Association for Information Systems

AIS Electronic Library (AISeL)

Research Papers | ECIS 2018 Proceedings

11-28-2018

THE EFFECT OF SOCIAL REPUTATION ON RETENTION: DESIGNING A SOCIAL REAL-TIME DELPHI PLATFORM

Simon Kloker
Karlsruhe Institute of Technology (KIT), simon.kloker@kit.edu

Tim Straub
FZI Forschungszentrum Informatik, straub@fzi.de

Stefan Morana
Institute of Information Systems and Marketing (IISM), stefan.morana@kit.edu

Christof Weinhardt
Karlsruhe Institute of Technology, weinhardt@kit.edu

THE EFFECT OF SOCIAL REPUTATION ON RETENTION: DESIGNING A SOCIAL REAL-TIME DELPHI PLATFORM

Research paper

Kloker, Simon, Karlsruhe Institute of Technology, DE, simon.kloker@kit.edu

Straub, Tim, FZI Forschungszentrum Informatik, DE, straub@fzi.de

Morana, Stefan, Karlsruhe Institute of Technology, DE, stefan.morana@kit.edu

Weinhardt, Christof, Karlsruhe Institute of Technology, DE, weinhardt@kit.edu

Abstract

Forecasting with high uncertainty and long-time horizons still challenges researchers and practitioners. A widely adopted method in knowledge sharing and forecasting based on experts is the Delphi method and its offspring, the Real-Time Delphi. While the traditional Delphi method already is intensely investigated, the Real-Time Delphi is still evolving, and no dominant design has been found yet. A problem arising in both variants of the Delphi method, are high drop-out rates between rounds. This paper applies a design science research approach to motivate the need for social design elements from literature and derives design principles for Real-Time Delphi platform. Based on the design, we implement and evaluate a prototype in an online experiment as well as an IT artifact in a field study. We find significant supporting evidence, that (the promise of) positive social reputation increases commitment, and therefore subsequent platform engagement of our Real-Time Delphi survey. Our findings, therefore, contribute valuable design knowledge for Real-Time Delphi platforms. Moreover, we provide advice on how to raise retention in knowledge sharing systems.

Keywords: Real-Time Delphi, Delphi Method, Knowledge Sharing, Retention, Crowd-based Forecasting, Decision Support Systems, Knowledge Management, Online Social Interaction, Social Reputation, Drop-Out Rates, Design.

1 Introduction

Reliable forecasting and assessment of future developments has always been a key success factor of governments, companies, and organizations (Durand, 2003). Especially in the long-view, researchers and decision makers are confronted with high uncertainty, new technologies, and new ideas frequently. In these situations analytical techniques and basic scenario methods often cannot be used, due to a lack of specific information or technical and historical data (Linstone & Turoff, 2002). Employees or members within organizations often carry insights and have a gut-feeling about such topics that is beyond mere historical and technical data (Styhre, 2002). Modern knowledge management approaches in companies established systems to involve employees on all hierarchical levels in the knowledge generation and management process. Such are, e.g., enterprise social media to search for help regarding certain issues (Martensen et al., 2016) or prediction markets to forecast the success of ideas or project milestones (Remidez Jr & Joslin, 2007). In areas where the strategic decisions in organizations addresses expertise from different fields and/or where conflicting goals and values have to be considered, the Delphi method offers experts a structured communication method free of “inter-personal” and hierarchical effects and social pressure (Linstone & Turoff, 1975; Strauss & Zeigler, 1975). The round-based character of such studies helps the panel of experts to revise their opinion and estimation several times and, therefore, find a consensus on questions and strategic decisions with a far horizon. With the combination of focused feedback on relevant information and statistical values on the one hand and the flexible methodology with valid results on the other hand, the Delphi method stands for a very effective technique (Landeta, 2006). In knowledge management they are also applied to, e.g., develop measurement scales (Boulkedid et al., 2011) or estimate trends and developments (Gnatzy et al., 2011). However, it also requires, as many other systems for knowledge management, high commitment and retention of the participants to the survey, as there may be long time-gaps between the rounds. This leads to the problem of drop-outs (the lack of retention) of participants during the process, which became more severe with the growth of the panels, made possible by the internet technology (Kloker et al., 2016; Lampe et al., 2010; Reid, 1988). Therefore, to further develop Real-Time Delphi for large panels, raising retention is a key challenge. We define retention as the recurring participation of participants/experts in multi-round or asynchronous (Real-Time) Delphi studies.

