:

SEW

4. Verwendung eines Computers

1. Welche Art von Personal Computer verwenden Sie am häufigsten?

[ ] Keinen [ ] Microsoft Windows [x] Apple Macintosh [ ] Unix [ ] sonstige _____

2. Wie lange benutzen Sie bereits einen Personal Computer?

12 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie einen Personal Computer?

60 *Stunden*

5. Verwendung eines Tablets

1. Welche Art von Tablet verwenden Sie am häufigsten?

[ ] Keines [ ] Android [x] Apple (iOS) [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Tablet?

7 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie ein Tablet?

1 *Stunden*

6. Verwendung eines Smartphones

1. Welche Art von Smartphone verwenden Sie am häufigsten?

[ ] Keinen [ ] Android [x] Apple (iOS) [ ] Windows Phone [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Smartphone?

7 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie ein Smartphone?

6 *Stunden*

7. Verwendung des Web

1. Wie viele Stunden pro Woche benutzen Sie das World Wide Web?

40 *Stunden*

2. Welches Gerät verwenden Sie am häufigsten zum surfen?

[ ] Desktop PC [x] Laptop [ ] Tablet
[ ] Smartphone [ ] sonstige _____

3. Welche Art von Internet Zugang verwenden Sie normalerweise?

[ ] xDSL [ ] Kabel Modem [x] Glasfaser
[ ] 3G mobiles Internet [ ] LTE mobiles Internet [ ] sonstige _____

4. Welchen Web-Browser verwenden Sie normalerweise?

[x] Chrome [ ] Firefox [x] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] sonstige _____

5. Haben Sie Erfahrung als Web Site Administrator?

Wenn ja, wie lange? _____ *Jahre*

8. Domänenspezifische Fragen

1. Haben Sie schon einmal eine Informationshierarchie entworfen?
(z.B. Produktkategorien und Unterkategorien, hierarchische Menüstruktur, Website-Navigationshierarchie)

[ ] Nein [x] Ja, Web

2. Haben Sie schon einmal als Testperson an einer Studie zum Testen einer Informationshierarchie teilgenommen?

[x] Nein [ ] Ja, _____

3. Haben Sie schon einmal ein Online-Tool zur Durchführung einer Nutzerstudie benützt? Wenn ja, welche(s)?

[x] Nein [ ] Ja, _____

9. Erfahrung mit Usability Tests

1. Haben Sie schon an einer Usability Studie teilgenommen?

[ ] als Testperson [ ] als Mitglied des Testteams

Wenn ja, was war das für eine Studie?

[ ] Thinking Aloud [ ] Formal Experiment [ ] sonstige _____

Datum: 18.11.19 Uhrzeit: 19:31 Test User ID: 7

Hintergrundbefragung

1. Allgemeine Informationen

Geschlecht: [x] *männlich* [ ] *weiblich*

Alter: 42

Beruf: Sozialpädagoge

2. Sehvermögen

1. Verwenden Sie eine Sehhilfe bei der Arbeit am Computer?

[ ] Keine [x] Brille [ ] Kontaktlinsen [ ] sonstige _____

2. Sind Sie farbenblind?

[x] Nein [ ] Ja, und zwar _____

3. Ausbildung

1. Höchste abgeschlossene Ausbildung:

[ ] Lehre [x] Matura [ ] Studium [ ] Doktorat Kolleg

2. Wenn Sie studieren oder studiert haben, beschreiben Sie bitte Ihr Hauptstudiengebiet:

_____

4. Verwendung eines Computers

1. Welche Art von Personal Computer verwenden Sie am häufigsten?

[ ] Keinen [x] Microsoft Windows [ ] Apple Macintosh [ ] Unix [ ] sonstige _____

2. Wie lange benutzen Sie bereits einen Personal Computer?

15 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie einen Personal Computer?

2 *Stunden*

5. Verwendung eines Tablets

1. Welche Art von Tablet verwenden Sie am häufigsten?

[ ] Keines [ ] Android [ ] Apple (iOS) [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Tablet?

_____ *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie ein Tablet?

_____ *Stunden*

6. Verwendung eines Smartphones

1. Welche Art von Smartphone verwenden Sie am häufigsten?

[ ] Keinen [X] Android [ ] Apple (iOS) [ ] Windows Phone [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Smartphone?

6 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie ein Smartphone?

42 *Stunden*

7. Verwendung des Web

1. Wie viele Stunden pro Woche benutzen Sie das World Wide Web?

38 *Stunden*

2. Welches Gerät verwenden Sie am häufigsten zum surfen?

[ ] Desktop PC [ ] Laptop [ ] Tablet
[X] Smartphone [ ] sonstige _____

3. Welche Art von Internet Zugang verwenden Sie normalerweise?

[ ] xDSL [ ] Kabel Modem [ ] Glasfaser
[ ] 3G mobiles Internet [X] LTE mobiles Internet [ ] sonstige _____

4. Welchen Web-Browser verwenden Sie normalerweise?

[X] Chrome [ ] Firefox [ ] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] sonstige _____

5. Haben Sie Erfahrung als Web Site Administrator?

Wenn ja, wie lange? _____ *Jahre*

8. Domänenspezifische Fragen

1. Haben Sie schon einmal eine Informationshierarchie entworfen?
(z.B. Produktkategorien und Unterkategorien, hierarchische Menüstruktur, Website-Navigationshierarchie)

[X] Nein [ ] Ja, _____

2. Haben Sie schon einmal als Testperson an einer Studie zum Testen einer Informationshierarchie teilgenommen?

[X] Nein [ ] Ja, _____

3. Haben Sie schon einmal ein Online-Tool zur Durchführung einer Nutzerstudie benützt? Wenn ja, welche(s)?

[X] Nein [ ] Ja, _____

9. Erfahrung mit Usability Tests

1. Haben Sie schon an einer Usability Studie teilgenommen?

[X] als Testperson [ ] als Mitglied des Testteams

Wenn ja, was war das für eine Studie?

