EESW09

Usability principles in testing questions for establishment surveys

Petri Godenhjelm

Statistics Finland

FIN-000022, Statistics Finland

www.stat.fi

petri.godenhjelm@stat.fi

Recently, establishment surveys have adopted approaches developed in cognitive psychology and information process studies. In this household survey based model, questioning and answering is seen as a four-phase process, which include understanding what a question means, retrieval of relevant information, judgment formation and lastly editing and reporting answer (Sudman et al., 1996, Tourangeu et al., 2000). All this has enhanced understanding of establishment surveys and lead to the additional phases to the response process specific to establishment surveys. The first new phase is encoding in memory/record formation, second is selection and identification of the respondents and third step before understanding is assessing establishment's priorities and the last step in the whole process is releasing of the data (Willimack, 2004).

The theoretical development steps have generated novel testing methods for establishment surveys. The most conventional methods are focus groups and cognitive interviewing, which seem to be the mostly used method. An overall goal in this process is to gather higher quality data on establishments and minimize response burden. This is all about a high quality survey questionnaire design, which has impact also to a response rate of surveys.

New ideas to the area of questionnaire design and testing come from the usability theories and practises. The focus of usability testing in a survey context is an interaction between computer and respondent, and especially between software of a questionnaire and respondents. The point of view in this presentation is that an electronic questionnaire is an electronic application which should be designed with same usability principles as many other applications, which people use in a daily life. Of course a survey questionnaire has its own specialities and conventions, but these should not be in a discrepancy with basic usability principles.

The aim of this paper is to depict the usability as a broad frame and how usability principles have been implemented in a questionnaire design and evaluation. Some words are also written on plans Statistics Finland has in usability tests of web-surveys in a two year project started in 2009.

Usability and user experience

The concept of usability consists of three dimensions: effectiveness, efficiency and satisfaction in a specified context of use. Here this context is seen as a response situation in establishment internet surveys. These aspects are considered during conducting and analyzing cognitive interviews on enterprises. Reflecting the general approval of these aspects, also the ISO standard on usability recognizes each of them. Usability approach emphasizes the importance of empirical usability tests. Designers of an application are blind to the aspects of novice user has. This is also valid in a questionnaire design of web-surveys.

The related concept to usability is user experience which refers to feelings a person has in using an application in hand. Some see that this belongs to the concept of usability, but Sinkkonen et al. (2009, 18) see that then usability concept represent one desirable feature which belongs to an application, a user experience refers to a quality of experience user has. They say that during 80's and 90's people talked only about usability, in which one aspect was satisfaction. But in a new millennium it was not enough to talk about satisfaction only, people expect more from applications. At the same as internet has become an everyday environment for plenty of services, it is not enough that an application is internet, it has to be pleasing to use and also match up to expectations people has up to a service and an application.

Usability and survey design

Usability theories and new visual science theories on a questionnaire design has its roots on Gestalt principles of psychology. These seven principles are 1) Proximity (placing objects closer together will cause them to be seen as a group), 2) Similarity (objects that are the same color, shape, size, or otherwise the same will be seen as group, 3) Elemental connectedness (elements connected by other elements tend to be viewed together, 4) Continuity and continuation (elements that continue smoothly across a field will be seen as grouped together, 5) Common region (elements with the same closed region of space will be viewed together, 6) Closure (elements forming a "closed" figure will appear grouped together, 7) Pragnaz (elements that are simpler and the same shape are easier to perceive and remember. (Dillman, 2008, 465-470.)

Dillman (2008, 470) speaks also that respondents draw information not only from words, but from other visual languages as well. He states that respondent use verbal language, numerical language, symbolic language and graphical language. In a survey context this means that respondents must first perceive the information presented them visually and give it their attention. This lead Redline and Dillman (2002, 191) to present that in to the question-answering model should be added another phase, perceiving and attending. Information in questionnaires that is not perceived or attended to cannot be used in the other steps of answering. The main point should be that a survey application supports a question-answering process of respondents.

In interview surveys the standardization and an interviewer play an important role and take care that answers are comparable. In a self-administered survey, there is not an interviewer available who could help respondents to a right navigation. This means that a questionnaire itself should attend that respondents answer all questions and will offer to all respondents same questions and definitions. All this makes a questionnaire design a little bit more challenging.

Census Bureau has developed guidelines on questionnaire design where is also many concrete specifications on visual solutions. These include for example guidelines on the display of answer spaces and response options and guidelines on eliminating visual clutter. Consistency of fonts and other elements in questionnaires is important in developing user friendly questionnaires and means that every organization has to develop their own standards. One important guideline to follow is to locate question-specific instructions into the questionnaire where they are needed and avoid placing them in a separate web page. (see more Census Bureau, 2008.)

In Appendix is written down Nielsen's heuristics, which are well known rules of thumb in usability approach. In each heuristic are some preliminary thoughts how they could be connected to a survey. There is also presented some usability problems we found in our project and which could illustrate Nielsen's heuristics in a questionnaire design.

