AIS SIGPRAG
Pre-ICIS Workshop 2015
"Practice-based Design
and Innovation of
Digital Artifacts"

Improving Online Surveys through Respondent Behavior Logging

Mohammad Hafijur Rahman [a] and Jonas Sjöström [a]

[a] Department of Informatics and Media, Uppsala University, Uppsala, Sweden

Abstract

This paper introduces the concept of Respondent Behavior Logging (RBL), consisting of static and dynamic models that conceptualize respondent behavior when filling in online questionnaires. It is argued that web-based survey design may benefit from logging as a technique for evaluation, since such data may prove useful during re-design of questionnaires. Although other aspects of online surveys have attracted considerable attention both in industry and in literature, how the Web may leverage new and innovative techniques to support survey design is still underexplored. The RBL concept and associated evaluation techniques are presented and evaluated through the use of a focus group. Preliminary results are reported in the paper, and we conclude by a discussion that relates RBL to other questionnaire evaluation techniques, and raises concerns regarding ethical aspects and potential observer effects of RBL.

Keywords: Questionnaire design, online surveys, evaluation, behavior logging

1 Introduction

Research in information systems, as well as in other disciplines, often includes data collection through surveys (Newsted, Huff, & Munro, 1998; Sivo, Saunders, Chang, & Jiang, 2006). As stated by Krosnick & Presser (2010, p. 263), "the heart of a survey is its questionnaire". The importance of well-designed questionnaires cannot be understated. Flaws in questionnaire design may lead to response errors that may negatively impact the entire study (Krosnick & Presser, 2010).

There are several methodological guidelines to support questionnaire design. These can be divided into three broad categories (Krosnick & Presser, 2010): (i) The design of questions and the options provided to answer the questions. Question design focuses on different types of open and closed questions, and how to appropriate suitable rating scales given the knowledge interest of the researcher. (ii) The structure of the questionnaire as a whole. Such guidelines address the order of questions, and various strategies to improve the interpretability of respondents' answers, e.g. through the use of vignette questions. (iii) Guidelines for the process of testing questionnaires formatively, as a means to prototype and improve the questionnaires before going live in a full-scale survey.

For online surveys, special design guidelines have been discussed (Lumsden & Morgan, 2005). In this paper, we argue that questionnaire evaluation has not yet exploited the potential of the online medium. We propose the novel concept of Respondent Behavior Logging (RBL) that allows us to scrutinize the respondents' pro-

cess of answering questions. Through analysis of respondent behavior data we may be able to scrutinize and improve questionnaire design. The idea is to log respondents' activities (e.g. answer a question, switch pages, save, submit, etc.) during filling in surveys online so that we can analyze and evaluate surveys based on their log data. Logged respondent behavior may also serve as an additional source of knowledge to interpret respondent answers.

Our work adds to the general knowledge base on questionnaire design and evaluation in the online context. This is especially important in order to support the evolution of commonly used instruments, such as the Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983) in psychology. Respondent behavior logging - if used repeatedly in online research where common instruments are applied - supports fine-tuning of instruments over time, based on large data sets of respondent behavior.

The paper proceeds as follows. First, we provide an overview of existing knowledge and artifacts in the domain of online surveys. Second, we give a conceptual view of RBL and show the overarching design process and the staged design process that have facilitated the continual development of RBL as of now. Furthermore, a dynamic model and a static model are presented to describe the mechanism of RBL. Third, we present an instantiation of the artifact, along with insights from the design process and preliminary evaluation results. We conclude by discussing our findings and the anticipated value of RBL as conceived at this stage.

2 Knowledge Base

In this section, we account for existing artifacts for online surveys, a pragmatic foundation for our design, and a brief summary of IS literature that relates to conceptualizations of and evaluation techniques for respondent behavior.

2.1 Existing artifacts for online surveys

There is no shortage of online tools and services to conduct online surveys. Several vendors (e.g. Google, SurveyGizmo and SurveyMonkey) provide features for designing questionnaires, collecting data and analyzing responses. Open source alternatives emerge, both in the form of software to power online surveys (e.g. LimeSurvey) and as open source schemas for describing and managing questionnaires (e.g. queXML and SUML). These emerging technologies and standards highlight various aspects of the complexities of conducting online surveys. In addition to questionnaire design, they support data collection and data analysis. As an example, SurveyGizmo provides the option to embed decision logic within questionnaires, and an application programming interface (API) for developers and social plug-ins for Facebook and Twitter. LimeSurvey supports multilingualism as well as 'skip logic' (e.g. showing question Y if and only if the user chose a specific optional answer for question X). As an additional example, there are typically features for data export/import to various formats, including CSV, PDF, SPSS, queXML and MS Excel. Third-party modules are available to, for example, connect LimeSurvey to Wordpress (SurveyPress) and Drupal (LimeSurvey Sync).

The existing tools and schemas include a variety of features to set up questionnaires, and to collect and analyze data. There is, however, little emphasis on the process of designing and evaluating questionnaires.

Although guidelines for online questionnaire design exist (Lumsden & Morgan, 2005), there is a lack of knowledge on how to design IT artifacts that support evaluation and re-design of questionnaires. Our ambition to develop such knowledge is so far influenced by two branches of literature, as outlined below.

2.2 Pragmatism as a Foundation for Design

First, our design builds on the philosophy of pragmatism, which leads us to seek an understanding of the world as an ongoing process of action (Blumer, 1969). The RBL concept is an attempt to understand the respondent's situation when filling in a questionnaire. Our view is based on Dewey's (1938) conception of a situation as a contextual whole, consisting of objects and events: "What is designated by the word 'situation' is not a single object or event or set of events. For we never experience nor form judgments about objects and events in isolation, but only in connection with a contextual whole. This latter is what is called a situation." We do not claim that we can get a full understanding of a 'contextual whole' through RBL, but our aim is to provide a concept that can bring us more understanding of the respondent's interpretation of the questionnaire by facilitating inquiry into the events through which their answers to the questionnaire emerged.

