COMMUNICATION &LANGUAGE at work

Usability Testing of User Manuals

Margrethe H. Møller

(pp. 51-59)

Subscribe:

http://ojs.statsbiblioteket.dk/index.php/claw/notification/subscribeMailList

Archives:

http://ojs.statsbiblioteket.dk/index.php/claw/issue/archive

Publishing:

http://ojs.statsbiblioteket.dk/index.php/claw/about/submissions#onlineSubmissions

Contact:

http://ojs.statsbiblioteket.dk/index.php/claw/about/contact

| Bridging Theory and Practice |
http://ojs.statsbiblioteket.dk/index.php/claw

Usability Testing of User Manuals

Margrethe H. Møller

Associate Professor, Ph.D.

Institute of Design and Communication, University of Southern Denmark

Abstract

Many guidelines and several standards exist for the development of good user manuals. But even though technical writers comply with all guidelines, problems will typically arise when users apply the manual in practice. Therefore, it is useful to have real users test the manual before it is published. This article discusses user tests in the form of think-aloud tests, with examples from the research project "User Manuals for older adults".

Evaluation methods from software and web development

Within software and web development, it is good practice to test the user interface of the product before it is released. This testing often takes place as part of an iterative design process which may include an expert review, followed by one or more user tests. The expert review may be a heuristic evaluation, where one or more experts evaluate the user interface according to a set of heuristics or design principles. In the user test, representative users from the target group try to solve typical tasks using the product.

It is generally acknowledged that the two evaluation methods – expert review and user test – have different strengths and weaknesses.

The expert review is characterized by its thoroughness and includes many aspects of the user interface that tend to cause problems. But at the same time, the review is influenced by the fact that it is conducted by experts, who are familiar with user interfaces and who have definite ideas about how they should be designed. As a consequence, the experts might identify 'false' errors which would not bother ordinary users, or miss problems which arise from ordinary users choosing unorthodox paths through the systems.

The user test, on the other hand, is good at revealing which problems will be experienced by ordinary users, but has the drawback that only those parts of the system needed to solve the tasks given, will be tested. In most cases, it will not be possible to cover all parts of the system (see e.g. Rubin & Chisnell (2008) on usability testing in software and web development projects).

Consequently, it is good practice to combine expert reviews with user tests.[1]

A similar procedure could be used in connection with user manuals: first an evaluation of the user manual according to a set of guidelines (an "expert review"), and subsequently a test carried out by representative users who are asked to solve tasks by means of the manual and the product.[2]

An example from our project "User Manuals for Older Adults"[3]

Studies show that older adults find it difficult to read the user manuals supplied together with domestic technological products. One example was a study cited in Danish newspapers in 2008 (Grønvald 2008). However, domestic technology and the accompanying manuals should be designed in a way that allows older adults to use them without depending on help from younger relatives.

In the project "User Manuals for Older Adults", we investigated whether user manuals written according to "best practice" in technical writing and document design would fulfil the needs of older users, or whether for example additional explanations and feedback would be necessary. This investigation was based on literature studies and usability tests of user manuals with test persons from our target group. The product described in the manual was a digital photo frame – a technological product which would likely be used as a gift to older adults from their grown-up children.

In the project, we applied user tests to evaluate the original user manual as well as two test manuals developed by us: Test manual no. 1 was developed according to general guidelines from the literature on document design and technical writing. Test manual no. 2 was developed according to the same general guidelines, supplemented with special guidelines from the literature on ageing and user manuals for older adults.

---

[1] When Denmark introduced a new web-based land registration system in 2010, the project management assumed that expert reviews had proven that the user interface was ok, and therefore chose to leave out the user tests. This resulted in trouble, inconvenience and many support calls from lawyers, real estate agents and ordinary citizens who were intended to use the system. (www.version2.dk, 2010)

[2] As early as in 1997, Karen A. Schriver writes about the advantages of applying user tests to technical documentation in her classic book "Dynamics in Document Design" (ch. 7: "What document designers can learn from readers").