[X] Thinking Aloud [ ] Formal Experiment [ ] sonstige _____

Datum: 19.11.19 Uhrzeit: 09:06 Test User ID: 8

Hintergrundbefragung

1. Allgemeine Informationen

Geschlecht: [x] *männlich* [ ] *weiblich*

Alter: 22

Beruf: Student

2. Sehvermögen

1. Verwenden Sie eine Sehhilfe bei der Arbeit am Computer?

[ ] Keine [x] Brille [ ] Kontaktlinsen [ ] sonstige _____

2. Sind Sie farbenblind?

[x] Nein [ ] Ja, und zwar _____

3. Ausbildung

1. Höchste abgeschlossene Ausbildung:

[ ] Lehre [x] Matura [ ] Studium [ ] Doktorat

2. Wenn Sie studieren oder studiert haben, beschreiben Sie bitte Ihr Hauptstudiengebiet:

Informatik

4. Verwendung eines Computers

1. Welche Art von Personal Computer verwenden Sie am häufigsten?

[ ] Keinen [x] Microsoft Windows [ ] Apple Macintosh [ ] Unix [ ] sonstige _____

2. Wie lange benutzen Sie bereits einen Personal Computer?

9 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie einen Personal Computer?

50 *Stunden*

5. Verwendung eines Tablets

1. Welche Art von Tablet verwenden Sie am häufigsten?

[ ] Keines [ ] Android [ ] Apple (iOS) [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Tablet?

_____ *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie ein Tablet?

_____ *Stunden*

6. Verwendung eines Smartphones

1. Welche Art von Smartphone verwenden Sie am häufigsten?

[ ] Keinen [X] Android [ ] Apple (iOS) [ ] Windows Phone [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Smartphone?

7 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie ein Smartphone?

35 *Stunden*

7. Verwendung des Web

1. Wie viele Stunden pro Woche benutzen Sie das World Wide Web?

80 *Stunden*

2. Welches Gerät verwenden Sie am häufigsten zum surfen?

[ ] Desktop PC [X] Laptop [ ] Tablet
[ ] Smartphone [ ] sonstige _____

3. Welche Art von Internet Zugang verwenden Sie normalerweise?

[ ] xDSL [ ] Kabel Modem [ ] Glasfaser
[X] 3G mobiles Internet [ ] LTE mobiles Internet [ ] sonstige _____

4. Welchen Web-Browser verwenden Sie normalerweise?

[ ] Chrome [X] Firefox [ ] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] sonstige _____

5. Haben Sie Erfahrung als Web Site Administrator?

Wenn ja, wie lange? ____ *Jahre*

8. Domänenspezifische Fragen

1. Haben Sie schon einmal eine Informationshierarchie entworfen?
(z.B. Produktkategorien und Unterkategorien, hierarchische Menüstruktur, Website-Navigationshierarchie)

[X] Nein [ ] Ja, _____

2. Haben Sie schon einmal als Testperson an einer Studie zum Testen einer Informationshierarchie teilgenommen?

[X] Nein [ ] Ja, _____

3. Haben Sie schon einmal ein Online-Tool zur Durchführung einer Nutzerstudie benützt? Wenn ja, welche(s)?

[X] Nein [ ] Ja, _____

9. Erfahrung mit Usability Tests

1. Haben Sie schon an einer Usability Studie teilgenommen?

[ ] als Testperson [ ] als Mitglied des Testteams

Wenn ja, was war das für eine Studie?

[ ] Thinking Aloud [ ] Formal Experiment [ ] sonstige _____

Datum: 12.11.19 Uhrzeit: 10:57 Test User ID: 9

Hintergrundbefragung

1. Allgemeine Informationen

Geschlecht: [ ] *männlich* [x] *weiblich*

Alter: 58

Beruf: Bürokauffrau

2. Sehvermögen

1. Verwenden Sie eine Sehhilfe bei der Arbeit am Computer?

   [x] *Keine* [ ] *Brille* [ ] *Kontaktlinsen* [ ] *sonstige* _____

2. Sind Sie farbenblind?

   [x] *Nein* [ ] *Ja, und zwar* _____

3. Ausbildung

1. Höchste abgeschlossene Ausbildung:

   [x] *Lehre* [ ] *Matura* [ ] *Studium* [ ] *Doktorat*

2. Wenn Sie studieren oder studiert haben, beschreiben Sie bitte Ihr Hauptstudiengebiet:

   _____

4. Verwendung eines Computers

1. Welche Art von Personal Computer verwenden Sie am häufigsten?

   [ ] *Keinen* [x] *Microsoft Windows* [ ] *Apple Macintosh* [ ] *Unix* [ ] *sonstige* _____

2. Wie lange benutzen Sie bereits einen Personal Computer?

   20 *Jahre*

3. Wie viele Stunden pro Woche verwenden Sie einen Personal Computer?

   8 *Stunden*

5. Verwendung eines Tablets

1. Welche Art von Tablet verwenden Sie am häufigsten?

[ ] Keines [ ] Android [x] Apple (iOS) [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Tablet?