In a traditional Delphi study, experts receive a questionnaire in paper-and-pencil form from a monitoring team. During a defined time horizon, they give their responses while assessing the specific scenario alternatives by scales of measurement and optionally provide reasons for their estimation. Subsequently, the questionnaires are sent back, and the individual arguments and judgments are summarized by the monitoring team. The group results serve as a basis for a new questionnaire, which is sent back to the respondent group. The experts then have the chance to examine the group feedback and, thereon, reevaluate their original answers. This process can be repeated several times resulting in the occurrence of “structured communication” between the participants. According to Dalkey et al. (1969); Rowe et al. (1991) the Delphi method displays four key elements: 1. *anonymity* of the survey participants, 2. *controlled feedback*, 3. *statistical group response*, and 4. *iteration*. Despite the great popularity, the conventional Delphi method also shows some weaknesses, such as the large time frame needed to perform studies, the lack of Real-Time presentation of results, and the elaborate tasks of the facilitator. The first and second have been addressed by the introduction of the Real-Time Delphi (Gordon & Pease, 2006). Real-Time Delphi allows participants to access the answers and estimation of other participants not only in the upcoming round, but immediately after submitting (or even before) their own estimation, and, therefore, in “Real-Time”. Access to the study is usually also possible at any time. Further challenging factors of the traditional Delphi method and Real-Time Delphi are the availability of experts and the drop-out-rate/lack of retention (Mullen, 2003; Walker & Selfe, 1996), as the likelihood of drop-outs increases the longer the process is ongoing (Gnatzy et al., 2011). For both challenges there is still a lack of applicable solutions, see e.g. D. A. Rowe et al. (2015). After the development of Real-Time Delphi by Gordon & Pease, similar concepts have been evolved by other researchers. However, none of them has yet addressed retention of participants in their design (Kloker et al., 2016).

The work at hand conducts a literature review and gives an overview of existing implementations of the Delphi method with Real-Time character and points out important similarities and differences. Based on this literature review, we derived eight design principles for a Real-Time Delphi platform. Based upon, we propose and evaluate a ninth design principle in a Design Science Study, which enhances the current design by facilitating social interaction as suggested by Kloker et al. (2016). The approach of Kloker et al. suggests the introduction of labels and generated user names in the discussion to enable participants to build up social reputation and a sense of community while also ensuring anonymity. We argue that social reputation raises retention. Overall, our research addresses the research question: *How to design social elements in anonymous knowledge sharing and forecasting platforms, in order to stimulate retention in multi-round settings?* Our results show that labeling comments has a significant influence on retention.

The remainder of the paper is as follows: Section 2 gives an overview of literature on Real-Time Delphi and social interactions on online platforms to give a foundation to our hypothesis. Section 3 introduces our Design Science Research project. Section 4 formulates the design principles. Section 5 reports the methodology and results of the two-fold evaluation strategy with an online experiment and a field test. We conclude with a summary of our findings and its limitations, as well as a brief outlook.