The pragmatic starting point has implications for conceptual modeling (Ågerfalk, 2010), a topic central to the IS field (Wyssusek, 2006). A common perspective in the literature is that information systems are symbolic representations of reality (Wand & Wang, 1996). An alternative view, rooted in pragmatism, constitutes a critique to the representational view (Hirschheim, Klein, & Lyytinen, 1995; Holm, 1996; P. Ågerfalk, 2010). Drawing on the principles of pragmatism, language philosophers (Habermas, 1984; Searle, 1969) have argued that language (and by extension also information systems) does not just Get the (state of the) world but also brings about change to the world. From such a pragmatic viewpoint, IT artifacts are instruments that are used in performance of action. From the pragmatic viewpoint, IT artifacts are instruments used to perform action (Goldkuhl & Lyytinen, 1982; Sjöström, 2010) and mediators of actors' intentions. In the case of RBL, we want to create a formalized representation of respondent behavior. That representation, however, must account for the actions the respondents have taken in the process of providing answers to a questionnaire. These actions reveal to us when the respondent hesitated, changed their mind, got tired of answering questions, *etc*.

2.3 Previous IS research on user logging

A literature review was conducted in order to identify previous IS research on logging and conceptualization of user behavior, and techniques to evaluate user behavior. The following eight journals were selected as the sources for the present literature review: European Journal of Information Systems, Information Systems Journal, Information Systems Research, Journal of the AIS, Journal of Information Technology, Journal of Management of Information Systems, Journal of Strategic Information Systems, and Management Information Systems Quarterly.

The literature review revealed several studies that track respondent behaviors in numerous ways. A common denominator between the identified studies is that they phrase research questions aiming at identifying the effects that a certain behavior or perception on the part of the respondent has on other outcomes. Some of the retrieved research articles have very specific focuses in this regard. For example, Dewan and

Ramaprashad (2012) sought to identify relationships between social media use and music consumption; Ruth (2012) has investigated the effects that "conversation" in a question-and-answer site can have on user satisfaction; and Koch et al. (2012) have explored the effects that social media use in the workplace can have on employee outcomes. Perhaps the most salient of the retrieved articles is the one written by Sen et al. (2006), in which the researchers attempted to delineate the search strategies of online buyers. All of these studies had the purpose of developing conceptual maps of what qualities in the participants corresponded to what specific selected outcomes.

Web based surveys provide a convenient means for collecting data during surveys, and has been used in many present research surveys (Picoto, Bélanger, & Palma-dos-Reis, 2014). However, in their research on m-business value determination and usage factors, there is no mention of logging user activities to evaluate respondent behavior. The research used the traditional questionnaire pre-test methods and pilot testing; these were done to evaluate possible user response in order to make changes to the questionnaire before final administering.

According to the organizational behavior typology, individuals exhibit behavior by developing psychological bonds based on their own affects, needs and obligations. Every type of community commitment exerts a distinctive effect on every behavior; need based commitments can predict whether a certain user will read a given thread or not. Obligations based commitment can predict the moderating behavior of users of online applications (Bateman, Gray, & Butler, 2011). Studies of online based forums and communities have not come up with a concrete explanation as to their nature; some studies conclude that individuals in these (online) forums are motivated by their own self-interest. Using the organization psychology typology and psychological affiliation, it can be predicted how users will behave in online discussion forums. A users' likelihood to respond to a thread, for instance, can be predicted by evaluating the users' needs.

Bhattacherjee (2001b) evaluate how the expectations of users affect their engagement with Information Systems; and Bhattarcherjee (2001a) has also evaluated the factors that contribute to users of electronic commerce services achieving the selected outcome of continuance. A conceptually similar study was conducted by Oliver (1980) regarding the factors that result in users of a technology, product, and/or service experiencing satisfaction, as well as the effects that follow from a consumer feeling satisfied. In another study, Raban and Rabin (2009) used statistical tests to evaluate patterns in web-based interactions. These studies had the purpose of developing conceptual maps of what qualities in the participants corresponded to what specific selected outcomes. On the basis of the findings of the various studies, business and IT professional could implement processes that are meant to change the independent variables (respondents' qualities) and thereby change the dependent variables (selected outcomes such IT use or morale in the workplace).

The general insight that can be gleaned from this literature review is that there is potential value in carefully conceptualizing human behavior with respect to IT use. Structurally, the idea of using respondent behavior tracking in order to improve the design of a survey is not so different from identifying antecedents of other types of commitment to IT use. In a narrower sense, though, it seems clear that there is virtually no literature that has focused on the use of respondent behavior tracking to improve the design of surveys *per se*. All of the reviewed literature focuses on achieving a concrete outcome that would be valuable out in the "real" world; none of the literature

has focused on how the concept of respondent behavior logging could be used in order to improve the quality of surveys – thus, by implication, the quality of the research that the survey is part of. The main subject of the present research idea is thus a unique one that has not been previously addressed by the Information Systems community. The identified related research uses surveys in order to track respondents' behaviors; the presented idea seeks to take a step back and track respondent behavior regarding the very process of taking the surveys.

Literates from other mainstream disciplines have not yet been addressed. But it is well understood that, designs that involve human interaction and understanding of human behavior in online context must be well investigated not only within Information Systems (IS) but also into other areas of research specially HCI (Human Computer Interaction). There are numerous indications suggest that, studying HCI among other areas will benefit us tremendously. Therefore, we have already taken into consideration to investigate the area of HCI among other.