[3] This section is mainly based on the chapter "Test af brugervejledninger" ("Test of user manuals") in Møller, Christoffersen, Toft & Norlyk (2012). The project "User manuals for older adults" was carried out by a group of researchers at the Department of Business Communication and Information Science, University of Southern Denmark, from 2009 to 2011 with the purpose of developing methods of optimizing the creation of technical user manuals for older adults. See also Møller & Christoffersen (2010).

The two test manuals were checked for compliance with the guidelines several times. This step was equivalent with the "expert review" in software and web development. Subsequently, the test manuals were exposed to user tests with test persons from the target group.

User testing with think-aloud protocol

The user tests were carried out according to the think-aloud method.

Participants in a think-aloud test are typically a test person, a test moderator, and an observer (Figure 1). The test can be carried out in a professional test lab, but often an ordinary meeting room or the place where the test person would normally use the product will be suitable.

The test person will be asked to solve a number of tasks by means of the product being tested. Simultaneously, he is asked to think aloud, i.e. say out loud what his thoughts are or what he is uncertain about. As a main rule, the test person should try to solve the tasks

Figure 1: *Test setup: the Test Moderator, the Test Person with the digital photo frame and the manual, and the Observer taking notes*

without help from the test moderator, unless help is necessary in order for the test to proceed.

The test moderator will be sitting opposite to or beside the test person and will manage the session, i.e. introduce the test, present the test tasks, remind the test person to think aloud, and wind up the session. The test moderator may also ask questions, prompting the test person to articulate his thoughts. It is important to reassure the test person that it is the product that is being tested, not him. He should also be convinced that the test moderator is neutral and did not participate in the development of the product, so that the test person can freely criticise the product, if he wishes to do so.

The observer sits discretely behind the test person so that she can see what is going on and take notes about the verbalizations and actions of the test person, perhaps in forms prepared for that purpose. If the test person gives his consent, the test session may be audio or video recorded to support the written notes[4].

---

[4] Normally, the test person is asked to fill in an 'informed consent' form.

Test tasks should be relevant and realistic, and should be presented in a realistic sequence. 6 to 10 test tasks will normally be sufficient, and they should be solvable within 1 hour. After that, test persons will normally become tired and unfocused. In our experience, it is a good idea to present the test task to the test person orally as well as in writing, and to present one test task at a time, so that the test person is not aware how many test tasks there are. This is less stressful to the test person, and furthermore, the test moderator will be able to interrupt the test after one hour or sooner, without giving the test person a sense of not being clever enough to solve all test tasks.

Normally, test tasks are phrased as scenarios so that test persons see the task in a context.

Below you can see examples of test tasks from our user tests in the "User manuals for older adults" project. The task is outlined in a box, and the passages in italics are intended for the test moderator to remind him of the purpose of the question.

1.

Note: When asking this question, make sure that the test person does not have the photo frame at hand.

Your daughter has just given you a digital photo frame as a gift. **Please try to get an overview of the functionality of the photo frame – look in the manual. Don't forget to think aloud.**

Focus points: Does the test person look in the table of contents? And in the version for older adults: does he read the introduction where the product is explained? This task is also intended to give the test person an idea of the purpose of the product so that he will have a context for the subsequent tasks.[5]

2.

The word "signal source" occurs in the manual several times. **What is a signal source?**

Focus points: Does the test person use the Glossary or the Index?

---

[5] Some users are accustomed to reading the whole manual before beginning to use the product. The test team could consider giving them the possibility to do so in the test situation, as well.

9.

> Until now, you have only looked at pictures from the internal memory of the photo frame. Now, your daughter has given you a USB plug with photos of your grandchildren. You want to see them in the photo frame. **Display the pictures from the USB plug.**

Focus points: In the Index, there is a reference from USB plug to USB key, which is the term used in the manual. Does the test person look up USB plug in the Index?

The important thing here is to phrase the tasks in such a way that focus will be on testing the manual and the way in which the test persons use it – and not on testing the photo frame as such.

The think-aloud test requires a considerable amount of planning and preparation in the form of a) contact to representative users who are willing to participate as test persons, b) phrasing and testing of test tasks, and c) phrasing of interview questions which can be put after the test. All this should be included in a test plan which can serve as a kind of script for the test. In addition, a pilot test is recommended to determine if any adjustments are necessary before the actual user test.

