



HAL
open science

Think-Aloud Surveys

Lene Nielsen, Joni Salminen, Soon-Gyo Jung, Bernard J. Jansen

► **To cite this version:**

Lene Nielsen, Joni Salminen, Soon-Gyo Jung, Bernard J. Jansen. Think-Aloud Surveys. 18th IFIP Conference on Human-Computer Interaction (INTERACT), Aug 2021, Bari, Italy. pp.504-508, 10.1007/978-3-030-85607-6_67 . hal-04291264

HAL Id: hal-04291264

<https://inria.hal.science/hal-04291264>

Submitted on 17 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Think-Aloud Surveys

A Method for Eliciting Enhanced Insights During User Studies

Lene Nielsen¹ Joni Salminen² Soon-Gyo Jung² Bernard J Jansen²

¹IT University Copenhagen, Copenhagen S, 2300, Denmark

²Qatar Computing Research Institute, HBKU, Doha, Qatar
Lene@itu.dk

Abstract. In a user experiment, we tried out a novel data collection approach consisting of combining surveys with the think aloud method. We coin the phrase “think-aloud survey method”, where participants think-aloud while completing a questionnaire. We analyzed the transcripts and found that the think aloud survey provides deeper insights into the reasoning behind the participants’ Likert scale choices and responses to open-ended questions, along with an enhanced understanding of the tasks and prompting the participants to explicate their choices. All resulting in more insights for design. The think- aloud survey method further offers other insights into usability issues and encourages the participants to provide system improvement proposals. The think-aloud survey method is useful for researchers and practitioners applying a large number of usability evaluations and wanting to dig deeper into the motivations for choice.

Keywords: Surveys, tests, interface design.

1 Introduction

This work reports from an instance of combining two well-known data collection techniques: surveys and the think aloud (TA) method resulting in something greater than the sum of its parts. Alongside an experiment testing of the interactive persona analytics system [1] in a real-world setting of an organization that uses the system for marketing and communication, we applied what we dub the “think aloud survey” (TAS) that combines quantitative and qualitative research by the use of the think-aloud protocol with a questionnaire. The novelty of the method is that, typically, user experiment participants are asked to complete surveys in silence. In contrast, our suggestion is to ask the participants to explicate their cognitive processes verbally while completing the survey, which we found have multiple benefits for data richness and quality.

In our experience, the TAS method provides nuanced insights into the participants’ understanding of the questions and spurs them to reflect upon their scale choices and

open answers, providing perceptions into why a certain answer was given. This encouraged us to look into if this method can provide information otherwise difficult to access. In the following, we report on our experiences of employing the TAS in an actual user study. We introduce the TAS as an additional qualitative analysis that answers the why of questionnaires while conducting a user study.

2 Surveys in small scale research

As stated by [9], surveys are a frequently used method within HCI research. The self-reported data provides information about the users' perception of the system, their interaction with it [16], understanding of the different user groups, and what opinions they hold [4]. It enables researchers to gather information from a small set of users and infer to a larger group [10]. The advantages of surveys are that they are relatively easy to generate and analyze. The challenges lie mainly in research mistakes and participant misunderstandings of survey questions and scales [12]. A common criticism of surveys is that the questions are subject to misinterpretation and fail to measure what is intended [14]. To eliminate the misinterpretations it is suggested to do a few initial tests of participants' understanding of the questions using the TA protocol [13] and let subjects verbalize their thoughts as they attempt to answer the survey questions. The proposed survey response process involves testing the comprehension, retrieval of information, judgment or estimation, and selection of a response to the question [15].

3 The Think Aloud method in usability evaluation

TA in HCI is known as one of the protocols for usability testing [6]. The participant is asked to vocalize his or her thoughts, feelings, and opinions while interacting with a system [8]. The concurrent TA method is expected to provide information during the participant's interaction with a system to identify the areas or features that cause problems for the user [3].

We propose to use the TA protocol to capture the active scale decision (qualitative analysis) and at the same time measure the participants' attitude towards and willingness to use, in this case, the interactive persona system (quantitative analysis). In our experimental setting, we found that the combination can provide answers to both how many have a particular attitude and why they express the attitude, as well as the reasoning and context behind a specific quantitative answer the participant gives.

4 Data collection

Our data collection site is a large non-profit organization with employees of multiple nationalities. In total, there were 37 participants, of which ten (27%) were women. The average age of the participants was 32.9 years (SD = 6.9). The participants' experience of personas included 'Conceptual experience' (71%, n = 26), 'Some practical experience' (27%, n = 10), and 'Extensive experience' (3%, n = 1).

In an experimental setting, we wanted to test a persona analytics system. The participants were encouraged to speak during the completion of the questionnaire, thus providing richness to the participant survey responses. The survey questions (SQs) were applied in a questionnaire using a seven-point Likert scale with 1 = Strongly disagree, 7 = Strongly agree. After the survey were completed, participants were asked questions about the process and the survey responses retrospectively. All of this was recorded and later transcribed for analysis. The questionnaire resulted in initial research on the participant's understanding of the system and its usefulness.

For the purpose of this research and to analyze the TA transcripts, one researcher applied a coding process focusing on the questions reported above. NVivo, a software package for qualitative analysis, was used for the coding.

5 Findings

Overall, we identified nine possible benefits associated with the use of the TAS to complement the quantitative survey results. In the following sections, we present the benefits.

1. Correcting Misunderstandings

Participants in doubt about a question, can ask the researcher to explain the question. The ambiguity of a question becomes visible.
Individual perceptions of the Likert scale becomes visible

2. Forcing Argumentation and Affecting Scoring

Upon what the participants base their judgment, becomes clear as the contextual judgement is vocalized. The researchers, get an understanding of the reasoning behind a score.

