Usability testing as part of the questionnaire design process

Petri Godenhjelm¹

¹Statistics Finland, e-mail: petri.godenhjelm@stat.fi

Abstract

Statistics Finland had a two-year project on the evaluation and testing of business web questionnaires, and to redevelop them and give guidelines for future work on questionnaire design and testing. The usability testing was done as cognitive interviews with the concurrent think-aloud method and recording of all interviews. The screen recordings were done with the Dream Screen software, which produced video clips with audio on the answering process. In usability literature, the concept of usability consists of three dimensions: effectiveness, efficiency and satisfaction. These aspects were considered during the conducting and analysing of the test interviews.

This paper discusses the experiences gained from the testing and from redesigning work, which appeared to be the most challenging part of the work. The usability testing was an input into the redesigning process. This was a new kind of situation in the whole organisation and created the need to learn new ways of organising this work. As of 2007, it has been possible to answer all major and permanent surveys of establishments via the web, mostly with an in-house XCola application and with software from an outside provider. This means that the web questionnaire designing was done mostly in the early years of the 2000s. Since then the web questionnaires have been developed in many ways and people’s expectations on the web applications have risen. Although the data collections via the Internet have been regarded as successful in many ways, continuous evaluation and development of the web questionnaires are still needed to ensure high quality of the collected data and to keep the respondents motivated.

Keywords: usability, questionnaire design, testing

1 Driving force for usability testing

Statistics Finland conducted its first electronic collection of data in the late 1990s. The actual transfer towards Internet-based data collection began in 2001 when the building cost index implemented the first web-based data collection. A large-scale web-based data collection was introduced in the statistics for the years 2005 and 2006, when the electronic response option could be used for all enterprise data collections. An Internet data collection...
method has been available to all over 65 major and permanent data collections from the end of 2006.

A decade later, in 2010, the web surveys for enterprises have grown to approximately 70 per cent and the share of the paper form has declined from year to year. In reality, the move was done very quickly and the design was still based more or less on paper questionnaire images. At that time, it was not thought so much that the web survey design has its own specialities and how user-centred design principles should be followed. As a whole, web collections have been interpreted as a success in many ways. The average response time of surveys has reduced and the received data have been of better quality.

1.1 Under the umbrella of the business data collection programme

The European Commission published an agenda to reduce administrative burdens on business in January 2007. The EU’s action plan was launched in spring 2007. The action plan aims to measure costs arising from EU legislation and its implementation and to reduce the administrative burden by 25 per cent by 2012. The Finnish Government adopted a decision in principle on a national action plan for the reduction of the administrative burden on business for the years 2009-2012. The goal is to reduce the administrative burden on enterprises in 2006, by 25 per cent by the end of 2012. The proportion of statistics in the administrative burden in Finland has estimated to be at slightly over one per cent.

As part of the European Commission and national action plan, Statistics Finland has run a programme for developing business data collections in 2007-2011. The purpose of the programme is to reduce the response burden and develop response relations. Under the programme there are plenty of different kinds of projects, such as co-operation with large companies, promotion of direct data collections, co-ordination of samples, development of receiving administrative data, development of the Business Register and measurement of response burden. In 2010 the measured response burden was 181 staff-days or EUR 8 million.

Under the work on response burden measurement, the survey questionnaire on response burden was pre-tested at enterprise sites. To be able to test the questionnaire draft, it was essential that the respondents answered to the actual questionnaire at hand. The test produced valuable information on developing the web questionnaire on response burden. Besides the main focus of the test, there were new findings on how the respondents behave when they answered to the web questionnaire. It was concluded that there is still some work to be done to design user friendly and high-quality questionnaires. And many web survey elements and technology have also developed from the way the first waves of web questionnaires were designed.
The main impetus for the need of usability testing and developing electronic questionnaires came after Statistics Finland’s personnel (the IT unit, the responsible statistics unit and management) saw the actual response process of their questionnaires. The screen capture software brought the experience of respondents to the office. The response process was made visible, and this concrete experience was the signal that led to an action in the electronic questionnaire development. The development programme on Data Collection from Enterprises decided to start a project at the beginning of 2009 on the usability of electronic questionnaires. The aim of the project was to do usability testing with business web questionnaires in businesses’ own environment. The results of the response burden were an important input to the decision process on which questionnaires will be tested. Even so, in the process of redesigning this also produced the most challenges. The most burdensome questionnaire also has the most difficult design implementation process. Also, expert reviews and a survey responsible unit keen on the testing process were important factors to be included in the project.