2.4 Conclusion drawn from the existing knowledge base

There is a trend in web development to use data about user behavior to enable new types of application. This trend is clearly related to a pragmatist interest in human action. In the process of giving answers, users may 'give off' signs about their behavior that may be utilized in various ways by designers (Ågerfalk & Sjöström, 2007). An example of this is the well-established use of 'clickstream' analysis on the World Wide Web, used (for instance) to tailor the display of ads on web pages (Martin, Wu, & Alsaid, 2003). Such tailoring is based on aggregated knowledge drawn from behavior of a large set of users in a community. The proposed RBL concept is analogical to other user behavior logging concepts. We thus conceive of RBL as an exaptation (Gregor & Hevner, 2013) of old questionnaire evaluation techniques ('prototyping') in combination with logging techniques, allowing us to design a new type of questionnaire evaluation technique in the online context. By measuring respondent behavior when filling in a questionnaire online, we mimic the established version of letting pilot respondents fill in paper-based questionnaires while observing them.

In essence, these two branches of literature (pragmatism and logging of web behavior), tells us to conceive of RBL as a concept to support keeping track of how respondents act while responding to a questionnaire. Further, we should be able to analyze respondent behavior for individual situations as well as aggregates. Such analysis should be useful to both (i) improve the questionnaire design and (ii) interpret the answers provided by respondents.

3 Research Approach

The design of RBL is a part of a larger design research program in which a fairly complex piece of software is built in the context of eHealth and eHealth research. The overarching project is a multi-disciplinary effort including researchers and practitioners from psychology, medicine, information systems, caring sciences, and economics. It aims a supporting people with potentially lethal somatic diseases to cope with their situation. In the next two sections we account for (i) the overarching multi-disciplinary design process and (ii) the staged design process for RBL, which was conducted as an integrated part of the overarching design process.

3.1 Overarching Design Process

The integration of IS research into the multi-disciplinary context promotes the relevance of our research. The collaboration with senior researchers as well as practitioners in other fields than IS has been conceived as imperative in understanding the problem domain (psychosocial care) and the clinical needs. The multi-disciplinary approach influenced the discussions in the group. The representatives from academic traditions (e.g. caring sciences vs. economics) approach research from different angles, and different ideals are put into play both with regard to health issues and research issues. This environment tends to emphasize issues of relevance and rigor, since there is a continual need within the group to provide an agreed-upon rationale for design decisions.

We categorize the research program as multidisciplinary, however, there are chances a transdisciplinary culture may also be found in our collaborative work. But it would be incorrect to solely describe the communication and collaboration efforts produced results only from blended knowledge from the experts of different disciplines. Rather, experts in their respective fields shared knowledge with each other to achieve their respective goals.

The design process was set up in accordance with agile values [4], characterized by sprint reviews approximately every two weeks. The review meetings had several recurring members representing different professions and academic disciplines. In addition, external specialists and patient groups were invited to explore the software, followed by workshops in which they provided feedback to the design team. In total, 100+ design workshops have been organized, engaging a great variety of stakeholders. With respect to rigor, researchers from the included disciplines have contributed to the design process. IS researchers contributed with knowledge from the IS field and its sibling disciplines (primarily interaction design and software engineering). The IS input is based on a pragmatic stream of IS research, focusing social interaction through instrumental use of technology (Sjöström, 2010; Ågerfalk, 2010). The idea to log respondent behavior in the survey context as born in the larger context of logging user behaviors for research purposes, and to ensure accountability in the process of dealing with sensitive data. The software was equipped with extensive logging functionality to enable retrospective analyses of user actions.

3.2 The Staged Design Process for RBL

The RBL work was conducted in three stages as outlined below. Essentially, the process is a design science research (Gregor & Hevner, 2013; Hevner, March, Park, & Ram, 2004) approach consisting of (i) drawing relevance from practice (i.e. the eHealth practice), (ii) performing design and evaluation and (iii) applying methods and drawing theoretical influences from the knowledge base. In doing so, we propose constructs and models that shape the static and dynamic models of RBL, and corresponding methods / techniques to evaluate online questionnaires. In addition, we instantiate these concepts into software. Through the software and focus group evaluations, we assess the RBL concepts and techniques as outlined below.

3.2.1 Stage I – Fundamental RBL design

Drawing mainly from a pragmatist philosophy and a general interest in extensive action logging, a first version of RBL was design (see sections 4.1 – 4.2). Basically,

the design aimed at logging respondent behavior in a manner that allowed a fairly detailed reconstruction of how respondents filled in questionnaires. The RBL concepts were implemented into the software, i.e. a proof-of-concept implementation. Data has been collected in the trials in three years using the original RBL implementations.

3.2.2 Stage II – Design of Visualization Techniques

As a means to make RBL data interpretable and meaningful for stakeholders in the design practice, a set of visualization techniques were crafted and implemented into the software. The knowledge base on questionnaire evaluation was factored in, primarily the idea of being able to assess time aspects, phrasing/quality of single questions, and structural issues with the questionnaire.

3.2.3 Stage III – Evaluation

On basis of collected data, a series of visualizations were rendered using the software. A focus group consisting of experts from the design practice was organized. The visualizations were presented to the focus group, and the participants were encouraged to discuss the utility of these representations to (i) identify design flaws and suggestions for improvements of the questionnaire and (ii) discuss if they would interpret the collected data differently given the way they made sense of the RBL visualization. Our evaluation of RBL is based on the qualitative data collected in the focus group session.

4 Respondent Behavior Logging

In this section, we present the RBL concept and related visualization techniques. First, we present a dynamic model to frame user actions related to filling in questionnaires. Second, we introduce a static model intended to keep records of these user actions. Third, we introduce a selected set of techniques to visualize respondent behavior, i.e. make it accessible for various stakeholders for interpretation.

4.1 Dynamic model of respondent behaviour

A logical first step in development is to identify what user actions to log. During the design process, we identified a set of user actions that could be traced at the server side through the HTTP requests client browsers make to the web server. The flowchart in Figure 1 shows the dynamic model of user behavior.