4 Jahre

3. Wie viele Stunden pro Woche verwenden Sie ein Tablet?

6 Stunden

6. Verwendung eines Smartphones

1. Welche Art von Smartphone verwenden Sie am häufigsten?

[ ] Keinen [x] Android [ ] Apple (iOS) [ ] Windows Phone [ ] sonstige _____

2. Wie lange benutzen Sie bereits ein Smartphone?

3 Jahre

3. Wie viele Stunden pro Woche verwenden Sie ein Smartphone?

7 Stunden

7. Verwendung des Web

1. Wie viele Stunden pro Woche benutzen Sie das World Wide Web?

9 Stunden

2. Welches Gerät verwenden Sie am häufigsten zum surfen?

[ ] Desktop PC [ ] Laptop [x] Tablet
[ ] Smartphone [ ] sonstige _____

3. Welche Art von Internet Zugang verwenden Sie normalerweise?

[ ] xDSL [ ] Kabel Modem [x] Glasfaser
[ ] 3G mobiles Internet [ ] LTE mobiles Internet [ ] sonstige _____

4. Welchen Web-Browser verwenden Sie normalerweise?

[x] Chrome [ ] Firefox [ ] Safari
[ ] Microsoft IE (Edge) [ ] Opera [ ] sonstige _____

5. Haben Sie Erfahrung als Web Site Administrator?

Wenn ja, wie lange? _____ *Jahre*

8. Domänenspezifische Fragen

1. Haben Sie schon einmal eine Informationshierarchie entworfen?
(z.B. Produktkategorien und Unterkategorien, hierarchische Menüstruktur, Website-Navigationshierarchie)

[X] Nein [ ] Ja, _____

2. Haben Sie schon einmal als Testperson an einer Studie zum Testen einer Informationshierarchie teilgenommen?

[X] Nein [ ] Ja, _____

3. Haben Sie schon einmal ein Online-Tool zur Durchführung einer Nutzerstudie benützt? Wenn ja, welche(s)?

[X] Nein [ ] Ja, _____

9. Erfahrung mit Usability Tests

1. Haben Sie schon an einer Usability Studie teilgenommen?

[ ] als Testperson [ ] als Mitglied des Testteams

Wenn ja, was war das für eine Studie?

[ ] Thinking Aloud [ ] Formal Experiment [ ] sonstige _____

Appendix B

Task Slips for Study Participants

B.1 Administrator Task Slips

The two task slips provided to administrators are shown in Figures B.1 and B.2.

B.2 Study Owner Task Slips

The task slips provided to study owners are shown in Figures B.3, B.4, B.5 and B.6. Task slip 5 shown in Figures B.7 was not given to the study owners. This task consisted of questions about the study results.

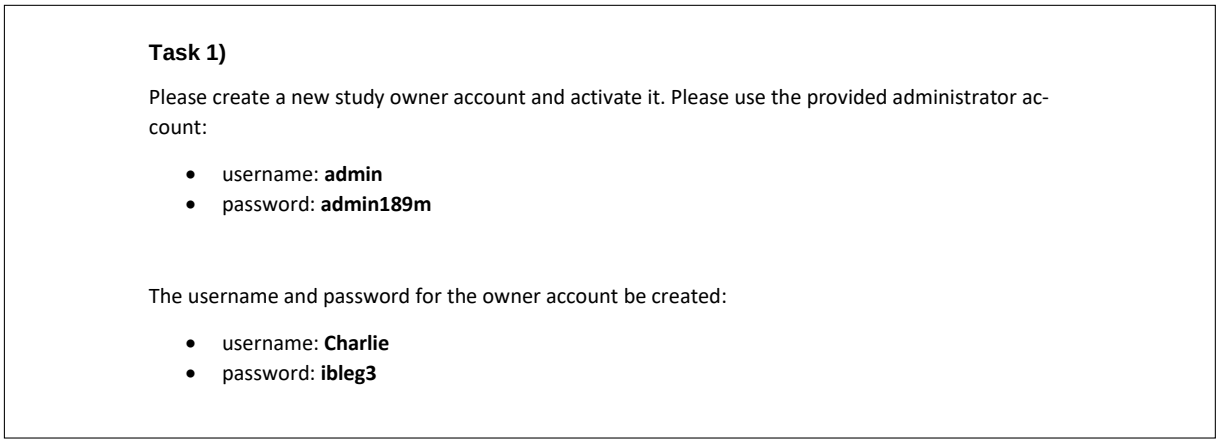

Figure B.1: Administrator Task 1.

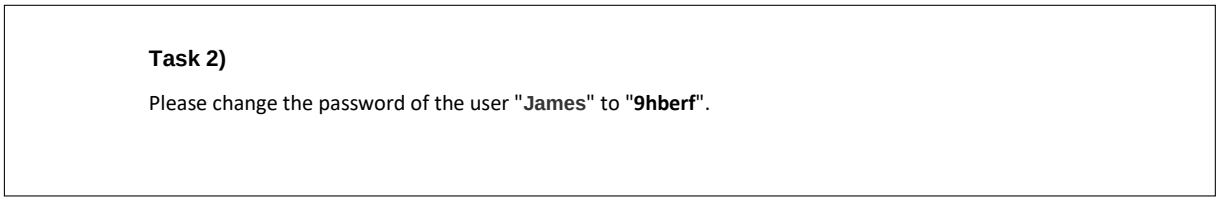

Figure B.2: Administrator Task 2.

Task 1)

Pease log in to the app as a study owner:

- user name: **Renee**
- password: **awesome**

Please create the following study:

- name: **TU-Graz-2019**
- password: **tu19**
- Upload the tree (information hierarchy) to be tested. The file containing the test tree can be found on the desktop: **tugraz-2019-at-en.csv**
- Define the following tasks:
    1. Where would you expect to find information about the history of Graz University of Technology?

       Home -> TU Graz -> University -> **History**

    2. Where would you expect to find information about the Faculty of Electrical and Information Engineering?

       Home -> Faculties and Institutes -> Overview: Faculties and Institutes -> **Faculty of Electrical and Information Engineering**

    3. Where would you look for information on master's programmes?

       Home -> Studying and Teaching-> Degree and Certificate Programmes-> **Master's Degree Programmes**

    4. Where would you look for job vacancies?

       Home -> TU Graz-> Working at TU Graz -> **Job Vacancies**

    5. Where would you look for information about Accessible Learning?

       Home -> Studying and Teaching-> Studying at TU Graz-> **Accessible Learning**

    6. Where would you look for research collaborations?

       Home -> Research-> Research at TU Graz-> **Research Cooperation Ventures**

    7. Where would you look for information about studying abroad?

       Home -> Studying and Teaching -> International: Studying and Teaching -> **Study Abroad**

    8. Where would you look for events at Tu Graz?

       Home -> TU Graz -> Services -> **TU Graz events**

Figure B.3: Study Owner Task 1.