2 Foundations and Related Work

2.1 Classic vs. Real-Time Delphi

The Real-Time Delphi technique is an advanced concept based on the conventional Delphi method, which is already covered by an extensive body of literature. The conventional Delphi method was first introduced by the RAND Corporation in the 1950-1960s. In 1975, Linstone & Turoff published a book titled "The Delphi Method" discussing the technique and its applications so far (Linstone & Turoff, 2002). Today it is considered to be the basic literature for defining the main characteristics and key elements of the conventional Delphi method. Though there is a "basic scheme" and key characteristics to the Delphi method, it often has to be adapted to real-world problems and context (Landeta, 2006). Nevertheless, Landeta (2006) summarizes "*[...] that the Delphi method continues to be used and is a valid instrument for forecasting and supported decision-making*" (Landeta, 2006, p. 478). Rowe et al. (1991) and Rowe & Wright (1999), however, criticizes that a real evaluation of the Delphi method is difficult and leads to a distorted estimation of the quality of the forecasts and conclusions. It is highlighted, that future research needs to elaborate on the mechanics of judgment change within the groups during the rounds and the underlying processes. Especially, as the Delphi method is referred to as an approach to reach consensus. It is also mentioned, that the forecasting accuracy increased with each round (Rowe & Wright, 1999). The named publications structure findings and standards in the conventional Delphi method and point out interesting strengths and weaknesses of the technique; which to some extent can be identified in Real-Time Delphi as well.

Research on Real-Time Delphi is based on many studies performed in the last decade. Several publications report the implementation of Real-Time Delphi platforms and present the basic procedures (Abadie et al., 2010; Gary & von der Gracht, 2015; Schuckmann et al., 2012). However, there is hardly any related work that carries out a systematic literature review to present different approaches of Real-Time Delphi or discuss similarities and differences between existing implementations. Gordon (2009) refers to other applications while describing his developed version of Real-Time Delphi. In Gordon et al. (2015), the same authors enlarge this view by presenting further techniques. This overview is useful to get an impression of the application areas for Real-Time Delphi, but does not provide any comprehensive comparison of the approaches. In 2011, Gnatzy et al. drew up a modified Real-Time Delphi technique based on the idea of Gordon & Pease (2006). After describing the developed methodology of Real-Time Delphi, Gnatzy et al. (2011) compared their approach to the one of Gordon & Pease. Since Gnatzy et al. (2011) wants to point out detailed improvements of features and process design, the comparison of the implementations is extensive and described thoroughly. The main focus of Gnatzy et al. is laid on the visual statistical group feedback and a higher level of expert guidance through the survey by a one-screen-one-question design.

To summarize, there is currently no standard definition of Real-Time Delphi. However, in order to perform a structured literature review and identify design principles (see Section 4), we need to define a minimum criteria set to decide whether an application is a Real-Time Delphi. Based on this literature review, we apply the following working definition: "A Real-Time Delphi is an online implementation of the Delphi method, where users can interact with the platform online and at any given time."

One problem with the traditional Delphi method as well as the Real-Time Delphi is the lack of retention of participants over multiple rounds. According to Mullen (2003); Walker & Selfe (1996) the response rate in Delphi studies needs to be at least 70%. Reid (1988) notes that the panel size has a strong influence on the drop-out rate. Large panels tend to lead to less retention of participants than small panels with less than 20 members. In Real-Time Delphi we expect the problem to be much larger, as individual's contribution and involvement becomes smaller. This effect is known as "Social Loafing" and is present in online communities as well (Lampe et al., 2010). Okoli & Pawlowski (2004) argue that the researcher has the possibility to contact the drop-outs and ask them to participate, but this can be, depending on the budget, related to a disproportionate effort (Ishikawa et al., 1993). However, the technological concept of Real-Time Delphi and its asynchronous character would allow distinctly larger panels. To draw upon this potential, it is necessary to bind users strongly to the platform and the survey. This can be accomplished by enabling participants to experience an online presence. The experience to feel "present" in the online community is a prerequisite to attribute actions and reactions on the platform to oneself and build up "reputation". Bolger & Wright (2011) found that in "traditional" Delphi studies the promise of gaining social reputation raises motivation to commit to the study and raises retention. Therefore, the focus of current work is to raise the retention of the participants over multiple rounds, as a key success factor of Real-Time Delphi.

