Using Internet based questionnaires as a method in media technology research

DM2572 - Theory and Method for Media Technology

2012-11-30

Joel Arnkvist
arnkvist@kth.se
Terese Nothnagel
tereseno@kth.se
Johan Olsson
johaolss@kth.se
Gustav Rannestig
gusran@kth.se

Introduction

Web based questionnaires is a convenient and commonly used research method. It has several advantages but also disadvantages to other types of research methods. Here we aim to assess how and when internet-based questionnaires should be used and designed in a suitable way for media technology research. This article can be seen as a guideline for using web based questionnaires as a research method.

Research questions

What are the general key-points to keep in mind when using internet-based questionnaires as a research method? What are the specific advantages/disadvantages compared to other ways of conducting questionnaires?

When to use Internet based questionnaires?

In this article one aim is to conclude when it is appropriate to use internet-based questionnaires as a research method. With internet-based questionnaires we in this case mean questionnaires that are available on a website for the target group to fill in. Online questionnaires can be used to an advantage when time and resources is an issue thanks to the easy and rapid spread through social media, forums etc. It can potentially reach a global audience which for example can be useful in finding reasonable numbers of respondents for a rare target group. (Wyatt, 2000) In addition it can also give a larger more diverse sample. Online tools like Google form or alike also makes data analysis easier than for regular questionnaires used as research method. Another advantage is the ability to provide a high level of anonymity which can increase self-esteem whilst reducing social anxiety. (Fox et al, 2003)

Internet based questionnaires is mainly to be used as a quantitative research method since one has a good opportunity to reach many respondents. It can however also be used for collecting qualitative data if the questions are designed in that way but one should bear in mind that the researcher cannot see the respondent's reactions, body language, gestures or alike. Nor are spontaneous follow-up questions a possibility.
The use of internet based questionnaires can also be appropriate when the researcher has a very dynamic set of questions to be asked; for instance if the respondent answers "Yes" to a specific question, the following 10 questions in a paper questionnaire may be irrelevant for the respondent, an internet based form can then in many cases skip them automatically for the respondent.

How to distribute the questionnaires and determine the audience?

How a questionnaire should be distributed is depending on many factors. One factor is how much control one would like to have over who is answering the questionnaire. If one would like strict control over who is answering the questionnaire, one could email a link to the questionnaire to known recipients in order to make sure that only those who received the email can answer the questionnaire. If a more modest degree of control is wanted, the recipients of an email with a link to a questionnaire can be asked to forward the link to their respective friends or colleagues. (Wyatt, 2000) This is called the "snowball" technique. (Forsström et al. 1999) When using the "snowball" technique, one need to keep in mind that the calculation of the response rate can be harder or even impossible since the number of people receiving the email with the link to the questionnaire will remain unknown. If one is concerned about *who* is answering the questionnaire, and *that* people answer it, reminders can be sent out if the questionnaire is distributed by email. Reminders are a key way to enhance response, but there is always a concern that the quality of data returned in the second or third rounds may have been reduced by the irritation of participants receiving reminders. (Oppenheim, 1991) Sending out reminders is more difficult if the "snowball technique" is used compared to if only known recipients of an email address are answering the questionnaire.

Another way of distributing questionnaires is by posting a link to them in internet discussion groups such as newsgroups or bulletin boards. (Fox et al, 2003) This is more of a passive method than emailing people directly. One can choose to post the link in a specific forum or in a more general forum depending on who one primarily wants to answer the questionnaire. This way of distributing the questionnaires is more suitable if one wants anyone and everyone to answer the questionnaire, with no control over the type of participants. One could also choose to send an email with a link to a questionnaire to random people of a general mailing list. (Wyatt, 2000)

One need to keep in mind that the web is a very public place and unless steps are taken to limit access to a questionnaire, it may be found and responded to by people who are not among those sampled by the researcher. This can either happen by accident or maliciously. Since one only has to click the submit button to respond to a web-based questionnaire once it is filled out, it is also quite possible for respondents to either mistakenly or purposefully submit multiple copies of their responses. (Solomon, 2001)

How to design the questionnaire and asking the right questions?