Task 2)

Call up the preview of the study you have created and follow the instructions. Use the provided name.

Use the following name: **Riley**

Figure B.4: Study Owner Task 2.

Task 3)

Please change the question and the path of the study task 4 „Where would you look for job vacancies? To:

- "Where would you look for information to post job vacancies?"
- Home -> TU Graz -> University -> Services -> **Publish Job Vacancies**

Finally, please launch the study.

Figure B.5: Study Owner Task 3.

Task 4)

Please look at the results of the "**TU Graz**" study and analyse them for a few minutes.

Figure B.6: Study Owner Task 4.

Task 5)

Which task was answered correctly by most participants? [correct: Task 7]

Which task could only be answered correctly by a few participants? [correct: Task 6]

Not every participant has found the right answer for task 8, what was chosen by them as their answer? [correct: Destinations for Task 8]

Not every participant directly found the right answer for task 6, where did the participants search in the tree? [correct: Task analysis- > Path Tree for Task 6]

Figure B.7: Study Owner Task 5.

Task:

Please log in to the study using the following name:

- **Niki**

and then follow the instructions.

Figure B.8: Study Participant Task.

B.3 Study Participant Task Slip

The task slip provided to regular study participants is shown in Figure B.8.

Appendix C

Feedback Questionnaires

A summary of the two different feedback questionnaires was already presented in Section 6.7. In the following, the feedback questionnaires filled out by the individual test users are presented. The feedback questionnaire was taken from the course material of the Human-Computer Interaction course by Andrews [2019a]. The content of the feedback questionnaires has been partly adapted.

Datum: 18.11.19 Uhrzeit: 9:53 Test User ID: 1

Feedback Fragebogen (Study Owner)

Bewerten Sie bitte anhand folgender Aspekte ihre Zufriedenheit mit der App. Markieren Sie dazu die passende Nummer mit einem Kreis.

1.	Die Benutzeroberfläche der App ist übersichtlich und logisch.	Sehr klar	(3)	2	1	0	1	2	3	Sehr unklar
2.	Texte sind leicht zu lesen.	Sehr einfach	3	2	1	0	1	(2)	3	Sehr schwer
3.	Graphische Gestaltung der App, inkl. Farben und Grafiken.	Sehr gut	3	(2)	1	0	1	2	3	Sehr schlecht
4.	Es ist einfach, zum richtigen Teil der App zu gelangen.	Sehr einfach	3	(2)	1	0	1	2	3	Sehr schwer
5.	Konsistenz der App.	Sehr konsistent	(3)	2	1	0	1	2	3	Sehr inkonsistent
6.	Reaktionsgeschwindigkeit der App.	Sehr schnell	3	(2)	1	0	1	2	3	Sehr langsam
7.	Wie gut führt Sie die App durch den Prozess zur Durchführung einer Studie?	Sehr gut	(3)	2	1	0	1	2	3	Sehr schlecht
8.	Wie einfach ist es, eine Studie zu erstellen?	Sehr einfach	(3)	2	1	0	1	2	3	Sehr schwer
9.	Wie einfach ist es, einen Baum zu erstellen/hochzuladen?	Sehr einfach	(3)	2	1	0	1	2	3	Sehr schwer
10.	Wie einfach ist es, Aufgaben zu erstellen?	Sehr einfach	(3)	2	(1)	0	1	2	3	Sehr schwer
11.	Wie einfach ist es, durch den Baum zu navigieren, um die richtige Antwort auszuwählen?	Sehr einfach	3	2	(1)	0	1	2	3	Sehr schwer
12.	Wie gut werden die Ergebnisse einer Studie präsentiert?	Sehr gut	3	(2)	1	0	1	2	3	Sehr schlecht
13.	Wie nützlich ist der Pfadbaum?	Sehr nützlich	3	2	1	(0)	1	2	3	Überhaupt nicht
14.	Diese App kümmert sich um meine Zufriedenheit.	Sehr	3	2	1	(0)	1	2	3	Überhaupt nicht
15.	Halten Sie diese App für vertrauenswürdig?	Auf jeden Fall	3	2	(1)	0	1	2	3	Überhaupt nicht
16.	Wie relevant ist die Funktionalität dieser App für Sie?	Sehr relevant	3	2	1	0	(1)	2	3	Überhaupt nicht
17.	Würden Sie diese App in Ihrer eigenen Arbeit verwenden?	Auf jeden Fall	3	2	1	0	1	(2)	3	Niemals
18.	Gesamteindruck der App.	Sehr gut	3	2	(1)	0	1	2	3	Sehr schlecht

Date: 18.11.19 Time: 11:44 Test User ID: 2

Feedback Questionnaire (Participant)

Please rate your satisfaction with these aspects of the web app you have just finished working with, by circling the most appropriate number.