2 Procedures on testing and evaluation

The overall objective of testing was to determine how the respondents fulfil the tasks the questionnaire entrusts to them, how they navigate and how they understand the questions and response options on the data collection application, the web questionnaire. An important goal was also to find out how the respondents use different kinds of information systems and other records in providing answers. For example, it was examined whether the information was stored in the company in the same way as it was asked in the questionnaire and how the person responded to the questionnaire. The approach was in line of a hybrid response process model, which brings new steps to the traditional cognitive response model (Willimack and Nichols, 2010). Even if the web questionnaire and usability issues were in main focus, the other parts of the response process emerged as well in test situations. In the first round of empirical testing the focus was not on any special part of the questionnaire, but in the second round it was more on aspects that were found and developed between the testing rounds.

The surveys tested were Finnish affiliates abroad, Business Register inquiry for multi-establishment enterprises, Financial statements inquiry for enterprises, Inquiry on business services, Inquiry on private sector wages and salaries, Information technology and electronic commerce in enterprises, and Finance of housing corporations. As a whole, 46 enterprise visits were done.

2.1 Video recording

The testing method was a usability test, which is similar to cognitive interviews. Respondents were instructed to think aloud during answering the questionnaire. The aim
was to go through the whole questionnaire. In some occasion, this was not possible because all the information or the personnel needed were not available or there were time constrictions. In those situations, a test procedure was arranged so that all parts of the questionnaire were evaluated at least partially. Thus it was possible to get an overview of the entirety of answering tasks. In the second round of testing, the focus was in the parts that were redesigned.

The answering process in computer screen was recorded with audio. Test videos could be exploited easily by many people involved in process in terms of usability issues as well as substance problems and communication problems. In line with response burden measurement results, an answer to the questionnaire produced colourful pictures of what is really happening and behind those quantitative figures. With real respondents in the middle of a response task, the message is realised in a very straightforward way.

2.2 Solutions behind questionnaire design

All interviews clearly showed how the layout and symbols guided the navigation on the questionnaire and, at the same time, the comprehension of the questions as well. However, instructions contain also information that is relevant to answering, the designer should consider thoroughly what information is the most important to communicate to ensure that the received answers will be comparable. At this point, the designer must make a strategic decision about what, how and how much information should be communicated to the respondent. The meaning of concepts is communicated not only in questions, but in definitions and instructions as well. In the web questionnaire design, it is possible to use more powerful tools than in the paper mode in presenting out all the relevant material. Perceptions on the test interviews encourage to follow in the design process the guidelines which tell to consider reformulating important instructions to questions and to incorporate instructions into the question where needed and to avoid placing them out of sight of the respondent and also to convert narrative paragraphs into a bulleted list (Morrison et. al. 2010)

Vagueness in the navigation path, meaning here the intended order in which answering should progress, caused problems in some of the tested questionnaires. In the web questionnaire design there are efficient ways to direct the response process, but it should be taken care at the same time that you do not design things that will easily make the respondent get lost in the web questionnaire structure. So a good design is needed in the way that the respondent is well aware where she/he is in the navigation process and how the design of different buttons is taken care of. It is also important to test the questionnaire with different kinds of browsers, because your plans may not show in the same in different browsers.
A questionnaire designer must make far-reaching decisions. On the one hand, he/she must make sure that the essential content of the question is communicated, but on the other, ensure that the question and the related instructions are not read superficially or bypassed altogether. How definitions and instructions are used and placed in a web questionnaire have important implications for the comparability of responses. It is also essential to keep up good motivation to continue the answering process and give profound answers, which will be reproduced in the quality of statistics. One conclusion of the project was that a design process needs different range of expertise like in designing questionnaires for multipopulation research (Harkness et al. 2010). This means expertise of substance, software, testing and also knowledge from literature of relevant research on usability, questionnaire design and data collection.

