A usability study of a website prototype designed to be responsive to screen size

Sophie Rundgren
Interactive Media Design
Södertörn University
Alfred Nobels allé 7
141 89 Huddinge
sophierundgren@gmail.com

ABSTRACT
This paper describes a usability study of a responsive website prototype. A heuristic evaluation was conducted with three evaluators with a background in HCI, and user tests with four potential users were conducted as a control method to test the relevance of the heuristic evaluation result. The study showed that the website prototype failed the visibility of system status, user control and freedom and error prevention heuristics in the heuristic evaluation. However, the users responses contradict that result by showing that the users had no problems with those areas in the user tests. In fact, the users detected no usability problems at all. This could mean that those heuristics are not relevant for a responsive website in the future, and that designing for responsiveness leads to higher usability. This study shows some indications but more testing needs to be done, preferably on several fully developed responsive websites, to be able to make any real conclusions about the usability of a responsive website and the relevance of the heuristics used in this study.

Keywords
Responsive web design, Usability, Flat design, Interaction design, human-computer interaction, heuristic evaluation, user test

INTRODUCTION
Usability (successful interaction between humans and computer systems) has been extensively studied in human-computer interaction (HCI) and it is emphasized as the most important factor when designing a variety of systems, including websites [2]. However, up to this day it doesn’t seem to exist any research about the usability of a website that has been designed to be responsive to screen size.

In the field of web design, it is getting impossible to keep track of the endless new resolutions and devices that comes along [1]. Creating a version of a website for each of them is even more impossible for most websites. Responsive design is an approach in web design that has emerged in recent years as a way of dealing with all the different devices and screen sizes. It suggests that design and development “should respond to the user’s behavior and environment based on screen size, platform and orientation” [1, p.234] and is achieved with “a mix of flexible grids and layouts, images and an intelligent use of CSS media queries” [1, p.234].

The flexible, or fluid, layout is essential in responsive web design. In practice it means that the content adapts automatically to whatever screen size and device the website is viewed on, for example by changing in size, looks and even functionality. It is the reason why developers doesn’t have to develop a website version for each specific device and screen size – the code does it for them. [1]

However, responsive design is more than just a few lines of code that adapts the content automatically – it is a whole new way of thinking about design [1]. To create a responsive design is not only a job for the developers, but creates challenges for the designers as well; it is not only one layout that has to be created anymore, but instead several at the same time, and the designer need to have a clear understanding about how to arrange everything on each device. Web design nowadays is not one fixed design, but instead a design that changes and the designers need to design those changes as well. [9]

No research could be found about this, but some designers thinks that the reappearance of the design style ‘flat design’ is the result of the need for websites to work on several devices and screen sizes [8]. Flat design is the opposite of skeuomorphism (design that imitates reality); instead it is about minimalism and design that focuses more on the content than effect and decoration (which is why it is so suitable for multi-platforms). Most responsive websites adopts this design style because of this.

Prior to this study, a design project was carried out where a prototype of a responsive website was developed. The website prototype was developed as a small attempt to remedy the problems of overconsumption of clothes. It took the form of a service called SwapShop, an online community where users can upload clothing items, shoes and/or accessories that they don’t want anymore, which other users can send swap requests upon. The website prototype was designed with the goal of being responsive to
screen size in mind when designing. The features and interaction design were therefore designed to better suit a fluid layout, so that consistency between the variety of devices would be achieved. Also, the choice was made early in the project to use flat design, just because its suitability for responsive websites, and even though the prototype lacks the visual design, the features and elements are very affected by this choice – i.e. too keep the interaction design simple and focus on the content.

Though designed with the goal of being responsive to screen size, the prototype was developed in the size of a common computer screen. This prototype, in the form of interactive wireframes, later served as the test object in this study.

This study focuses on the following research questions; a) how does designing for responsiveness to screen size, affect usability (when viewed in the size of a common computer screen)? b) How relevant are the heuristics developed for web 2.0 websites for a usability evaluation of a website designed to be responsive to screen size?

Nothing suggests that there would be some sort of usability problem with a website designed to be responsive, but since there has been such a dramatic change in the web design area of HCI, it is highly relevant to test the affect it may have had on the user experience as well as the methods that are used to test the user experience. And the most important aspect in the user experience area is the usability [2], which is why this study was conducted.