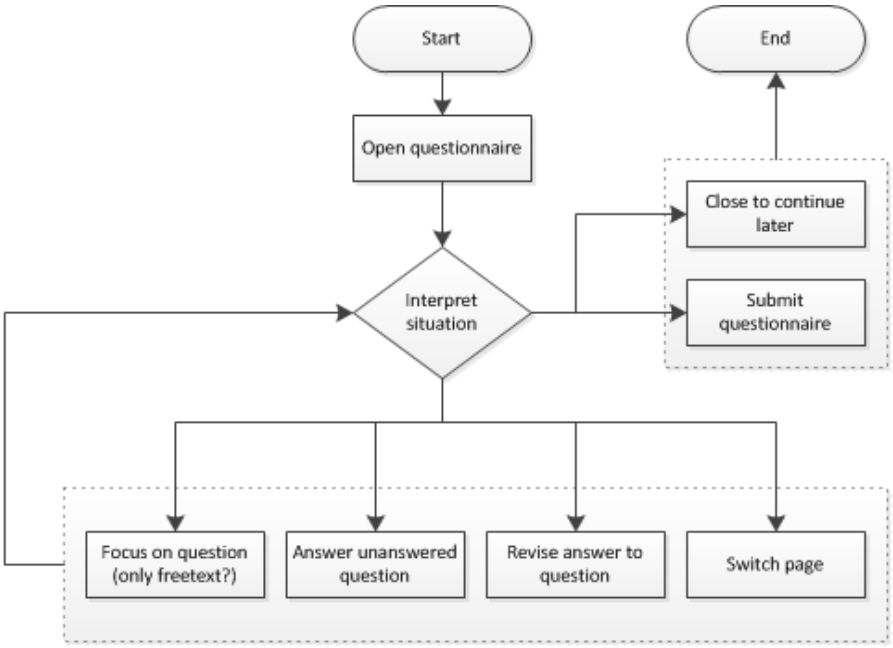


Figure 1: Dynamic Model of Respondent Behavior.

Seven action types are outlined in Figure 1. All are possible to track in a web context. The action types 'open questionnaire', 'close to continue later' and 'submit questionnaire' need no further explanation. The 'focus on question' action type means that there is some indication that a question is currently in the focus of the user, e.g. by hovering a question area with the mouse pointer or setting focus on a text field by clicking it. The 'answer unanswered question' occurs when a user answers a question that has not been previously answered, differentiating it from the 'revise answer to question' action type. It is also possible to track when users switch to another page in the questionnaire. The dynamic model, despite its simplicity, is an important concept to better understand what types of user actions we may trace in the context of online questionnaires. A main idea of tracking these actions is that a rich account of actions taken serves to indicate how the respondent interprets various situations when filling in a questionnaire.

4.2 Static model of Respondent Behavior

The entity/relationship diagram (Figure 2) shows a static model to keep track of user actions as shown in Figure 1. The purpose of the design is to allow traceability, i.e. that we are able to reconstruct the manner in which a respondent filled in a questionnaire. In addition, there is a need to keep track of the respondent's (user's) identity, to allow us to make queries related to a particular individual.

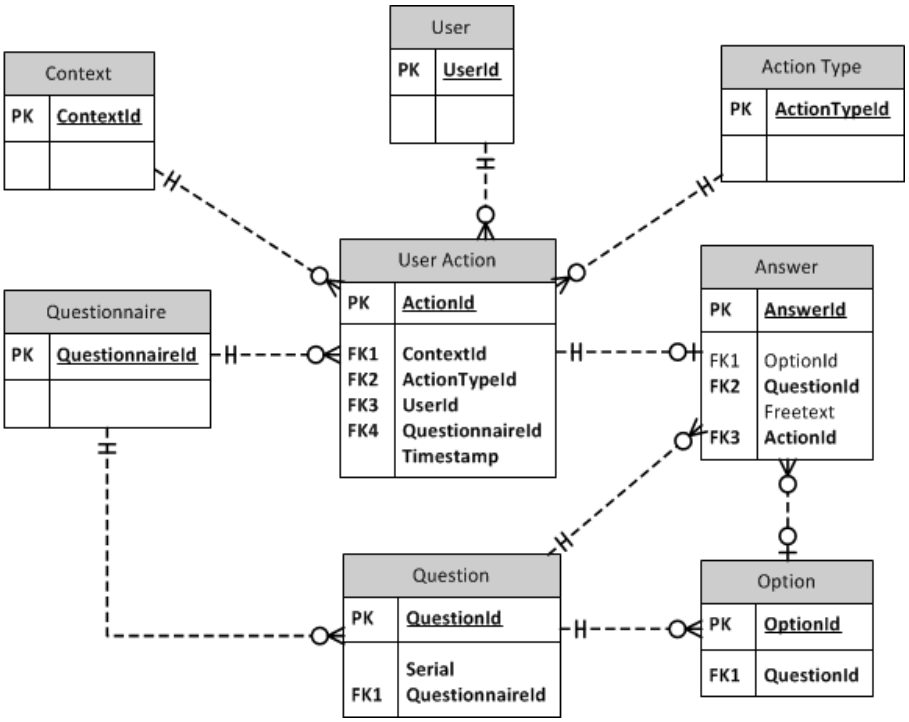


Figure 2: Static Model of Respondent Behavior

'User Action' defines the actual log record. It shows how a user, at a certain point in time, performs an action of some type in relation to a specific questionnaire. This action may include an answer to a question, which may be free text, but it may also point out an optional answer to a question. Keeping track of context is imperative, since one questionnaire may be used at different times. This is the case in many randomized controlled trials where the same questionnaire is typically used repeatedly to measure the same construct (e.g. depression or anxiety) at different times in the study. By including action type, we facilitate queries of user behavior based on grouping by the type of action performed.