3. Revealing Same Arguments for Different Scores

Different participants might have the same argument but provide different scores. The difficulty of relying on the value without the contextual information becomes clear.

4. Providing Information Beyond the Scale

Participants elaborate on the middle score, as weighing between the positive and the negative sides, but also something missing. The answers open for much more information than a neutral value.

5. Providing Contextual Information about Participants' Understanding of the System

In surveys, it can be difficult to access if the participants understand what it is they are evaluating. An example shows that the participant have a fairly accurate understanding of the persona method, what personas are, and what to expect from the persona method. Thus providing information about the understanding of the system.

6. Understanding Use Contexts

The information that the participants provide, makes the use context known to the researchers, thus enables them to improve the IT system. These insights adds richness to the findings and actions that occurred during the actual user study preceding the questionnaire.

It becomes clear that if the participant lacks something that the system should be able to provide for their daily tasks.

7. Identifying Missing Content for Design

The TAS includes issues concerning the interface and the content that is not provided in an unmoderated survey. This is valuable for the further development of the interactive system.

8. Exposing Interface Issues

The TAS method exposes problems in the interface design and how participants relate to and their expectations of an interface. In this case that parts of system are confusing.

9. Yielding Proposed Solutions

A critique often leads to suggestions for improvement and the participant suggests improvement.

In summary: The analysis of the TAS transcripts revealed different patterns that connects to the participants misunderstanding survey questions, revealing arguments for the scores they gave and at the same time forcing the participants to consider their motivation for scoring. It uncovered that the same argument can result in different scores, providing the researchers with information beyond the scores. Results also tested if the participants understood the persona analytics system, and provided information about a future use context and uncovered usability issues, as well as suggestion for solutions.

The positive effect of the TAS is the possibility to do both a qualitative analysis of the transcripts and a quantitative analysis of the quantitative scoring carried out by the participants. This provides information on the participants understanding of the system, information needs and use context, usability issues, and solutions to design. A quantitative analysis of the scores provides where the main problems are, and is relatively fast to extract.

The downside is that it requires a moderator and is time consuming.

6 CONCLUSION AND FUTURE WORK

The proposed method of combining surveys with the think aloud protocol – the think aloud survey – provides insights similar to usability testing and insights into survey questions and responses. The method provides in-depth insights into the survey questions and can answer the why that is typically difficult to elicit from quantitative research. For practitioners, this method can be a faster way to get both quantitative and qualitative data, and it can be a test bed for setting up large surveys. In this research, we did not anticipate the usefulness of the method in terms of enhanced system design. In the future, it will be necessary to set up studies that are designated to investigate the application of the think aloud survey.

7 REFERENCES

1. An, J. et al. 2018. Imaginary people representing real numbers: Generating personas from online social media data. *ACM Transactions on the Web*. 12, 4 (2018). DOI:<https://doi.org/10.1145/3265986>.
2. [Ericsson, K.A. and Simon, H.. 1984. *Protocol analysis: Verbal reports as data*. MIT Press,.
3. Frandsen-Thorlacius, O. et al. 2009. Non-universal usability? Proceedings of the 27th international conference on Human factors in computing systems - CHI 09 (New York, New York, USA, 2009), 41.
4. Goodman, E. et al. 2013. *Observing the User Experience: A Practitioner's Guide to User Research (Second Edition)*. IEEE Transactions on Professional Communication.
5. Haeger, H. et al. 2012. Cognitive Interviews to Improve Survey Instruments. Annual Forum of the Association for Institutional Research. (2012).
6. Hertzum, M. 2020. *Usability Testing: A Practitioner's Guide to Evaluating the User Experience*.
7. Johnstone, C.J. et al. 2006. Using the think aloud method (cognitive labs) to evaluate test design for students with disabilities and english language learners (NCEO Technical Report). National Center on Educational Outcomes. (2006).
8. Mack, R. and Nielsen, J. 1993. Usability inspection methods. *ACM SIGCHI Bulletin*. (1993).
9. Müller, H. et al. 2014. Designing unbiased surveys for HCI research. Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14 (New York, New York, USA, 2014), 1027–1028.
10. Müller, H. et al. 2014. Survey research in HCI. *Ways of Knowing in HCI*.
11. Nielsen, J. et al. 2002. Getting access to what goes on in people's heads? Proceedings of the second Nordic conference on Human- computer interaction - NordiCHI '02 (New York, New York, USA, 2002), 101.
12. Ozok, A.A. 2020. *Survey Design and Implementation in HCI. The Human-Computer Interaction Handbook*.
13. Padilla, J.-L. and Leighton, J.P. 2017. Cognitive Interviewing and Think Aloud Methods. *Understanding and Investigating Response Processes in Validation Research*. 211–228.
14. Schwarz, N. 2001. *The Psychology of Survey Response*, Roger Tourangeau, Lance J. Rips, and Kenneth Rasinski. New York: Cambridge University Press, 2000. 401pp, ISBN 0-521-57246-0 (cloth) and 0-521-57629-6 (paper). *International Journal of Public Opinion Research*. (2001). DOI:<https://doi.org/10.1093/ijpor/13.1.80>.
15. Tourangeau, R. 1984. *Cognitive Sciences and Survey Methods. Cognitive Aspects of Survey Methodology : Building a Bridge*
16. *Between Disciplines*.
17. Tullis, T. and Albert, B. 2013. *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics: Second Edition*.
18. Willis, G.B. and Artino, A.R. 2013. What Do Our Respondents Think We're Asking? Using Cognitive Interviewing to Improve Medical
19. *Education Surveys. Journal of Graduate Medical Education*. (2013). DOI:<https://doi.org/10.4300/jgme-d-13-00154.1>.