DESCRIPTION OF SYSTEM
The design of the system is very simple and minimalistic. Nothing is there simply for the purpose of decoration or effect, instead the content and simple functionality is in focus. The content follows a grid-based layout to allow for a better and fluid rearrangement of the content on smaller screens. Additional pages were consciously avoided to keep the number of pages to a minimum. Instead the system is designed to respond to interaction while still on the same page. Information accessibility is lower when the screen space is smaller, and keeping the number of pages down to a minimum helps with that. For example, on the wardrobe page the ‘Upload New’ box is designed to be expanded on the same page. On the swaps page, both swap requests and swap messages are located on the same page. Functions such as ‘mark as favorite’ and ‘send swap request’ can be done both from the shop page and from the item page, to allow for a more accessible interaction. The search function on the shop page is designed so that the interaction never results in leaving that page, instead the search result is designed so it would update on the same page in real time.

As an example of a similar service, but not designed for responsiveness, see figure 2.
Figure 2. An example of a similar service to SwapShop, but not designed for responsiveness. Typical skeuomorphic style.

BACKGROUND
In 1990, Jakob Nielsen and Rolf Molich presented a set of heuristics (principles) of good interaction design, to help interaction designers point out usability problems in their designs [4,5]. Nielsen later revised the set of heuristics in 1994 based on a factor analysis on 249 usability problems, and it has since then served as a traditional set of guidelines when designing and evaluating interactive systems for HCI professionals [6].

<table>
<thead>
<tr>
<th>Visibility of system status</th>
<th>Match between system and real world</th>
</tr>
</thead>
<tbody>
<tr>
<td>User control and freedom</td>
<td>Error prevention</td>
</tr>
<tr>
<td>Consistency and standards</td>
<td>Flexibility and efficiency of use</td>
</tr>
<tr>
<td>Error prevention</td>
<td>Help users recognize, diagnose, and recover from errors</td>
</tr>
<tr>
<td>Recognition rather than recall</td>
<td>Interactive interface (replacing aesthetic and minimalist design)</td>
</tr>
<tr>
<td>Flexibility and efficiency of use</td>
<td>Data control and management</td>
</tr>
<tr>
<td>Aesthetic and minimalist design</td>
<td>Technologies</td>
</tr>
<tr>
<td>Help and documentation</td>
<td>Support for Web 2.0 paradigm</td>
</tr>
</tbody>
</table>

Table 1. Revised set of heuristics by Nielsen.

It is recommended to have between three and five evaluators when conducting a heuristic evaluation, since it is highly unlikely that one evaluator can detect all usability problems [5]. In fact, heuristic evaluation commonly produces hard-to-explain variations in the answers from different evaluators. Even if the evaluators have similar educational and professional backgrounds evaluation results can be very different – most likely due to differences in cognitive styles [3].

In 2009, Thompson and Kemp raised the question however the heuristics still was relevant for Web 2.0 sites [11]. The results of the study showed that some of the guidelines were not relevant anymore, and some needed to be changed or added, which led to proposing an updated version of the heuristics in their paper. Thompson’s and Kemp’s paper is deeply relevant in this study since the website prototype used in this study is a Web 2.0 site. The difference is that this website prototype has the addition of being designed for responsiveness, and therefore it is relevant to see if the updated set of heuristics by Thompson and Kemp still is relevant for a responsive Web 2.0 site. The first six heuristics are from the original set, developed by Nielsen and Molich, and the last four are new ones developed by Thompson and Kemp based on their study.

The omitted heuristics from the original set was omitted because they failed in the heuristic evaluation while having no equivalent response from the users in Thompson and Kemp’s study. For example, YouTube and Flickr websites failed the Recognition rather than recall heuristic, but the users did not find it difficult to remember their position on the website (which would be the equivalent to the heuristic in this regard). [11]

After comparing the usability inspection methods user tests and heuristic evaluation in a study, Tan, Liu and Bishu [10] concluded that they recommend using both methods in a development process. They mean that they are complementary but work best when used at different stages in the development process. They believe that heuristic evaluation works better at an early stage because evaluators could project potential usability problems even though the user interface is not fully developed, while user testing
requires a “well-developed test bed” [8, p.626] for it to be most useful. In their study it was found that the two methods detected very different and specific types of usability problems.