4.3 Visualization Techniques for Respondent Behavior

Based on the way information is stored in the static RBL model, there are various ways to query into the RBL database to render reports and visualizations of RBL data. In this paper, we will present three such visualization techniques. The visualization techniques are primarily guided partly through choices that were felt most relevant for the researchers who design surveys and partly because of data unavailability that were not as comprehensive as to focus on other objectives. However, the techniques of visualization are still in initial stage but the idea is of course to create as many diverse interpretation of the log data as possible. More the techniques are exposed to researchers, more these will entail which visual presentations are more beneficial than the other. Therefore, the design process of creating meaningful visual techniques will continue to maximize RBL's objectives. In this section, we only briefly describe three techniques and their intended purpose (Table 1).

Table headings should be placed above tables. This is to allow for tables that span more than one page. Table headings should be Times New Roman 10 points bold and table text Times New Roman 10 points but not bold (or equivalent). Both headings and text in tables should have 3 points space before and after, as shown in Table 1.

Table 1: RBL visualization Techniques.

Visualization Technique	**Description**	**Purpose**
Activity Chart	The activity chart shows the number of activities (answer, delete, change answer) per question in a questionnaire.	More than 1 activity indicates that the respondent shows some uncertainty how to answer or what to answer.
Time Chart	The time chart shows how much time taken to answer each question in a questionnaire.	If a question takes a long time to answer, it could indicate some design flaw in the questionnaire (e.g. the phrasing of the question).
Answer Matrix	The answer matrix shows how a respondent moves between questions in a questionnaire.	A lot of movement back and forth between questions may signal structural problems with the questionnaire.

We will then return to these techniques in the evaluation section and discuss how they were interpreted by focus group participants.

5 Evaluation

The focus group I have interviewed, consists of six participants. They are vital stakeholders who have firsthand interest in this research project. The reasons I have included these participants as of my first focus group are, they are directly involved in designing online studies, they have been using online surveys from the beginning when this research project (U-CARE) came into place, and they represents three ongoing online studies at that moment. The purpose of this setting is to get as many diverse perspectives on the application of the RBL tool as possible. Through this focus group process, the effectiveness or lack thereof of the RBL tool itself should come into focus, on the basis of whether the tool helps to produce meaningful insights that would have otherwise been unavailable to the researchers and scholars of the focus group.

The focus group has been asked to reflect on three basic scenarios. First, covers the scenario when respondents' answers were collectively shown in relation to each questionnaire. Second, focuses on individual responses in a questionnaire in relation to other respondents. Third, reveals the comparisons of a same questionnaire used in different times. The charts and the cases were selected based on those overarching scenarios. The extreme cases were taken out due to invalid or incomplete characteristics of user actions. For example, a user closes the browser in the middle of answering a survey, or the session of a browser dies due to longer response time.

In organizational research visual aids are powerful methods of articulation. Visual aids perform, they are not representations of realities or practices (Davison et al., 2012). Visual media compels performance, i.e. enforces a two way process of understanding and acting at the same time. Using visual media in qualitative research is a good way of giving the field alternative ways to understand and talk about social existence. Visual aids produce a more dynamic configuration of the world around us that leave a lasting effect on the observer. It helps the users to avoid preconceived ideas while at the same time creating a possibility to observe the object of study in a different light, unlike in a verbal or written presentation (Nulty, 2008).

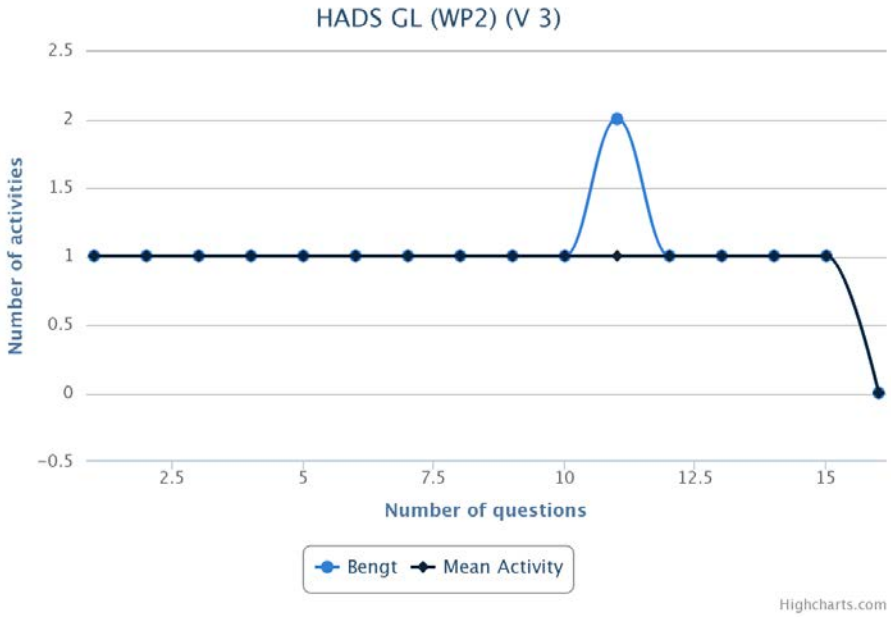


Figure 3: Activity Chart

Figure 3 shows the number of activities of the user Bengt (anonymous) took to answer a questionnaire (HADS – Has Anxiety or Depression Scale). There are 15 questions in HADS. The black line shows the mean number of activities to answer this questionnaire (for all users in the context). This is one of the examples to showcase that the ability to reinterpret user activities reveals knowledge about how a questionnaire has been generally used. In contrast to this case, there was another user who took quite a different path in answering the same questionnaire. So the question is, how we make sense of it? The user who took a different path, does that entail that, the user was very much enlightened/disturbed by the questions? Or, the user had difficulties in understanding some of the questions in the questionnaire.

One interpretation from one of the focus group members was, the HADS questionnaire is generally used for screening participants into different patient groups. So it is expected that, often we will discover some users who will take different approach to this questionnaire. Basically it suggests that, the users who often take unusual path the ones who need treatment.