1.	The app's user interface is clear and logical.	Very clear	3	2	1	0	1	2	3 Very unclear
2.	Text is easy to read.	Very easy	3	2	1	0	1	2	3 Very hard
3.	Appearance of app, including colours and graphics.	Very good	3	2	1	0	1	2	3 Very poor
4.	Consistency of app.	Very consistent	3	2	1	0	1	2	3 Very inconsistent
5.	Response speed of app.	Very fast	3	2	1	0	1	2	3 Very slow
6.	How well did the app guide you through the test?	Very good	3	2	1	0	1	2	3 Very poor
7.	How easy is it to navigate through the tree?	Very easy	3	2	1	0	1	2	3 Very hard
8.	This app cares about my satisfaction as a user.	Very much.	3	2	1	0	1	2	3 Not at all.
9.	Do you consider this app to be trustworthy?	Definitely	3	2	1	0	1	2	3 Not all all
10.	How relevant is the functionality of this app to you?	Very relevant.	3	2	1	0	1	2	3 Not at all.
11.	Would you consider using this app in your own work?	Definitely	3	2	1	0	1	2	3 Never
12.	Overall impression of the app.	Very good	3	2	1	0	1	2	3 Very bad

Date: 18.11.19 Time: 13:09 Test User ID: 3

Feedback Questionnaire (Study Owner)

Please rate your satisfaction with these aspects of the web site you have just finished working with, by circling the most appropriate number.

1.	The app's user interface is clear and logical.	Very clear	[3]	2	1	0	1	2	3	Very unclear
2.	Text is easy to read.	Very easy	[3]	2	1	0	1	2	3	Very hard
3.	Appearance of app, including colours and graphics.	Very good	[3]	2	1	0	1	2	3	Very poor
4.	Getting to the right part of the app is easy.	Very easy	[3]	2	1	0	1	2	3	Very hard
5.	Consistency of app.	Very consistent	[3]	2	1	0	1	2	3	Very inconsistent
6.	Response speed of app.	Very fast	[3]	2	1	0	1	2	3	Very slow
7.	How well does the app guide you through the process of conducting a study?	Very well	3	[2]	1	0	1	2	3	Very poor
8.	How easy is it to create a study?	Very easy	[3]	2	1	0	1	2	3	Very hard
9.	How easy is it to create/upload a tree?	Very easy	[3]	2	1	0	1	2	3	Very hard
10.	How easy is it to create tasks?	Very easy	[3]	2	1	0	1	2	3	Very hard
11.	How easy is it to navigate through the tree to select the correct answer?	Very easy	3	[2]	1	0	1	2	3	Very hard
12.	How well are the results of a study presented?	Very well	[3]	2	1	0	1	2	3	Very poor
13.	How useful is the path tree?	Very useful	[3]	2	1	0	1	2	3	Not at all
14.	This app cares about my satisfaction as a customer.	Very much	3	2	[1]	0	1	2	3	Not at all
15.	Do you consider this app to be trustworthy?	Definitely	3	2	1	[0]	1	2	3	Not all all
16.	How relevant is the functionality of this app to you?	Very relevant	3	2	1	[0]	1	2	3	Not at all
17.	Would you consider using this app in your own work?	Definitely	3	[2]	1	0	1	2	3	Never
18.	Overall impression of the app.	Very good	[3]	2	1	0	1	2	3	Very bad

Date: 18.11.19 Time: 14:56 Test User ID: 4

Feedback Questionnaire (Study Owner)

Please rate your satisfaction with these aspects of the web site you have just finished working with, by circling the most appropriate number.

1.	The app's user interface is clear and logical.	Very clear	3	2	(1)	0	1	2	3	Very unclear
2.	Text is easy to read.	Very easy	(3)	2	1	0	1	2	3	Very hard
3.	Appearance of app, including colours and graphics.	Very good	(3)	2	1	0	1	2	3	Very poor
4.	Getting to the right part of the app is easy.	Very easy	3	(2)	1	0	1	2	3	Very hard
5.	Consistency of app.	Very consistent	3	(2)	1	0	1	2	3	Very inconsistent
6.	Response speed of app.	Very fast	(3)	2	1	0	1	2	3	Very slow
7.	How well does the app guide you through the process of conducting a study?	Very well	3	2	(1)	0	1	2	3	Very poor
8.	How easy is it to create a study?	Very easy	3	2	(1)	0	1	2	3	Very hard
9.	How easy is it to create/upload a tree?	Very easy	3	2	(1)	0	1	2	3	Very hard
10.	How easy is it to create tasks?	Very easy	3	2	1	(0)	1	2	3	Very hard
11.	How easy is it to navigate through the tree to select the correct answer?	Very easy	3	2	1	0	(1)	2	3	Very hard
12.	How well are the results of a study presented?	Very well	3	2	1	(0)	1	2	3	Very poor
13.	How useful is the path tree?	Very useful	3	2	1	(0)	1	2	3	Not at all
14.	This app cares about my satisfaction as a customer.	Very much	3	2	1	(0)	1	2	3	Not at all
15.	Do you consider this app to be trustworthy?	Definitely	3	2	(1)	0	1	2	3	Not all all
16.	How relevant is the functionality of this app to you?	Very relevant	3	2	1	0	1	(2)	3	Not at all
17.	Would you consider using this app in your own work?	Definitely	3	2	1	0	1	(2)	3	Never
18.	Overall impression of the app.	Very good	3	2	(1)	0	1	2	3	Very bad

Date: 18.11.19 Time: 15:57 Test User ID: 5

Feedback Questionnaire (Study Owner)

Please rate your satisfaction with these aspects of the web site you have just finished working with, by circling the most appropriate number.