The mention of Nielsen was to give a background about how heuristic evaluation arose, and the description of Thompson and Kemp’s contribution in 2009 showed that the method of heuristic evaluation has been updated in recent time, and therefore suggest that using such a method would be appropriate to test the usability of a website. This is the why that method was chosen to test the usability of the website prototype in this study and it explains how the second research question became what it is.

METHODOLOGY

In addition to the conclusive remarks in the previous section, heuristic evaluation is a good usability inspection method of websites because it is cheap, it doesn’t require a lot of planning and it has been proven effective early in the design process [5]. The website prototype tested in this study is an early prototype, which again proves the method’s suitability. But to be able to test the relevance of the heuristic evaluation method and answer the second research question (how relevant are the heuristics developed for web 2.0 websites for a usability evaluation of a website designed to be responsive to screen size?) a second usability inspection method needed to be implemented as well, as a control method. In Thompson’s and Kemp’s study, they chose to use user testing as the control method [11] and since this study is similar to that (and since heuristic evaluation and user testing are two of the most popular usability inspection methods [10]) the decision to use user testing as the control method was made in this study as well. In addition, it is the users that would eventually use the website, so potential users’ opinions are the ones that matter the most and would therefore make sense as a control method. However, research has shown that these two methods are complementary in a usability study, pointing out different usability problems and having different strengths [10], which would benefit the first research question (how does designing for responsiveness to screen size, affect usability (when viewed in the size of a common computer screen)?, but less so for the second research question.

In both test situations, the participants were asked to perform the following tasks on the prototype; a) find information on how to use the service, b) upload a new item in your wardrobe, c) find a women’s jacket in the shop, mark it as a favorite and send a Swap Request, d) follow up on a received Swap Request by selecting something from the sending user’s wardrobe, e) answer an unread Swap message. These tasks were designed to cover the most prominent functions of the service. As mentioned, the prototype is in the form of wireframes that are made interactive in InDesign, and the file format is interactive PDF, so the participants viewed the prototype in an application that can view PDF files. The interactivity level of the prototype made it possible for the participants to actually click on different parts of the prototype to reach information (as you would do on an actual website), instead of scrolling (as you would do on a normal PDF). Not all links were clickable, just the ones needed to be able to complete the tasks.

Heuristic evaluation

For this study, the decision was made to exclude the interactive interface and technologies heuristics that Thompson and Kemp [11] presented in their article, because they would not make sense when testing a prototype that doesn’t have everything that a finished website has. Interactive interface has to do with the interactive aspects of Web 2.0 sites, for example the use of JavaScript and Ajax, as well as the overall visual design, and technologies covers the media and interactive quality, such as looking at the quality of images, video and alternative text. Since the prototype tested in this study lacks both of those aspects due to being more low-fi, it was best to exclude those heuristics altogether. Also, the decision to incorporate the consistency and standard heuristic by Nielsen was made, even though it was removed in the set presented by Thompson and Kemp. Why this decisions was made is because consistency was one of the main focuses when designing the website prototype, therefore it was highly relevant to include this heuristic.

Three evaluators with a background in HCI completed the heuristic evaluation. Each evaluator received a short brief about the website prototype and the set of tasks to complete on it. After they completed the tasks they were asked to complete a survey where they measured the usability by either agreeing or disagreeing (or neither) with statements (items) related to the heuristics on a 5-point Likert-scale. The scale went as follow: 1: strongly disagree, 2: disagree to some extent, 3: neither agree nor disagree, 4: agree to some extent and 5: strongly agree. See the full list of heuristics and their items in appendix B.