Then a combination of activity charts was shown in which the same questionnaire was plot but used in two different occasions. However, the results were very much different in terms of users attending first time HADS compare to second time. The designer of that questionnaire said *"I am not sure this brings any ambiguity because users when attending HADS for the second time, they already have learned about the questions,.. but what should be interesting is that, the users who consistently had same results from both HADS"*. That means, the user who scored the same, is important to study in this case since it raises concerns that the user may not be improving or reacting properly to treatment.

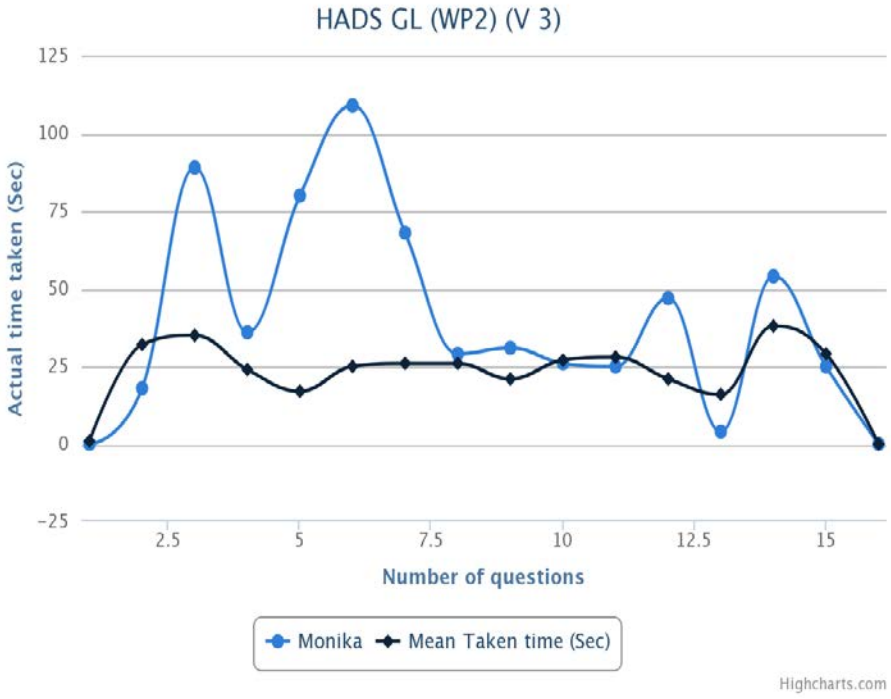
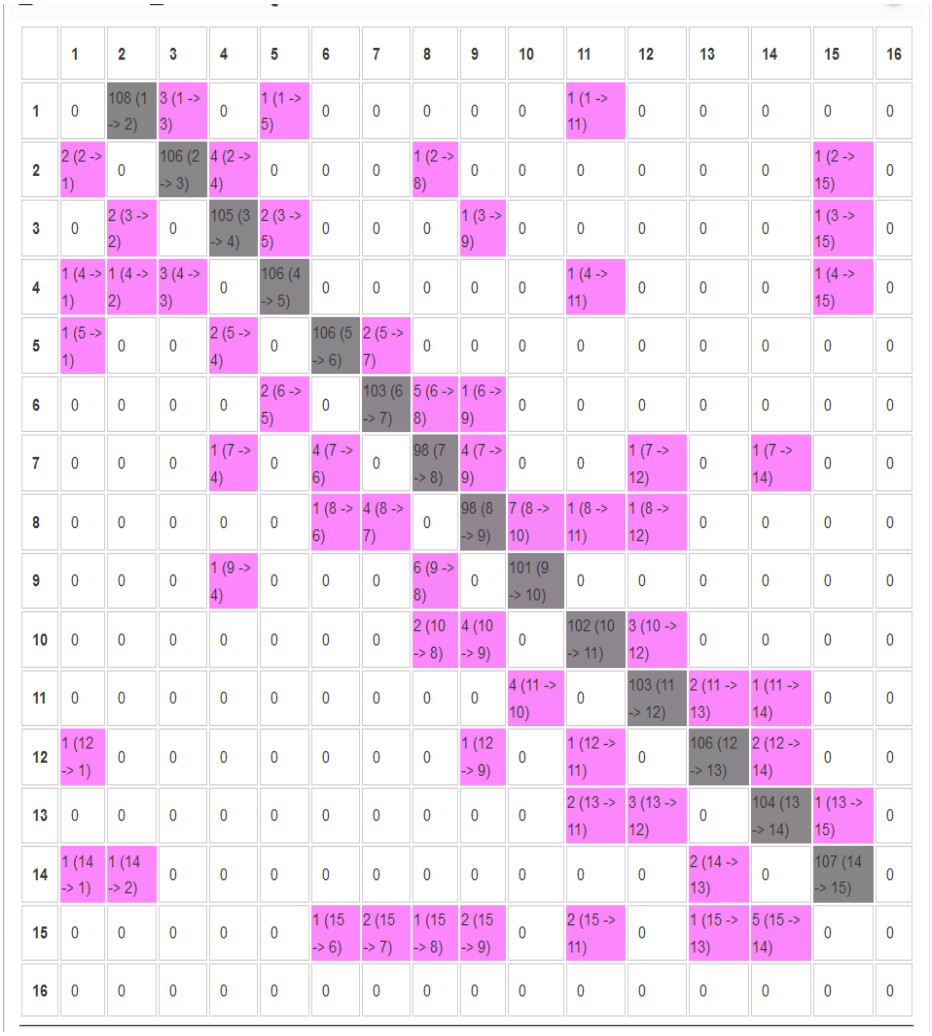


Figure 4: Time Chart

Figure 4 is an example of how much time a participant consumed to fill in HADS. The black line represents the mean time for participants, the blue line represents the Monika's time. It seems that, Monika took some time than normal around the questions 3 to 7 and the rest of the questions were filled in following the same time which most of the participants actually took. There is nothing more to prevail from Monika except the fact that she may have been very thoughtful around the first 7 questions. The HADS questionnaire usually consist of first seven questions related to measuring depression and the rest of the questions related to measuring anxiety. One specific comment from one of the members of the focus group was, *"this chart can be very useful if we combine the same chart time to time of a specific patient in order to compare a patient's situation when the same patient takes HADS later in their treatment process"*.