1.	The app's user interface is clear and logical.	Very clear	3	(2)	1	0	1	2	3	Very unclear
2.	Text is easy to read.	Very easy	(3)	2	1	0	1	2	3	Very hard
3.	Appearance of app, including colours and graphics.	Very good	(3)	2	1	0	1	2	3	Very poor
4.	Getting to the right part of the app is easy.	Very easy	3	(2)	1	0	1	2	3	Very hard
5.	Consistency of app.	Very consistent	3	(2)	1	0	1	2	3	Very inconsistent
6.	Response speed of app.	Very fast	(3)	2	1	0	1	2	3	Very slow
7.	How well does the app guide you through the process of conducting a study?	Very well	3	2	(1)	0	1	2	3	Very poor
8.	How easy is it to create a study?	Very easy	3	(2)	1	0	1	2	3	Very hard
9.	How easy is it to create/upload a tree?	Very easy	(3)	2	1	0	1	2	3	Very hard
10.	How easy is it to create tasks?	Very easy	3	(2)	1	0	1	2	3	Very hard
11.	How easy is it to navigate through the tree to select the correct answer?	Very easy	3	(2)	1	0	1	2	3	Very hard
12.	How well are the results of a study presented?	Very well	3	2	(1)	0	1	2	3	Very poor
13.	How useful is the path tree?	Very useful	3	(2)	1	0	1	2	3	Not at all
14.	This app cares about my satisfaction as a customer.	Very much	3	2	(1)	0	1	2	3	Not at all
15.	Do you consider this app to be trustworthy?	Definitely	(3)	2	1	0	1	2	3	Not all all
16.	How relevant is the functionality of this app to you?	Very relevant	3	2	1	0	1	(2)	3	Not at all
17.	Would you consider using this app in your own work?	Definitely	3	2	1	0	1	(2)	3	Never
18.	Overall impression of the app.	Very good	3	2	(1)	0	1	2	3	Very bad

Datum: 18.11.19 Uhrzeit: 17:37 Test User ID: 6

Feedback Fragebogen (Participant)

Bewerten Sie bitte anhand folgender Aspekte ihre Zufriedenheit mit der App. Markieren Sie dazu die passende Nummer mit einem Kreis.

1.	Die Benutzeroberfläche der App ist übersichtlich und logisch.	Sehr klar	3	(2)	1	0	1	2	3	Sehr unklar
2.	Texte sind leicht zu lesen.	Sehr einfach	(3)	2	1	0	1	2	3	Sehr schwer
3.	Graphische Gestaltung der App, inkl. Farben und Grafiken.	Sehr gut	3	2	1	0	1	2	(3)	Sehr schlecht
4.	Konsistenz der App.	Sehr konsistent	3	(2)	1	0	1	2	3	Sehr inkonsistent
5.	Geschwindigkeit der App.	Sehr schnell	3	(2)	1	0	1	2	3	Sehr langsam
6.	Wie gut hat Sie die App durch den Test geführt?	Sehr gut	3	2	1	0	1	2	(3)	Sehr schlecht
7.	Wie einfach ist es, durch den Baum zu navigieren?	Sehr einfach	3	2	(1)	0	1	2	3	Sehr schwer
8.	Die App kümmert sich um meine Zufriedenheit.	Auf jeden Fall	3	2	(1)	0	1	2	3	Gar nicht
9.	Halten Sie diese App für vertrauenswürdig?	Auf jeden Fall	3	2	1	0	1	(2)	3	Gar nicht
10.	Wie relevant ist die Funktionalität dieser App für Sie?	Sehr relevant	3	(2)	1	0	1	2	3	Überhaupt nicht
11.	Würden Sie diese App bei Ihrer eigenen Arbeit verwenden?	Auf jeden Fall	3	2	1	0	1	(2)	3	Niemals
12.	Gesamteindruck der App.	Sehr gut	3	2	1	0	1	(2)	3	Sehr schlecht

Datum: 18.11.19 Uhrzeit: 20:30 Test User ID: 7

Feedback Fragebogen (Study Owner)

Bewerten Sie bitte anhand folgender Aspekte ihre Zufriedenheit mit der App. Markieren Sie dazu die passende Nummer mit einem Kreis.

1.	Die Benutzeroberfläche der App ist übersichtlich und logisch.	Sehr klar	3	2	(1)	0	1	2	3	Sehr unklar
2.	Texte sind leicht zu lesen.	Sehr einfach	3	2	(1)	0	1	2	3	Sehr schwer
3.	Graphische Gestaltung der App, inkl. Farben und Grafiken.	Sehr gut	3	2	1	0	1	(2)	3	Sehr schlecht
4.	Es ist einfach, zum richtigen Teil der App zu gelangen.	Sehr einfach	3	2	1	(0)	1	2	3	Sehr schwer
5.	Konsistenz der App.	Sehr konsistent	3	2	1	0	(1)	2	3	Sehr inkonsistent
6.	Reaktionsgeschwindigkeit der App.	Sehr schnell	(3)	2	1	0	1	2	3	Sehr langsam
7.	Wie gut führt Sie die App durch den Prozess zur Durchführung einer Studie?	Sehr gut	3	2	1	(0)	1	2	3	Sehr schlecht
8.	Wie einfach ist es, eine Studie zu erstellen?	Sehr einfach	3	(2)	1	0	1	2	3	Sehr schwer
9.	Wie einfach ist es, einen Baum zu erstellen/hochzuladen?	Sehr einfach	3	(2)	1	0	1	2	3	Sehr schwer
10.	Wie einfach ist es, Aufgaben zu erstellen?	Sehr einfach	3	(2)	1	0	1	2	3	Sehr schwer
11.	Wie einfach ist es, durch den Baum zu navigieren, um die richtige Antwort auszuwählen?	Sehr einfach	(3)	2	1	0	1	2	3	Sehr schwer
12.	Wie gut werden die Ergebnisse einer Studie präsentiert?	Sehr gut	3	2	1	(0)	1	2	3	Sehr schlecht
13.	Wie nützlich ist der Pfadbaum?	Sehr nützlich	3	2	1	(0)	1	2	3	Überhaupt nicht
14.	Diese App kümmert sich um meine Zufriedenheit.	Sehr	3	2	1	0	(1)	2	3	Überhaupt nicht
15.	Halten Sie diese App für vertrauenswürdig?	Auf jeden Fall	3	2	1	(0)	1	2	3	Überhaupt nicht
16.	Wie relevant ist die Funktionalität dieser App für Sie?	Sehr relevant	3	2	(1)	0	1	2	3	Überhaupt nicht
17.	Würden Sie diese App in Ihrer eigenen Arbeit verwenden?	Auf jeden Fall	3	2	1	0	1	(2)	3	Niemals
18.	Gesamteindruck der App.	Sehr gut	3	2	1	0	(1)	2	3	Sehr schlecht

Datum: 19.11.19 Uhrzeit: 09:31 Test User ID: 8

Feedback Fragebogen (Participant)

Bewerten Sie bitte anhand folgender Aspekte ihre Zufriedenheit mit der App. Markieren Sie dazu die passende Nummer mit einem Kreis.