2.3 Examples of how to help respondents in their tasks

A few examples of how to help the response task were developed in the survey questionnaire for Finnish affiliates abroad when the questionnaire was re-developed after usability tests. One important part in the answering process is when the respondent actually starts to make an overview of the whole task ahead of her/him. Many things are told already in the pre-information letter, but an image of the questionnaire is not perceived before the answering starts and the navigation has gone to the end. Still the respondent might want to know beforehand what is ahead for many reasons. To know if help is needed from colleagues, what kind of information she/he needs, how much time it will take, etc. One tool for this is a pdf file where the respondent can see all the modules of the questionnaire and average lengths of time they will take to answer in the name of one model enterprise (appendix 1). This is also important because the questionnaire is dynamic as the remaining parts of the questionnaire are generated only after the second part is ready. This means that the order of responding is predetermined. As a whole, this is a helping tool to the respondent on how to design the whole response task, not only answering to the questions themselves.

Another example is to offer an alternative way to give answers by using a fixed-format csv file (csv is comma separated values). The respondent needs to download a pre-filled file on the affiliates that we are asking about. Then the respondent can save the file onto her/his own computer in Excel or csv format. After this, the respondent can upload the file completed to the questionnaire page from her/his own computer. Answering is done by filling in the data indicated by the column headings. The data on each affiliate show on a separate row. CSV is an alternative to fill in the same part of the questionnaire in some cases hundreds of times. It needs some preparations and capability to use spreadsheet tools, but will save a huge amount of time in big enterprises, which will also motivate them to use it (end of appendix 1).
3 Good design behind motivated respondents and preventing response burden

The response burden could be lessened by high-quality design of web questionnaires. Compared to the paper mode, where the user interface is much more stable than the web, the management of future design work will have an important role. Now we have already seen that there are so many changes in the web environment and technology that conventions in the web have already changed in some areas. This means that not everything is done when you have once moved your questionnaires to the web. It means that survey practitioners should all the time be aware of the situation in the web environment and the changes it brings. And every now and then also to react to these changes with evaluations and testing procedures. This work is a clear statement for taking care of respondent relations with businesses. This hopefully will also lessen the response burden and raise the motivation to give good-quality answers.

References

Appendix 1: Example Finnish affiliates abroad 2009

Please familiarise yourself with the different parts of the questionnaire and the estimated response times. Views of an imaginary group *Esimerkkikonserni* were used in the images.

1 Information about respondent

Please check, complement or correct your contact details and those of the group you represent as well as the first and last days of the accounting period ending in 2009.

It typically takes one to five minutes to respond.
2 Affiliates abroad

Please check, complement or correct the data on groups located abroad during the accounting period ending in 2009. Groups include affiliates, branches and associate companies in which the Finnish group parent has at least 10 per cent of the voting rights.

The length of time for answering depends on the number of units located abroad and on how easily the data inquired can be found in the group's internal systems. When the data are available, it typically takes one to three minutes to answer per enterprise.
3 R&D investments

Please respond to the questions about research and development (R&D) conducted abroad during the accounting period ending in 2009 and on investments into targets abroad. The length of time for answering depends on how many countries the activity inquired has concerned and how easily the data inquired can be found in the group's internal systems. When the data are available, it typically takes 15 to 30 minutes to answer per enterprise.
4 Activity of affiliates

Please choose first whether you answer with separate enterprise-specific questionnaires or by one fixed-format file. It typically takes three to five minutes to familiarise yourself with the alternatives.

The length of time for answering with both separate enterprise-specific questionnaires and with a fixed-format file depends on how many enterprises have to be responded for and how easily the data inquired can be found on the group's internal systems. In addition, the time of answering with a fixed form file is highly dependent on how used the respondent is to using the file format in question. When the data are available, it typically takes 1 to 5 minutes to answer per enterprise.
When the data are available, it typically takes one to three minutes to answer per enterprise one fixed format csv-file. If the data are combined mechanically, for example by a unique company group id, the response time is much shorter.