Figure 5 is a matrix of answers that shows how many times users have 'jumped' from one question to another given the questions are shown sequentially and at once.

Both axis represent the question numbers as they appear to the users. The gray boxes show the most common path or sequence the users have answered the questions.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	108 (1 -> 2)	3 (1 -> 3)	0	1 (1 -> 5)	0	0	0	0	0	1 (1 -> 11)	0	0	0	0	0
2	2 (2 -> 1)	0	106 (2 -> 3)	4 (2 -> 4)	0	0	0	1 (2 -> 8)	0	0	0	0	0	0	1 (2 -> 15)	0
3	0	2 (3 -> 2)	0	105 (3 -> 4)	2 (3 -> 5)	0	0	0	1 (3 -> 9)	0	0	0	0	0	1 (3 -> 15)	0
4	1 (4 -> 1)	1 (4 -> 2)	3 (4 -> 3)	0	106 (4 -> 5)	0	0	0	0	0	1 (4 -> 11)	0	0	0	1 (4 -> 15)	0
5	1 (5 -> 1)	0	0	2 (5 -> 4)	0	106 (5 -> 6)	2 (5 -> 7)	0	0	0	0	0	0	0	0	0
6	0	0	0	0	2 (6 -> 5)	0	103 (6 -> 7)	5 (6 -> 8)	1 (6 -> 9)	0	0	0	0	0	0	0
7	0	0	0	1 (7 -> 4)	0	4 (7 -> 6)	0	98 (7 -> 8)	4 (7 -> 9)	0	0	1 (7 -> 12)	0	1 (7 -> 14)	0	0
8	0	0	0	0	0	1 (8 -> 6)	4 (8 -> 7)	0	98 (8 -> 9)	7 (8 -> 10)	1 (8 -> 11)	1 (8 -> 12)	0	0	0	0
9	0	0	0	1 (9 -> 4)	0	0	0	6 (9 -> 8)	0	101 (9 -> 10)	0	0	0	0	0	0
10	0	0	0	0	0	0	0	2 (10 -> 8)	4 (10 -> 9)	0	102 (10 -> 11)	3 (10 -> 12)	0	0	0	0
11	0	0	0	0	0	0	0	0	0	4 (11 -> 10)	0	103 (11 -> 12)	2 (11 -> 13)	1 (11 -> 14)	0	0
12	1 (12 -> 1)	0	0	0	0	0	0	0	1 (12 -> 9)	0	1 (12 -> 11)	0	106 (12 -> 13)	2 (12 -> 14)	0	0
13	0	0	0	0	0	0	0	0	0	0	2 (13 -> 11)	3 (13 -> 12)	0	104 (13 -> 14)	1 (13 -> 15)	0
14	1 (14 -> 1)	1 (14 -> 2)	0	0	0	0	0	0	0	0	0	0	2 (14 -> 13)	0	107 (14 -> 15)	0
15	0	0	0	0	0	1 (15 -> 6)	2 (15 -> 7)	1 (15 -> 8)	2 (15 -> 9)	0	2 (15 -> 11)	0	1 (15 -> 13)	5 (15 -> 14)	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure: Answer Matrix

This picture was a surprise for the focus group, because it reveals some unanticipated jumps between questions by the users. For example, the picture shows that a user answers question number #14 after answering question number #2. Similarly there are other occurrences that are hard to explain. There is no problem if a user gives answer to second question and then jumps to the last question but the problem arises when so many users show the same tendency. The discussion in the focus group led to discover a bug in the questionnaire. The questionnaire designer then immediately recognized the buggy behavior of that questionnaire. She said, *"We were in the initial stage of designing a study and we wanted to use HADS for the screening purpose. We have mistakenly put branching condition on question number #2 instead of question number #14. But as soon as we found the problem we edited the questionnaire and repub-*

lished". This example was an indication of an issue that the structure of the questionnaire was misconfigured.

6 Concluding Discussion

We have designed the RBL concept on the basis of (i) existing artifacts, (ii) literature and (iii) a design process in the context of eHealth and eHealth research. We provided a proof-of-concept by implementing the RBL concept into software. The software, being used in live eHealth trials, is used to collect a large amount of data for respondent behavior while filling in questionnaires. So far, we have only conduct one focus group in which RBL visualizations were presented to stakeholders in the eHealth research practice. The results signal a potential value of RBL as a means to identify questionnaire design issues, but also as a means to make sense of data collected in the trials.

We do not consider RBL to be a substitute for other questionnaire design techniques, rather a complement that opens up novel possibilities for evaluation. While the implications of RBL at this stage are speculative, we find that it has good potential to become an important strategy for questionnaire evaluation and re-design. It is a new use of technology that may be integrated into the process of large-scale data collection online, supporting both the interpretation of collected data and the refinement of questionnaires.

As long as the RBL tool is activated, questionnaire evaluation occurs automatically while the user is taking the questionnaire. If RBL is 'activated', virtually no further work is needed on the part of the researchers to collect and monitor data for questionnaire evaluation. RBL thus has the potential advantage of being extremely cost-effective in comparison with other questionnaire evaluation techniques. Other methods typically require additional effort in order to carry out evaluation.