<table>
<thead>
<tr>
<th>Visibility of system status</th>
</tr>
</thead>
<tbody>
<tr>
<td>“The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.” [7]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Match between the system and the real world</th>
</tr>
</thead>
<tbody>
<tr>
<td>“The system should speak the users’ language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.” [7]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User control and freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Users often choose system functions by mistake and will need a clearly marked &quot;emergency exit&quot; to leave the</td>
</tr>
</tbody>
</table>
unwanted state without having to go through an extended dialogue. Support undo and redo.” [7]

### Consistency and standards

“Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.” [7]

### Error prevention

“Even better than good error messages is a careful design, which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.” [7]

### Flexibility and efficiency of use

“Accelerators -- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.” [7]

### Help and documentation

“Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user’s task, list concrete steps to be carried out, and not be too large.” [7]

### Data control and management

“This focuses on the idea that Web 2.0 sites are significantly motivated to achieve user driven and generated content. The user testing indicated that users were critical of the quality of much of the user content. Checks should be made to ensure that the information presented is relevant and accurate. Is it also clear how this information can be used?” [11]

### Support for Web 2.0 paradigm

Covers “aspects that are important for these sites, including whether users are able to tag content, join online communities and edit or change the content provided.” [11]

### Table 3. Set of heuristics used in this study

<table>
<thead>
<tr>
<th>User tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Four people relatively close to the author and part of the target audience of SwapShop were recruited as potential users. They received the same brief and set of tasks that the evaluators received, but instead of doing the heuristic evaluation survey after completing the tasks, they were asked to complete a survey containing items chosen from the Software Usability Measurement Inventory (<a href="http://sumi.ucc.ie/">http://sumi.ucc.ie/</a>). This is a way to capture the users thoughts in a more controlled manner, so comparisons to the heuristic evaluation results are easier to do. See the complete list of items in appendix C. The users answered the items in the same way that the evaluators answered the heuristic items, by either agreeing or disagreeing (or neither) on a 5-point Likert scale with the same grade of possible answers as stated in the previous section.</td>
</tr>
</tbody>
</table>

### RESULT

The website prototype was found to pass a heuristic if the evaluator answered 4 or 5 (‘agree to some extent’, or ‘strongly agree’) to at least half of the items related to a heuristic, otherwise the website prototype failed the heuristic. Since there were three evaluators, the prototype could both pass and fail a heuristic depending on the answers from each evaluator. Therefore, to get a fair final result of whether or not the prototype passed or failed the heuristics, the rule of majority complied – i.e. if it passed by one evaluator and failed by two, the overall result would be that it failed.

### Overall Heuristic evaluation results

The results showed proof of the hard-to-explain variations that heuristic evaluation commonly produces [3]. In the heuristic visibility of system status one evaluator’s score was 12/15 (pass) (12 is the sum of the answers put together and 15 is the highest possible score for that heuristic), the second evaluator’s score was 7/15 (fail) and the third evaluator’s score was 6/15 (fail). The second and third evaluator answered below 4 to at least half of the items, which resulted in the final conclusion that the prototype failed the visibility of system status heuristic. The second heuristic that the prototype failed was user control and freedom, where the scores were 24/25 (pass), 17/25 (fail) and 16/25 (fail). The third and last heuristic that the prototype failed was error prevention; the first evaluator answered 4/5 (pass) the second 2/5 (fail) and the third 3/5 (fail).

As shown, none of the heuristics that the prototype failed was so by all three evaluators. But for the most part of those that it passed, it passed by all three evaluators. The heuristics that the prototype passed by all three evaluators were consistency and standard, help and documentation, data control and management and support for the web 2.0 paradigm. The ones that it passed by two and failed by one evaluator (which equals to that it passed) were match between system and real world and flexibility and efficiency of use.

### Overall user test results

Overall, the users did not report any real usability problems. The only thing noticeable was that it was one user who sometimes wondered if s/he used the right function (s/he answered with a 4, ‘agree to some extent’). However, when reading that user’s response to similar items, as for example I can understand and act on the information provided by this software, the user had no problem. And when reading
the user’s comment about what can be improved, it seems that it was more about the form of the prototype (wireframes, with no visual design and not fully working) that made him/her a bit confused and questioned his/her own actions sometimes. Another user also brought up the form of the prototype in the free text comments, saying that though s/he thought the prototype was clear and understandable, it would probably be clearer when the visual design is in place. As for the rest of the users and items, the answers were somewhat similar.