After the actual test, test results must be analysed and presented to the colleagues who need them in order to improve the product.

A thorough usability test may run over 4 weeks: 2-3 weeks for preparation, 1-2 days for execution and 4-5 days for analysis and documentation of results (Gregersen & Wisler-Poulsen (2009).

The think-aloud test is a qualitative method suitable for investigating which problems typical users will experience in typical use cases. Normally, 4-6 test persons will be able to find the majority of problems in the parts of the product being tested. Adding more test persons will, in most cases, not reveal any additional problems (Nielsen 2000).

Think-aloud tests are <u>not</u> suitable for measuring how fast users can solve various tasks, as the very act of thinking aloud while solving a task requires extra time and mental focus. Due to the limited number of test persons, think-aloud tests are not suited for generalizations to all users of the product, either. However, they will inspire improvements to the user manual.

Results from our user tests

The 10 test persons were aged 55-77 years and had different educational and vocational backgrounds. Age-related changes are primarily genetic, but are also caused by external factors such as general living conditions and life style. Therefore, there will be large individual differences in the aging processes experienced by older adults. The age group

of 55+ was primarily chosen because it was used in other, similar projects – and not because specific age-related changes will necessarily occur at that time of life.

Test persons were initially categorized by us as novices or experts, based on their self-reported experience with technological products such as cell phones and computers.

Three versions of the manuals were tested:

1. The original manual (8 pages) was tested by one novice and one expert user
2. Test manual no. 1, designed according to general technical writing and document design guidelines (30 pages), was tested by 2 novice and 2 expert users
3. Test manual no. 2, designed according to guidelines for older adults, with additional motivating elements (Loorbach et al. 2006) such as explanations, feedback information and illustrations (72 pages), was tested by 2 novice and 2 expert users; see examples with and without motivating elements in Figures 2 and 3.

5.1 Indstilling af sprog
1. Tryk **Exit** for at komme til skærmen
   Slide show / Foto / Kalender
2. Tryk **Menu** for at åbne menuen
3. ...

Figure 2: *Example from the general version of the user manual ("Setting the display language"), with no motivating elements.*

6.1 Indstilling af sprog
Denne funktion får du kun brug for hvis der på skærmen vises engelsk tekst i stedet for dansk. Dette kan forekomme hvis foto-rammen er blevet nulstillet.

Nedenfor tages der udgangspunkt i den tilstand hvor der vises bil-leder på skærmen. For at starte det rigtige sted kan du evt. slukke og tænde fotorammen med knappen **Power Switch** ❼, se s. 10.
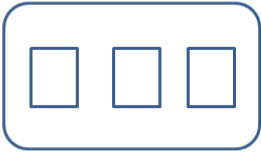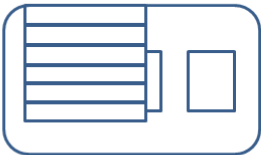
1. Tryk på knappen **Exit** for at komme til nedenstående skærm:

2. Tryk på knappen **Menu** for at åbne menuen:

3. ...

Figure 3: *Example from the older adults version of the user manual. The sequence of instructions is the same as in Figure 2, but motivational elements in the form of explanations and feedback information in the form of illustrations have been added (illustrations were anonymised for the purpose of this article).*

Communication & Language at Work

The three manuals tested had the same reader friendly font size.

Our general experience with the user tests was that on the one hand, they confirmed our expectations regarding users' preferences and behaviour, and on the other hand, they surprised us. In addition, they revealed concrete errors in the two test manuals which we had not noticed in proofreading.

The following expectations were confirmed:

a) **Step-by-step instructions are good**
Test persons expressed dissatisfaction with missing step-by-step instructions in the original version of the manual

b) **English and other foreign words are problematic**
According to both sets of guidelines – the general guidelines and those for older adults – English and other foreign words should be avoided as far as possible, or they should be explained in the text and in a Glossary.

c) **Test persons prefer short manuals**
This conflicts with the recommendations for manuals for older adults as regards readable fonts of a certain size and additional motivating elements as mentioned above.