The project on usability of establishment surveys

The program for developing business data collection is going on in Statistics Finland. In this frame the testing of a response burden questionnaire was done in the summer 2008, using two different ongoing questionnaires as a platform. The main impetus for a need of usability testing and developing electronic questionnaires came after personnel (it-section, responsible statistics unit and management) saw the actual response process of their questionnaires. The screen capture software brought the experience of respondents into the office. The response process was made visible, and this concrete experience was the message which led to an action in the electronic questionnaire development.

An internet data collection method has been available to all over 60 major and permanent data collections from the end of 2006. The aim of the project is to do usability testing with almost 10 data collections during 2009-2010. We use the results of the response burden as an input to the decision process on which questionnaires will be tested. A first round of empirical testing could be seen as a user research, because focus is not in any special part of a questionnaire. The second round is more like normal usability testing and focus is more on aspects which will be found and developed between testing rounds.

The screen recordings of test interviews so far have been an important part in bringing the response process to the eyes of statistics personnel. With real response process observations, it is easier to make personnel aware of problems and also to make development work smoother. When using recording software Dream Screen, video clips are in the server behind password and could be seen with internet just clicking the hyperlink of a recording. The questionnaire testing unit of Statistics Finland administers the use of links. The main problem found in first tests was the moving of respondents between pages and screens in questionnaires. The navigational path was not always clear to respondents and this produced unnecessary burden to respondents.

As a broader goal in the project is to develop a model of usability testing process in Statistics Finland and develop design standards in electronic questionnaires. A long term goal is to launch user-centered design methods culture and organization, which will be a tool to reach a good usability and a good user experience. In an electronic questionnaire, which is good in usability, questions are easy to find, to understand and to glance at. Navigation and functions are easy to use. Lay out is clear and constructed so that it supports an order of questions. Questions are grouped according to a topic and so that all relevant things are visible. Lay out should support answering and should be in line with a graphic image of a survey organization.

References:

Census Bureau. Economic Directorate Guidelines on Questionnaire Design (2008): http://www.census.gov/srd/Economic_Directorate_Guidelines_on_Questionnaire_Design.pdf

Dillman, Don A. (2007): Mail and Internet Surveys. The Tailored Design Method. 2007 Update With New Internet, Visual, and Mixed-Mode Guide. John Wiley & Sons, New Jersey.

Redline, Cleo K &, Dillman, Don A (2002): The Influence of Alternative Visual Designs on Respondents' Performance with Branching Instructions in Self-Administered Questionnaires. In Groves, Robert M, Dillman, Don A, Eltinge, John L & Little, Roderick J.A. (2002): Survey Nonresponse. John Wiley & Sons, New York.

Sinkkonen, Irmeli, Nuutila, Esko & Törmä, Seppo (2009): Helppokäyttöisen verkkopalvelun suunnittelu. (Design of User-Friendly Internet Services). Tietosanoma. Kariston Kirjapaino Oy.

Sudman, S., Bradburn, N. M. & Schwarz, N. (1996): Thinking about answers. The application of cognitive processes to survey methodology. Jossey-Bass Publishers. San Francisco.

Tourangeu, R., Rips, L. J. & Rasinski, K. (2000): The psychology of survey response. Cambridge University Press. Cambridge (UK).

Willimack, Diane K., Lyberg, Lars, Martin, Jean, Japec, Lilli & Whitridge, Patricia (2004): Evolution and Adaptation of Questionnaire Development, Evaluation, and Testing Methods for Establishment Surveys. In Presser, S., Rothgeb, J. M, Couper, M. P, Lessler, J. T., Martin E., Martin, J. & Singer, E. (2004): Methods for testing and evaluating survey questionnaires. John Wiley & Sons, New Jersey.

Appendix. Nielsen's heuristics and their links to survey and to the project
http://www.useit.com/papers/heuristic/heuristic_list.html

1) Visibility of system status. The system should always keep users informed about what is going on, through appropriate feedback within reasonable time. This is straight applicable to questionnaire context. First results in Statistics Finland's project illustrate this in questionnaire where respondents had to fill many sub questionnaires where was asked information about kind of activity units and NACE codes.

2) Match between system and the real world. 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. How to operationalize successfully is everlasting dilemma in production of statistics.

3) User control and freedom. Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo. This should always keep in mind in electronic questionnaires.

4) Consistency and standards. Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions. This should be a guideline in questionnaire design in all questionnaires an institution produces.

5) 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. This should be applied in questionnaires with care. You can't use too many confirmation options. And placing should be thought carefully.

6) Recognition rather than recall. Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate. This is connected to the use of instructions and definitions. Their placement is important part of a questionnaire design. Always should be in mind if respondents really are capable of recalling information from texts which are behind hyperlinks or in other documents.

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. This might be difficult to apply in questionnaires, because tailoring might weaken comparability of answers.

8) Aesthetic and minimalist design. Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility. This is an important guideline. It is well known in survey literature, what respondents use all the information available when they answer to a question.

9) Help users recognize, diagnose, and recover from errors. Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution. This should be valid in survey instrument design as well.

10) 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. If help is provided, it should not be an obstacle in smooth answering. Minimalistic design principle should be followed.