RBL potentially minimizes bias in the questionnaire evaluation process. In general, unwanted bias diminishes the quality of any research process. In other questionnaire evaluation methods such as cognitive interviews, there may be a gap between what the user's self-reported survey-taking behavior and their actual survey-taking behavior. Even if the researchers were to conduct independent quantitative analyses, there would still be a risk of selection bias regarding what data he chooses to focus on. RBL eliminates such risk by providing a one to one correlation between the data produced and the actual behaviors of the user. The value of this cannot be overstated.

A potential problem with RBL is that it adds an additional layer of logging to online research. Logging of respondent behavior, especially in the eHealth context, may be considered an intrusion of privacy. We are aware of two implications, namely (i) the use of RBL would require consent when applied in eHealth trials, and (ii) the fact that people are aware that their behavior while filling in the questionnaire is being logged may lead to a changed behavior, i.e. a type of observer effect occurring due to RBL. There is a need for further research on how to address the ethical aspects and the potential observer effects related to RBL.

Finally, we acknowledge the need to extend the literature review to include a wider range of sources, including but not limited to publication outlets for human-computer interaction, computer science, software engineering and research methodologies.

References

Bateman, P. J., Gray, P. H., & Butler, B. S. (2011). Research note-the impact of community commitment on participation in online communities. *Information Systems Research*, *22*(4), 841–854.

Bhattacherjee, A. (2001a). An empirical analysis of the antecedents of electronic commerce service continuance. *Decision Support Systems*, *32*(2), 201–214.

Bhattacherjee, A. (2001b). Understanding information systems continuance: an expectation-confirmation model. *MIS Quarterly*, 351–370.

Blumer, H. (1969). *Symbolic interactionism: Perspective and method*. University of California Press.

Davison, J., McLean, C., Warren, S., Bramming, P., Gorm Hansen, B., Bojesen, A., & Gylling Olesen, K. (2012). (Im) perfect pictures: snaplogs in performativity research. *Qualitative Research in Organizations and Management: An International Journal*, *7*(1), 54–71.

Dewan, S., & Ramaprasad, J. (2012). Research note-music blogging, online sampling, and the long tail. *Information Systems Research*, *23*(3-part-2), 1056–1067.

Dewey, J. (1938). *Logic: The theory of inquiry*. Henry Holt and Company, New York.

Goldkuhl, G., & Lyytinen, K. (1982). A language action view of information systems. In G. Ross (Ed.), *Proceedings of the 3rd international conference on information systems*.

Gregor, S., & Hevner, A. R. (2013). Positioning and Presenting Design Science Research for Maximum Impact. *Mis Quarterly*, *37*(2), 337–355.

Habermas, J. (1984). *The Theory of Communicative Action*. Cambridge: Polity.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in Information Systems research. *Mis Quarterly*, *28*(1), 75–105.

Hirschheim, R., Klein, H. K., & Lyytinen, K. (1995). *Information systems development and data modeling: conceptual and philosophical foundations* (Vol. 9). Cambridge University Press.

Holm, P. (1996). *On the design and usage of information technology and the structuring of communication and work*. Stockholm University Royal Institute of Technology.

Koch, H., Gonzalez, E., & Leidner, D. (2012). Bridging the work/social divide: the emotional response to organizational social networking sites. *European Journal of Information Systems*, *21*(6), 699–717.

Krosnick, J. A., & Presser, S. (2010). Question and Questionnaire Design. In P. V Marsden & J. D. Wright (Eds.), *Handbook of Survey Research* (2nd ed., pp. 263–313). Emerald Group Publishing Limited.

Lumsden, J., & Morgan, W. (2005). *Online-Questionnaire Design : Establishing Guidelines and Evaluating Existing Support. Proceedings of the 16th Annual International Conference of the Information Resources Management Organization (IRMA'2005)* (Vol. 31).

Martin, D., Wu, H., & Alsaid, A. (2003). Hidden surveillance by Web sites: Web bugs in contemporary use. *Communications of the ACM*, *46*(12), 264.

Newsted, P. R., Huff, S. L., & Munro, M. C. (1998). Survey instruments in information systems. *MIS Quarterly*, *22*(4), 553.

Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: what can be done? *Assessment & Evaluation in Higher Education*, *33*(3), 301–314.

Picoto, W. N., Bélanger, F., & Palma-dos-Reis, A. (2014). An organizational perspective on m-business: usage factors and value determination†. *European Journal of Information Systems*, *23*(5), 571–592.

Raban, D. R., & Rabin, E. (2009). Statistical inference from power law distributed web-based social interactions. *Internet Research*, *19*(3), 266–278.

Ruth, R. D. (2012). Conversation as a source of satisfaction and continuance in a question-and-answer site. *European Journal of Information Systems*, *21*(4), 427–437.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge Univ Press, London.

Sen, R., King, R. C., & Shaw, M. J. (2006). Buyers' choice of online search strategy and its managerial implications. *Journal of Management Information Systems*, *23*(1), 211–238.

Sivo, S. A., Saunders, C., Chang, Q., & Jiang, J. J. (2006). How low should you go? Low response rates and the validity of inference in IS questionnaire research. *Journal of the Association for Information Systems*, *7*(6), 17.

Sjöström, J. (2010). *Designing Information Systems*. Uppsala University.

Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, *39*(11), 86–95.

Wyssusek, B. (2006). On ontological foundations of conceptual modelling. *Scandinavian Journal of Information Systems*, *18*(1), 9.

Zigmond, A. S., & Snaith, R. P. (1983). *The Hospital Anxiety and Depression Scale. Acta Psychiatrica Scandinavica*, 67, 361–370.

Ågerfalk, P. (2010). Getting pragmatic. *European Journal of Information Systems*, *19*, 251–256. Retrieved from http://www.palgrave-journals.com/ejis/journal/v19/n3/abs/ejis201022a.html

Ågerfalk, P. J., & Sjöström, J. (2007). Sowing the Seeds of Self: A Socio-Pragmatic Penetration of the Web Artefact. In *Proceedings of the 2nd International Pragmatic Web Conference*. Tilburg, The Netherlands.