1.	Die Benutzeroberfläche der App ist übersichtlich und logisch.	Sehr klar	3	(2)	1	0	1	2	3	Sehr unklar
2.	Texte sind leicht zu lesen.	Sehr einfach	3	2	1	0	(1)	2	3	Sehr schwer
3.	Graphische Gestaltung der App, inkl. Farben und Grafiken.	Sehr gut	3	2	1	0	1	2	(3)	Sehr schlecht
4.	Konsistenz der App.	Sehr konsistent	3	(2)	1	0	1	2	3	Sehr inkonsistent
5.	Geschwindigkeit der App.	Sehr schnell	(3)	2	1	0	1	2	3	Sehr langsam
6.	Wie gut hat Sie die App durch den Test geführt?	Sehr gut	3	2	(1)	0	1	2	3	Sehr schlecht
7.	Wie einfach ist es, durch den Baum zu navigieren?	Sehr einfach	3	2	(1)	0	1	2	3	Sehr schwer
8.	Die App kümmert sich um meine Zufriedenheit.	Auf jeden Fall	3	2	1	0	(1)	2	3	Gar nicht
9.	Halten Sie diese App für vertrauenswürdig?	Auf jeden Fall	3	2	1	0	1	(2)	3	Gar nicht
10.	Wie relevant ist die Funktionalität dieser App für Sie?	Sehr relevant	3	(2)	1	0	1	2	3	Überhaupt nicht
11.	Würden Sie diese App bei Ihrer eigenen Arbeit verwenden?	Auf jeden Fall	3	2	1	(0)	1	2	3	Niemals
12.	Gesamteindruck der App.	Sehr gut	3	2	1	(0)	1	2	3	Sehr schlecht

Datum: 12.11.19 Uhrzeit: 11:15 Test User ID: 9

Feedback Fragebogen (Participant)

Bewerten Sie bitte anhand folgender Aspekte ihre Zufriedenheit mit der App. Markieren Sie dazu die passende Nummer mit einem Kreis.

1.	Die Benutzeroberfläche der App ist übersichtlich und logisch.	Sehr klar	(3)	2	1	0	1	2	3	Sehr unklar
2.	Texte sind leicht zu lesen.	Sehr einfach	(3)	2	1	0	1	2	3	Sehr schwer
3.	Graphische Gestaltung der App, inkl. Farben und Grafiken.	Sehr gut	(3)	2	1	0	1	2	3	Sehr schlecht
4.	Konsistenz der App.	Sehr konsistent	(3)	2	1	0	1	2	3	Sehr inkonsistent
5.	Geschwindigkeit der App.	Sehr schnell	(3)	2	1	0	1	2	3	Sehr langsam
6.	Wie gut hat Sie die App durch den Test geführt?	Sehr gut	3	(2)	1	0	1	2	3	Sehr schlecht
7.	Wie einfach ist es, durch den Baum zu navigieren?	Sehr einfach	(3)	2	1	0	1	2	3	Sehr schwer
8.	Die App kümmert sich um meine Zufriedenheit.	Auf jeden Fall	(3)	2	1	0	1	2	3	Gar nicht
9.	Halten Sie diese App für vertrauenswürdig?	Auf jeden Fall	(3)	2	1	0	1	2	3	Gar nicht
10.	Wie relevant ist die Funktionalität dieser App für Sie?	Sehr relevant	3	(2)	1	0	1	2	3	Überhaupt nicht
11.	Würden Sie diese App bei Ihrer eigenen Arbeit verwenden?	Auf jeden Fall	(3)	2	1	0	1	2	3	Niemals
12.	Gesamteindruck der App.	Sehr gut	(3)	2	1	0	1	2	3	Sehr schlecht

Bibliography

Andrews, Keith [2008]. *Evaluation Comes in Many Guises*. CHI 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization (BELIV'08). Florence, Italy, 05 Apr 2008. 3 pages. https://ftp.isds.tugraz.at/pub/papers/andrews-beliv08.pdf (cited on pages 35–36).

Andrews, Keith [2019a]. *Materials for HCI Practicals 2019*. 01 Jan 2019. https://courses.isds.tugraz.at/hci/practicals/materials/ (cited on pages xi, 61, 93).

Andrews, Keith [2019b]. *Writing a Thesis: Guidelines for Writing a Master's Thesis in Computer Science*. Graz University of Technology, Austria. 24 Jan 2019. http://ftp.iicm.edu/pub/keith/thesis/ (cited on page xi).

Andrews, Keith [2020]. *Human-Computer Interaction Course Notes*. 04 Jun 2020. https://courses.isds.tugraz.at/hci/hci.pdf (cited on pages 35, 37).

Babich, Nick [2020]. *How Tree Testing Can Improve Your Website Usability*. Adobe. 29 Jan 2020. https://xd.adobe.com/ideas/process/information-architecture/tree-testing-improve-website-usability/ (cited on page 3).

Kohavi, Ron, Diane Tang, and Ya Xu [2020]. *Trustworthy Online Controlled Experiments: A Practical Guide to A/B Testing*. Cambridge University Press, 02 Apr 2020. 289 pages. ISBN 1108724264. https://experimentguide.com/ (cited on page 36).

Krug, Steve [2009]. *Rocket Surgery Made Easy*. New Riders, 18 Dec 2009. 168 pages. ISBN 0321657292 (cited on pages 35, 37–38).