"Collaboration begins with interaction" (Murphy, 2004, p. 422). The experience of presence in online settings makes geographically separated persons behave as a group, which enriches interaction and the sense of community. *Experience of online presence* is hereby regarded as the degree a participant is feeling personally involved in an online task, which is crucial for forming collaborative communities (Lampe et al., 2010). In a collaborative community members do not only share perspectives, but are starting to challenge other opinions, reshape their own, and restructure their thinking. This process leads finally to a shared meaning – which is also characteristic to the Delphi method. However, online presence in online collaboration has the ability to start additional processes: New perspectives and meanings as well as shared goals can evolve (Roschelle & Teasley, 1995). Second leads to the production of shared artifacts and the intention to "add value" (Kaye, 1992). It is not yet discussed, if these processes lead ultimately to better results in every case, but intuitively, it may improve the result in some dimension. Leveraging this improvement for Real-Time Delphi has not yet happened and Linstone & Turoff (2011, p. 1718) says *"[...] the future of Delphi will be in collaborative organizational and community planning systems that are continuous, dispersed, and asynchronous."*

Online presence and the sense of community allow building up "social reputation". In order to do this, "labeling" is a widely applied approach in forums. Lampel & Bhalla (2007) emphasizes the "reputation (status) seeking" behavior of users in online communities. Tagging or labeling (tagging with a fixed set of labels) content in "social question answering" (e.g. Yahoo! Answers or Live Q&A), can open opportunities for richer user interaction (Rodrigues et al., 2008). According to Ames & Naaman (2007) there are mainly two reasons to tag social content[1]: i) Providing one's opinion on something (social interaction) and ii) help others/oneself to find something (self-organization). Additionally, Rainie (2007) puts that tagging allows groups to form around points of view and similarities of interest. If persons use the same tags, they may get the impression that they probably share some deep commonalities. Tagging or labeling can therefore contribute to Real-Time Delphi in multiple ways: First, it enables users to express their opinion about arguments and gain reputation. Second, it enables users to express "common sense". Both leads to higher experience of online presence and therefore raise commitment to the platform. Third, tagging and labeling are a strong instrument of (self-) organizing content. Especially for larger panels, online discussion can quickly become confusing, if there are no means to structure and distinguish the important from the unimportant or the interesting from the uninteresting. Lots of large online platforms as Twitter, Facebook, or GitHub use tagging or labeling as a mean to allow structuring and organizing the content. Turoff et al. (2004) already used labels to organize content in a study which he attested a "Delphi-structure". He did, however, not enable the users to label arguments or the inputs of other users, so we do not find the social character here. As his panel consisted of students of a lecture and participation was mandatory, we can also derive no assumption on retention. In addition, his implementation did not fit the anonymity criteria, as names of the authors of arguments were visible. Usually the Delphi method as well as Real-Time Delphi build on absolute anonymity (or quasi-anonymity as argued in Kochtanek & Hein (1999)). Gordon (2009) states the concern about spurious factors, such as (prior) reputation, status, or other social behavior that intrude in face-to-face interactions among experts. These concerns lead once to the feature of anonymity in the beginning of the Delphi method. Anonymity is a key feature of the Delphi method and it

[1] In case of Ames & Naaman (2007) photos.

was adopted in Real-Time Delphi. Therefore, key challenge of current research is to raise retention by increased online presence and the promise of social reputation with labels and, at the same time, not to allow tracking of single users long-term and, therefore, harm the anonymity criteria. Kloker et al. (2016) suggested generated user names to support the promise of social reputation and retention. They may induce the feeling of addressability (and therefore social presence) and that individual participants may be traced and, therefore, can collect reputation.

3 The Design Science Research Project Setting

Aiming at investigating and solving the challenge of low retention by the introduction of social interaction in Real-Time Delphi, we apply the Design Science Research (DSR) (Hevner et al., 2004) approach as described by Kuechler & Vaishnavi (2008). In our case we perceive the DSR approach as promising, because it helps us to understand the underlying design and at the same time evaluate an appropriate information system based on the design. To the best of our knowledge, there are no design principles for the class of Real-Time Delphi published yet. Therefore, our research will contribute to the existing body of knowledge within the IS community. Our DSR project is conducted in two consecutive design cycles (see Fig. 1).