Comparisons
The heuristic visibility of system status had no real equivalent among the user test items, but the ones that have the closest relation to it would be *it is easy to forget how to do things with this software and this software occasionally behaves in a way that can’t be understood* – which all users disagreed with. And the users did not react to this being a problem with the prototype in their free text comments either. The user test survey item that is related to the heuristic user control and freedom is *I feel in command when I use this website*, which all users did feel. The third heuristic that the prototype failed was error prevention and there are several user test survey items that are closely connected to that, such as the instructions and prompts are helpful, the way that system information is presented is clear and understandable and tasks can be performed in a straightforward manner using this software – all scored high by each user.

The user test item most related to the heuristic consistency and standard is *I think this software is inconsistent*, which all users answered with 1 – matching the unanimous result by the evaluators who all passed the consistency and standard heuristic. The heuristic help and documentation is equivalent to the user test items the instructions and prompts are helpful, the way that system information is presented is clear and understandable, I can understand and act on the information provided by this software (all users agreed to all of the items above) and either the amount or quality of the help information varies across the system (all users disagreed). As demonstrated, the users answers to all of these items are as unanimous as the evaluators’ agreement on the help and documentation heuristic.

There are no real equivalent user test items to the heuristics data control and management and support for web 2.0 paradigm, which makes it difficult to compare them to the user test.

DISCUSSION
Although Tan, Liu and Bishu [10] concluded that user tests and heuristic evaluation were complementary as usability inspection methods, pointing out different usability problems, this study shows signs of the contrary. The heuristics that was passed by all evaluators had equivalent aspects in the user tests that were just as unanimously approved by the users, showing that the evaluators and users thought alike in those regards. Knowing this, it is interesting to see that the heuristics that the prototype failed had no equivalent, but instead contradictory, result in the user tests.

This could, however, be the result of using the method of user tests on a prototype not developed far enough for the method to reach its full potential, as Tan, Liu and Bishu [10] described in their paper. Two out of four users brought up the form of the prototype in the free text comments, pointing out that the website surely would be more clear and understandable when the visual design and full functionality is in place. But then one could ask the question why the users still didn’t have any problems with the tasks and usability areas, if anything they would rate it worse because of the prototype being too under-developed for the method to reach its full potential – not better.

And since the evaluators and users thought alike in some regards, one could ask the question why they would not in other regards. It could mean that the heuristics that the prototype failed simply are not relevant anymore, for a responsive Web 2.0 website. It is worth mentioning that the heuristics that the prototype failed was all heuristics from the original set developed by Nielsen and all the new heuristics from the updated set by Thompson and Kemp passed. The heuristics that the prototype failed might not apply for a site that is designed with the goal of being responsive to screen size, which is, as mentioned earlier, often in a minimalistic style that keeps it simple and focuses on the content. The failed heuristics might simply be relevant for more complex (both interaction and visual design) sites, but are not needed when the website is designed like this prototype. The users clearly had no problem with the areas that the heuristic evaluation failed.

That the prototype passed all the heuristics from the updated set developed by Thompson and Kemp might have to do with that they are newer and adjusted just for the aspects of a Web 2.0 site, which the ones that failed are not.

CONCLUSION
The prototype failed the visibility of system status, user control and freedom and error prevention heuristics in the heuristic evaluation. However, the users responses contradict that result by showing that the users had no problems with those areas in the user tests. In fact, the users detected no usability problems at all. This could mean that those heuristics are not relevant for a responsive Web 2.0 site and needs to be removed when conducting a heuristic evaluation of a responsive website in the future.

On the other hand, since user tests reaches its full potential on a more well-developed test bed than this prototype according to [10] it could also mean that the result from the user tests are misleading and would not be the same if the prototype was more well-developed. Though, it is highly unlikely that the users would rate the usability so high like
they did if the prototype was too under-developed. One could argue that this means that designing for responsiveness simply leads to higher usability, since the users rated the usability so high even though, according to [10], the prototype isn’t developed far enough for user testing.

Regarding the research questions, it is difficult to answer either one of them because of the ambiguity of the study and the results.