Among the factors which we had not considered in the guidelines, were the following:

d) **Sequences of instructions should not span more than one page**
Test persons lost perspective when sequences of instructions spanned two or more pages. Also here there is a conflict with the recommendations for manuals for older adults as regards readable fonts of a certain size and additional motivating elements – all factors which require extra space

e) **Overviews which are needed all the time should unfold as a flap beside the manual**
If users need to flip back and forth between the sequence of instructions and the overview, it will be more difficult to keep a sense of perspective.

The factor which surprised us the most, because it contradicted the way we would use a manual ourselves, was the following:

f) **Users do not necessarily apply the glossary and the index of a manual**
Most test persons did use the table of contents, but to our surprise, no one used the glossary or the index – possibly because they were not highlighted in the table of contents. Glossary and index are meant for users who are not familiar with special vocabulary in the manual, and therefore one should consider how these tools can be highlighted in the manual, and how users can be motivated to utilize them.

The user tests indicated that the manual which served the users best was Test manual no. 1 which was designed according to general guidelines. In our view, this does not make motivating elements superfluous, it only means that motivating elements should be kept within limits so that the manual will retain a manageable size.

Experience with user testing of manuals

The test persons were very positive towards participating in the tests, but frustrations did occur during the tests.

The user tests were rather demanding. The test persons had to divide their attention between:

1. following a step-by-step instruction in the manual;

2. finding the right buttons at the back of the digital photo frame (by feeling their way or by turning the photo frame, and perhaps having to look up the names and positions of the buttons on a different page in the manual);

3. observing the outcome on the screen at the front of the photo frame; and

4. thinking aloud, because this was a usability test.

Another factor was previous experience with technology. Test persons will probably be more anxious the less familiar they are with the type of system being tested. In our tests with older users especially novice users of technology showed signs of emotional stress.

Therefore, it was important to reassure test persons that they were not being tested – they were contributing to the manuals being improved for the benefit of other users.

Summing up

In our view, applying user tests as a supplement to guidelines and proofreading when developing user manuals has many advantages. Testing with users from the target group is especially important with user manuals for consumer goods, as they are to be used by many different user groups with different backgrounds. User tests are resource demanding, but they may save the company the costs of support calls, and they will most certainly save users unnecessary frustrations.

Author

Margrethe H. Møller

Assoc. Prof., Ph.D., Institute of Design and Communication, University of Southern Denmark

Margrethe H. Møller holds a master's degree in modern business languages (German and English) and a Ph.D. in computational linguistics. Her current research interests are technical communication and user participation. Her teaching areas include German language, technical communication, and evaluation methods for usability and user experience of web applications.

Contact:

mhm@sdu.dk

Works Cited

Gregersen, I. & Wisler-Poulsen, O. (2009): *Usability. Testmetoder til mere brugervenlige websites på internettte.* Grafisk Litteratur.

Loorbach, N. et al. (2007): Adding Motivational Elements to an Instruction Manual for Seniors: Effects on Usability and Motivation. In: *Technical Communication* Vol. 54(3), 343-358.

Molich, R. (2003): *Brugervenligt webdesign*, Nyt Teknisk Forlag.

Møller, M. H., Christoffersen, E. (2010): "Technical Manuals for the Elderly - Can Controlled Language Principles Help?" In: Heine, Carmen/Engberg, Jan (eds.): *Reconceptualizing LSP*. Online proceedings of the XVII European LSP Symposium 2009.

Møller, M., Christoffersen, E., Toft, B. & Norlyk, B. (2012): *Brugervejledninger for ældre. En formidlingsrapport*.

Nielsen, J. (2000). Why You Only Need to Test with 5 Users. In: *Jakob Nielsens Alertbox*. http://www.useit.com/alertbox/20000319.html (last visited Apr. 15, 2012)

Rubin, J. & Chisnell, D. (2008): *Handbook of Usability Testing*. Wiley Pub.

Henrik Grønvald: "Vi forstår ikke brugermanualer". (2008, Nov. 22). *Jyllands-Posten Puls* p. 10.

"Tinglysningschef fjernede brugertest trods kontrakt-krav". (2010, Sept. 3). www.version2.dk.