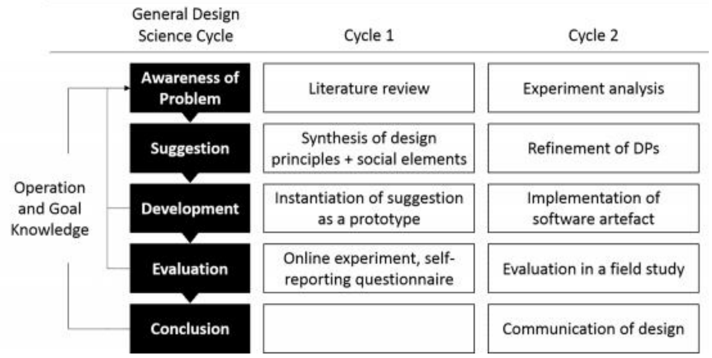

Figure 1: Two cycle DSR project (adopted from Kuechler & Vaishnavi (2008))

To create a complete and sound awareness of the problem, we started the first design cycle with an extensive literature review on prior findings in related fields, former applications of the Real-Time Delphi, and the relevance of the problem with retention in past applications. Based on this review, we derive design principles for the class of Real-Time Delphi platforms. The literature search took all publications into account, in which a Real-Time Delphi approach (or a comparable implementation), based on the working definition of a Real-Time Delphi from Subsection 2.1, was utilized (implemented) and described (or evaluated). Thereupon, we conducted a backward and forward search. Many of the relevant literature was found in the journal "Technological Forecasting & Social Change", in which publishing articles about the Delphi method and Real-Time Delphi had become common. By scanning available descriptions and discussions of different Real-Time Delphi approaches, we extracted features, organized them according to the key aspects of the Delphi method, and compared them according to these features. Then we differentiated between features that achieved "quasi-standard"-status and features with varying implementations. A preliminary online experiment instantiated a prototype of a Real-Time Delphi platform based on these findings in two treatments: (1) A control-group faced a Real-Time Delphi platform implementing the "quasi-standard" (RTD). (2) A treatment-group faced a Real-Time Delphi platform in the "quasi-standard", plus featuring social elements (sRTD).

In the second design cycle, we reconsidered the design based on the feedback and results of the first design cycle and instantiated the prototype into a full IT artifact within and as a component of a platform named "FAZ.NET-Orakel". The FAZ.NET-Orakel[2] is a gamified crowd-based forecasting platform utilizing a prediction market that is publicly available on the online magazine of our industry partner[3]. The social Real-Time Delphi tool shall complement the prediction market to additionally enhance the forecast with qualitative feedback and additional background information on the forecasting task as suggested by Kloker et al. (2017). This enables us to evaluate the artifact under real world conditions. The combination of our literature review, preliminary online experiment, and the field test together with our industry partner ensures both rigor and relevance of our research project (Hevner, 2007). Regarding Venable et al. (2016), we perform a formative evaluation, both ex ante and ex post to determine, how well our artifact achieves its expected environmental utility. The ex ante evaluation can be considered *artificial*, the ex post evaluation as *naturalistic*. In our context the "Human Risk & Effectiveness" DSR evaluation strategy is suitable (Venable et al., 2016), as our preliminary online experiment is an early formative evaluation of the later naturalistic evaluation in the field test.

4 Designing Real-Time Delphi platforms

First, we want to derive the design principles based on a literature review on Real-Time Delphi platforms. Wagner & Back (2008) defined principles for the class of Group Wisdom Support Systems (GWSS) as an extension of Group Decision Support Systems. Following design principles can be understood as a refinement of the principles of Wagner & Back (2008) for the class of Real-Time Delphi. Literature reports of many applications of Real-Time Delphi studies. However, only few studies reported about the implementation of the method. Probably the two most important "clusters of publications" are those around Gordon & Pease (e.g., Gordon & Pease (2006)) and the European Business School (EBS), resp. Gnatzy et al. (e.g., Gnatzy et al. (2011)). In addition, a cluster of studies are based on the eDelfoi platform (see e.g., Kuusi et al. (2000) or https://metodix.fi). A full comparison of the results is provided in Table 1, structured according to the key principles of the Delphi method. Hereinafter, general findings on an abstracted level in the form of design principles are discussed.