When you have decided to do a questionnaire it is very easy to create and distribute them, however it is hard and it requires a lot of work to get reliable and valid data from a questionnaires. (Boynton & Greenhalgh 2004, Fowler Jr. 1995, Murray 1999)
When doing questionnaires one tip, especially for those new to questioners or those that are not confident in their questioner skill, is to look for already validated and published questionnaires. (Boynton & Greenhalgh 2004) Then one can save time and resources with the benefits from also comparing the results with previously done studies.

When using questionnaires as a research method there is a lot you should think about because the validity of the results depends critically on the design of the questions. One of the most important things for a good question is that all the participants answering it should understand it in a consistent way and in the way that the researcher expected it to mean. But avoiding misunderstanding or ambiguous questions is also one of the hardest things to achieve. (Fowler Jr. 1995)
The simplest way to make it easier for the participants filling out the questionnaire is to make the questions easy to read. One should also use words that are appropriate for the target group, double barrel questions should be omitted or split and double negatives question should not be used. (Murray 1999)

In order to design a good questionnaire it is important to think about the aim for the questions, both because it helps to formulate the question and also because it helps with reducing redundant or unnecessary questions. (Fowler Jr. 1995)
You should also think about whether to use open or closed questions. Open questions produce better validity and reliability than closed questions (Krosnick 1999) because they allow participants to express themselves in their own words instead of choosing from predetermined answers. However open questions are hard to analyze and if you don't have the time, skills and resources for the analysis you may end up wasting the participants time. (Boynton & Greenhalgh 2004)
One thing to think about when using yes or no questions is to mix questions that are formulated positively and negatively because people can be more incline to say yes than no. (Boynton & Greenhalgh 2004, Krosnick 1999)
It is also important to think about the placing of the questions and the question alternatives. Questions should follow a logic structure but questions that are posted earlier are often believed to be more important than others, the same thing corresponds to questions with multiple choices. Doing a questioner can be a memory intensive task for the respondents and first and last option are many times picked more often than others because of memory load, a suitable answer can sometimes be given by the respondents instead of a sincerely answer due to fatigue. (Murray 1999, Krosnick 1999)

What can be done with the data collected from an internet based questionnaire?

Coding

If a questionnaire contains qualitative variables, it's most likely a good idea to find someway of translating these into numerical values to get data that's more useable. This can be called the process of coding the answers. There are some standards that are used in statistics, and it can be good to follow these to some extent, especially if the data is to be used by another researcher than you.

Two general rules when coding the answers are to be consequent and to always provide a code for a non-answer. (Dahmström, 2005) For example if you have many yes/no/don't know questions, these should be consistently coded; say 1 = yes, 2 = no and 9 = don't know.

If you use ranges or ask the respondent to list variables in a specified order, for example; best to worst, you should decide if the variables should be coded in descending or ascending order, so 1 is worst 2 is a little better etc.

If multiple answer questions are used, one have the choice of either combining the codes so that if the respondent choose alternative 1 and 2 the code becomes 12, or each alternative can be seen as a single yes/no question so yes or no becomes 1 or 0 respectively.

Its worth mentioning that Internet based surveys has the potential to drastically lower the amount of work and the cost spent on coding data, since this process can be automated to a greater extent.

Finding relations between variables

The statistical operations that can be performed on a dataset collected through an internet-based survey do not differ from any other data. The difference lies in the potential sampling and response rate errors that have to be taken into consideration when using data collected through such a survey.

The most basic operation one can do with collected data is to perform operations on individual variables. First of all we can summarize a survey by looking at the individual frequencies of variables, such as 45% of the sample n answered “yes” to the question X. We can also look at the type-value, median-value and mean-value of a variable. The type-value is the value that has the highest frequency. The median and mean-value requires the data to be coded in such a way so that individual answers can be compared against each other.