This study shows some indications but more testing needs to be done, preferably on several fully developed responsive websites, to be able to make any real conclusions about the usability of a responsive website and the relevance of the heuristics used in this study. The ideal situation would be to conduct a heuristic evaluation in the same part of the development process as the prototype is in this study, then continue to develop it into a fully working website (without changing anything based on the result of the heuristic evaluation) – and then conduct user tests. This way, the uncertainties prominent in this study would be non-existent, and it would be a way to truly test the relevance of the heuristics.

LIMITATIONS
As mentioned, this study uses a prototype as a test object, and therefore it is difficult to make any real conclusions and generalizations about the usability of a responsive website. Also, the sample sizes in this study are too low of an amount to make generalizations based on them.

REFERENCES
8. Rutherford, C. Flat Design: just a design trend or that little bit more? Cargo Creative, 2013. http://www.cargocreative.co.uk/2013/10/flat-design-just-a-design-trend-or-that-little-bit-more/.
11. Thompson, A. and Kemp, E. Web 2.0: extending the framework for heuristic evaluation. ... of the 10th International Conference NZ ..., (2009), 29–36.
The first wireframe represents the start page and the second represents the ‘How it works’ page, which is the last main menu item and can be reached from every page through there.

These wireframes represent the wardrobe of the user currently logged in. The first is the overview of the wardrobe, and the second is when you’ve clicked ‘Upload new’.
The left wireframe is the shop page, which you can reach on every page through the ‘Shop’ menu item. The right wireframe is the item page, which you reach if you click on the image of an item in the shop page.

These wireframes represent what happens when you’ve clicked the swap request symbol. The dialog boxes look the same if the symbol is clicked on an item in the shop page.
The left wireframe is the swaps page, which you can reach through the ‘Swaps’ service menu item. A little number is shown next to the ‘swaps’ service menu item to show the amount of new requests and new swap messages. The right wireframe shows the swap conversations with an unread message expanded.
APPENDIX B

Visibility of system status
The website keeps me informed about where I am
The website keeps me informed about what is going on
The website’s features change as I carry out tasks

Match between system and the real world
The website contains familiar terms and natural language
The website contains metaphors from real life
The website follows real life conventions
The website’s screen representations matches non-computer

User control and freedom
Undo and redo is supported on the website
The website has obvious ways to undo actions
The website has clearly marked exists
The website holds the possibility to cancel tasks
The website allows me to initiate/control actions

Consistency and standards
The website is consistent in expression between pages
The website is consistent in layout/looks between pages
The website show similar information at same place on each screen

Error prevention
The website is designed to prevent errors

Flexibility and efficiency of use
The website offers shortcuts to important information
The website is efficient to use
The website gives me the possibility to tailor the content
The interactions with the website feels natural

Help and documentation
The website offers sufficient documentation about how to use the service
The documentation is easy to reach
The documentation is focused on the user’s tasks
The documentation consists of concrete steps to be carried out

The documentation is not too large in size

Data control and management
The website holds the possibility for me to upload my own content
The website enables user-generated content

Support for the Web 2.0 paradigm
The website allows me to tag content
The website allows me join an online community
The website allows me to edit or change content

APPENDIX C

1. The instructions and prompts are helpful.
2. I sometimes don't know what to do next with this software.
3. It takes too long to learn the software functions.
4. I sometimes wonder if I am using the right function.
5. Working with this software is satisfying.
6. The way that system information is presented is clear and understandable.
7. There is never enough information on the screen when it’s needed.
8. I feel in command of this software when I am using it.
9. I think this software is inconsistent.
10. I can understand and act on the information provided by this software.
11. Tasks can be performed in a straightforward manner using this software.
12. Using this software is frustrating.
13. The software has helped me overcome any problems I have had in using it.
14. It is obvious that user needs have been fully taken into consideration.
15. The organization of the menus seems quite logical.
16. There are too many steps required to get something to work.
17. I will never learn to use all that is offered in this software.
18. The software hasn't always done what I was expecting.
19. Either the amount or quality of the help information varies across the system.
20. It is relatively easy to move from one part of a task to another.

21. It is easy to forget how to do things with this software.

22. This software occasionally behaves in a way that can't be understood.

23. It is easy to see at a glance what the options are at each stage.