MEAN [2019]. *MEAN.js*. 31 Oct 2019. https://meanjs.org/ (cited on page 27).

Mehic, Ajdin [2019]. *TreeTest: Online Tree Testing for Information Hierarchies*. Master's Thesis. Graz University of Technology, 14 Oct 2019. 96 pages. https://ftp.isds.tugraz.at/pub/theses/amehic-2019-msc.pdf (cited on pages IX, XI, 1, 7, 27).

Nielsen, Jakob [1992]. *Finding Usability Problems through Heuristic Evaluation*. Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '92) (Monterey, California, USA). ACM, 03 May 1992, pages 373–380. doi:10.1145/142750.142834 (cited on page 36).

Nielsen, Jakob [1994a]. *Enhancing the Explanatory Power of Usability Heuristics*. Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '94) (Boston, Massachusetts, USA). ACM, 24 Apr 1994, pages 152–158. doi:10.1145/191666.191729 (cited on page 36).

Nielsen, Jakob [1994b]. *Heuristic Evaluation*. In: *Usability Inspection Methods*. Edited by Jakob Nielsen and Robert L. Mack. Wiley, 1994. Chapter 2, pages 25–62. ISBN 0471018775 (cited on page 35).

Nielsen, Jakob [1994c]. *How to Conduct a Heuristic Evaluation*. Nielsen Norman Group. 01 Nov 1994. https://nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/ (cited on pages 36–37).

Nielsen, Jakob [1994d]. *Severity Ratings for Usability Problems*. Nielsen Norman Group. 01 Nov 1994. https://nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/ (cited on page 37).

Nielsen, Jakob [2012]. *Usability 101: Introduction to Usability*. Nielsen Norman Group. 03 Jan 2012. https://nngroup.com/articles/usability-101-introduction-to-usability/ (cited on page 35).

Nielsen, Jakob and Robert L. Mack, editors [1994]. *Usability Inspection Methods*. Wiley, 1994. ISBN 0471018775 (cited on page 35).

Nielsen, Jakob and Rolf Molich [1990a]. *Heuristic Evaluation of User Interfaces*. Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI '90) (Seattle, Washington, USA). ACM, 01 Apr 1990, pages 249–256. doi:10.1145/97243.97281 (cited on pages 36–37).

Nielsen, Jakob and Rolf Molich [1990b]. *Improving a Human-Computer Dialogue*. Communications of the ACM 33.3 (Mar 1990), pages 338–348. doi:10.1145/77481.77486 (cited on pages 35–36).

O'Brien, Dave [2020]. *Tree Testing for Websites*. 24 Jan 2020. https://treetesting.atlassian.net/ (cited on pages 4–5).

OW [2020a]. *Optimal Workshop*. 23 Aug 2020. https://optimalworkshop.com/ (cited on pages 7, 11).

OW [2020b]. *Treekjack*. Optimal Workshop. 23 Aug 2020. https://treejack.com/ (cited on page 7).

Ratcliff, Christopher [2019]. *What is Tree Testing and How Can It Improve Your Site's Usability?* userzoom. 31 Jul 2019. https://userzoom.com/blog/what-is-tree-testing-and-why-is-it-important-for-your-sites-ux/ (cited on page 4).

Rubin, Jeffrey and Dana Chisnell [2008]. *Handbook of Usability Testing*. 2nd Edition. Wiley, 25 Apr 2008. 384 pages. ISBN 0470185481 (cited on page 35).

Salazar, Kim [2016]. *Diary Studies: Understanding Long-Term User Behavior and Experiences*. Nielsen Norman Group. 06 May 2016. https://nngroup.com/articles/diary-studies/ (cited on page 35).

Sauro, Jeff [2018]. *Benchmarking the User Experience*. MeasuringUPress, 25 Jun 2018. 334 pages. ISBN 0692149090 (cited on pages 35–36).

Schroeder, Alana [2018]. *Tree Testing: Making Navigation Menus More Intuitive and Easier to Use*. UX Planet. 08 Feb 2018. https://uxplanet.org/tree-testing-making-navigation-menus-more-intuitive-easier-to-use-fbdc5be2e6bf (cited on page 3).

TUG [2020]. *Graz University of Technology*. Technische Universität Graz. 26 Jun 2020. https://tugraz.at/ (cited on pages 3–4, 45).

UXtweak [2020a]. *TreeTest Tool*. UXtweak. 23 Aug 2020. https://uxtweak.com/treetest-tool (cited on pages 7, 15).

UXtweak [2020b]. *UXtweak*. UXtweak. 23 Aug 2020. https://uxtweak.com/ (cited on pages 15, 19).

Walker, Rob [2003]. *The Guts of a New Machine*. The New York Times Magazine (30 Nov 2003). https:/nytimes.com/2003/11/30/magazine/the-guts-of-a-new-machine.html (cited on page 35).

Wharton, Cathleen, John Rieman, Clayton Lewis, and Peter Polson [1994]. *The Cognitive Walkthrough Method: A Practitioner's Guide*. In: *Usability Inspection Methods*. Edited by Jakob Nielsen and Robert L. Mack. Wiley, 1994. Chapter 5, pages 105–140. ISBN 0471018775. http://colorado.edu/ics/sites/default/files/attached-files/93-07.pdf (cited on page 35).

Whitenton, Kathryn [2017]. *Tree Testing: Fast, Iterative Evaluation of Menu Labels and Categories*. Nielsen Norman Group. 07 May 2017. https://nngroup.com/articles/tree-testing/ (cited on page 4).

Yuan, Sam [2019]. *The Things No One Tells You About Tree Testing*. UX Collective. 06 Jun 2019. https://uxdesign.cc/what-no-one-tells-you-about-tree-testing-ba00e7931fce (cited on page 3).