Finding relations between variables can also provide a good insight into what the collected data tells us, and can be a key point in producing findings. A method to find relations between variables is to set one of the variables as the dependant variable, and one or more as independent. If we set Y as a dependant variable and X as an independent variable, the general idéa is to find an approximation of Y expressed in X, for example a linear regression would be Y = a +bX. If such a relationship can be found we can then test its strength by looking at how much the real Y corresponding to X differs from the predicted Y we get from Y = f(X).(Dahmström, 2005)

Potential errors when using internet-based questionnaires

The errors that we have to take into consideration when dealing with Internet-based surveys deals with bias of sampling and bias of return. The bias of sampling deals mainly with the access to the internet, the comfort of using the internet and the fact that most internet surveys are self-selecting. (Zhang, 2000) Access to the internet provides a sampling bias if the target group of the survey might include people not having access to the internet. Comfort of using the Internet can also produce a sampling bias since some parts of the target group might be more likely to not feel comfortable with answering a survey through the use of a computer.

The problem of self-selecting is a hard one to overcome, and certain groups might be more likely to respond to a survey if they have the choice to do so or not. These limitations of Internet based surveys have led researchers to argue that Internet based surveys are limited in their use for targeting general populations, but can prove to be advantageous over other survey methods when the target group can be said to be tech-savvy. (Cook et al., 2000)

The bias of return occurs due to the limited control that researchers have over internet-based questionnaires. As mentioned in the section “how to distribute the questionnaires”, internet based questionnaires introduce problems of multiple answers and answers from people outside the target population.

Conclusion

In this report we have presented general key-points that should be kept in mind when using Internet-based questionnaires as a research method. Internet-based questionnaires provides the opportunity to reach a target group that is spread over a large geographical area, and has advantages in terms of data analysis, that it requires less resources and maybe cost less compared to other questionnaire distribution methods. Internet based questionnaires also has the benefit of providing the possibility of crafting surveys that are more responsive and adapt themselves to the respondents, which is hard to to with printed questionnaires. The disadvantages are mainly the lack of control of response rates and the risk of sampling bias occurring.

Reference list

Jon A. Krosnick (1999), Survey Research
http://www.annualreviews.org.focus.lib.kth.se/doi/abs/10.1146/annurev.psych.50.1.537

Petra M. Boynton & Trisha Greenhalgh (2004), Hands-on guide to questionnaire research: Selecting, design, and developing your questionnaire
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC420179/

P. Murray (1999), Fundamental issues in questionnaire design
http://www.sciencedirect.com/science/article/pii/S0965230299800745#

Floyd J. Fowler Jr. (1995): Improving survey questions: Design and evaluation
Sage Publication, Inc
http://www.sagepub.com/books/Book4994

Dahmström, K. (2005) Från datainsamling till rapport -*att göra en statistisk undersökning*
Studentlitteratur, Lund
http://libris.kb.se/bib/9892263

Zhang, Y. (2000) “Using the Internet for survey research: A case study” *Journal of the American society for Information Science* vol. 51, Issue 1 http://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-4571(2000)51:2%3C%3E1.0.CO;2-D/issuetoc

Cook, C., Heath, F., Thompson, RL. 2000 “A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys” *Educational and Psychological Measurement* vol. 60, issue. 6 http://epm.sagepub.com/content/60/6/821.abstract

Forsstrom J, Rigby M, Roberts R, et al. Key Recommendations of the TEAC-H Project: Toward Evaluation and Certification of Telematics Services for Health. Turku, Finland: University of Turku, 1999

Jeremy C. Wyatt (2000) “When to Use Web-based Surveys”. *Journal of the American Medical Informatics Association, Volume 7, Number 4, Jul / Aug 2000* http://www.ncbi.nlm.nih.gov/pmc/articles/PMC61446/

Oppenheim AN. “Questionnaire design, interviewing and attitude measurement”. London, UK: Pinter Publishers, 1991. http://onlinelibrary.wiley.com/doi/10.1002/casp.2450040506/abstract

Jezz Fox, Craig Murray & Anna Warm (2003) “Conducting research using web-based questionnaires: Practical, methodological, and ethical considerations”. *International Journal of Social Research Methodology, Volume 6, Issue 2, 2003. http://www.tandfonline.com/toc/tsrm20/6/2*

Solomon, David J (2001) Conducting Web-Based Surveys. [Journal (On-line/Unpaginated)] http://cogprints.org/2357/