Design Principle 1: *Ensure anonymity*

Throughout all publications of implementations of Real-Time Delphi platforms, we find that the participants remained concealed to the other participants (Gnatzy et al., 2011, i.a.).

Design Principle 2: *Provide meaningful statistical group response allowing self-location*

Researchers in the field of Delphi methodology agree that feedback is crucial to achieve results of high quality in Delphi and Real-Time Delphi studies (Best, 1974; Rowe & Wright, 1999; Rowe et al., 2005). Iterative examination of group response is essential to find valid consensus on a topic and also to form clusters around alternative positions (Best, 1974). Early Real-Time Delphi studies only displayed basic statistical group response, like the average. Soon many authors started to introduce more meaningful measures as the additional presentation of the median (Gordon, 2009), min-max values (Gordon et al., 2015), the interquartile range (IQR) (Gary & von der Gracht, 2015; Gnatzy et al., 2011; Schuckmann et al., 2012), or the standard deviation (SD) (Schuckmann et al., 2012). These measures help the participant to locate his own opinion within the overall range. Less conformity exists regarding the presentation of the number of responses. Gnatzy et al. (2011) mentions the problem, that the feedback pushes participants towards conformity. However, only few authors use Real-Time Delphi without the presentation of the group response, as e.g. Steinert (2009), who explicitly wants to find dissension.

Design Principle 3: *Use visual feedback to ease the understanding of the statistical group response*

[2] http://orakel.faz.net
[3] Frankfurter Allgemeine Zeitung, one of the largest German news magazines

Closely related to design principle 2, diverse authors decided to enrich the statistical group response with graphics (Gary & von der Gracht, 2015; Gnatzy et al., 2011; Schuckmann et al., 2012).

Design Principle 4: *Hide feedback before first estimation to avoid anchors*

We add this design principle as it fits more to the traditional Delphi process. Existing implementations show low conformity on this, as only the EBS researches stick to this principle. However, we share the argumentation of Gnatzy et al. (2011) that otherwise the experts become consciously or unconsciously influenced by other participants while forming their own estimation.

Design Principle 5: *Guide the expert but allow free navigation*

Literature reports of two different questionnaire structures: 1D and 2D layouts. In the 1D layout (Gnatzy et al., 2011) experts see one question per screen and proceed to the next question by a button. This reduces information overload and allows to set focus on one question. The order of the questions can be used intentionally, for example to take the expert mentally further in the future by every question. In contrast, in the 2D layout the experts face all questions (in a minimized form) at one glance and can therefore choose the order to answer questions. This is suitable for utility matrixes, decision models, input/output, or, for example, cross impact. The argumentation must be opened in an external window. Though both implementations exist, the 1D layout is implemented more often. In addition, the 1D instances provide the possibility to navigate to each question directly. So, we argue that the user should feel guided by the software but can access any question at any time. Both not necessarily must happen with the 1D or the 2D layout.

Design Principle 6: *Indicate dissension to highlight where other participants have different opinions*

Most implementations included an indicator showing the participant that his opinion is out of the group's estimation. This is achieved either by an "indicator flag" (Gordon, 2007-Gordon & Pease, 2006) or by a "color code" (Gary & von der Gracht, 2015; Gnatzy et al., 2011). Some authors also added a consensus portal, others suggested a dissension portal. Authors that did implement none of the above, usually used a 2D matrix layout that had a comparable function.

Design Principle 7: *Enable argumentation to allow qualitative discussion*

Argumentation is a key aspect of the Delphi method. Therefore this is included in all Real-Time Delphi implementations that are reported in literature. In one case each argument gets reviewed by an administrator before it appears for the other participants. The administrator's task is not to evaluate, but to check the arguments on two criteria: First, eliminate spelling mistakes, and second, delete duplicates to avoid information overload (Linstone & Turoff, 2011).

Design Principle 8: *Allow access at any time to make the process asynchronous*

Throughout all implementation we find the implementation of an "asynchronous process", so that the access was provided at any time. Gordon & Pease (2006) is arguing that this comes close to the iterative process in the traditional Delphi method. Not yet "standard" and therefore not a design principle, but interesting anyway, is the approach of Kuusi et al. (2000) to organize "hot periods", where all participants were additionally invited for specified one or two-hour slots to forecast together. This meets the iterative process even better, besides having other advantages (Gordon, 2009): First, this promotes active participation and keeps the discussion ongoing. Second, this ensures that the experts assess the questions simultaneously and also recognize their personal influence shown by immediate reactions of other participants.

With other design decisions that are reported in literature on Real-Time Delphi, we find both insufficient consensus and no need to use them as design principles. These are the realizations of the following elements: "deception of progress," "argument input," the "estimation input;" and additional information such as: "confidence," "access to survey," the "end" of the survey (after enough estimation or enough time), and "tutorials and introduction". An important issue, but not a real design principle, is that several authors are arguing that the experts should not start with a "null questionnaire" and that initial estimations should be provided by for example a beta test panel etc. (Gordon & Pease, 2006). An alternative is to provide extensive supporting material, definitions, references, or a supportive framework (Gordon & Pease, 2006; Steinert, 2009).

As previously mentioned in Subsection 2.2, we have assumed that more options for social interaction can raise retention. So, we formulate a new design principle "Enable social interaction to promise the gain of social reputation" that should not harm anonymity.

Design Principle 9: *Enable social interaction to promise the gain of social reputation*

To raise retention, it is necessary to stimulate experts to a higher commitment to the survey (Linstone & Turoff, 2011). Therefore, we want to create a stronger bond to the platform by enabling the experience of an online presence. This is achieved by increasing addressability or the promise of gaining social reputation (Bolger & Wright, 2011). The first one must be handled with caution, as anonymity must be ensured. The second may introduce easy to use functionality for feedback. The introduction of, for example, labels that can be added to arguments by all participants, enables each participant to give and receive social reputation (Rainie, 2007).

A Real-Time Delphi, implementing the ninth design principle, can be regarded as a social Real-Time Delphi (sRTD) as suggested by Kloker et al. (2016).

5 Instantiating the Design

5.1 Cycle 1: Prototype Implementation and Evaluation Study

For the first design cycle a prototype was instantiated implementing design principles 1-5, 7, and 8 and is evaluated in an online experiment. As this was a one-shot experiment, design principle 6 was skipped, since dissension indicators first become relevant in a second visit.

5.1.1 Method

A two-treatment, between-subject online experiment was conducted, following experimental procedures from experimental economics (Roth et al., 1986). Therefore, two instances of the platform were set up. From a technical perspective, the online experiment uses a customized web-application following guidelines for online experiments as proposed by Mason & Suri (2012). The first instance implemented a standard Real-Time Delphi (RTD) setup. The second instance implemented a social Real-Time Delphi (sRTD) and, therefore, also implemented Design Principle 9. It is suggested that Design Principle 9 raises retention in a Real-Time Delphi context.

Stimuli & Experiment Design: Both groups participated in a Real-Time Delphi survey, from which the control group was confronted with the standard version (RTD) and the treatment group with the social version (sRTD). The social elements offered together in the sRTD are illustrated in Fig. 2:

- **Labeling:** The arguments and opinions that have been added to the responses by the participants can be marked with labels such as "good," "bad," or "helpful."
- **Generated user names:** In each question participants receive a generated user name, which is used to mark their arguments as their own while maintaining their anonymity. These user names can then be used to reference a single person in other arguments or to see which arguments were provided by the same person.

Our research model is illustrated in Fig. 3. According to our theoretical considerations in Subsection 2.2 and for the Design Principle 9, we argued that the *(perceived) Promise of Social Reputation* (introduced by labeling, abbr. "*Reputation*") has a positive effect of *Retention* (H_{12}). As this was a one-shot experiment and there was no additional round, we operationalized our dependent variable *Retention* with a self-defined proxy construct *Commitment*. Based on our considerations in Subsection 2.2 we also assumed that the generated user names raise *(perceived) Addressability*, which influences both *Commitment* (H_{11}) and *Reputation* (H_{13}) positively.

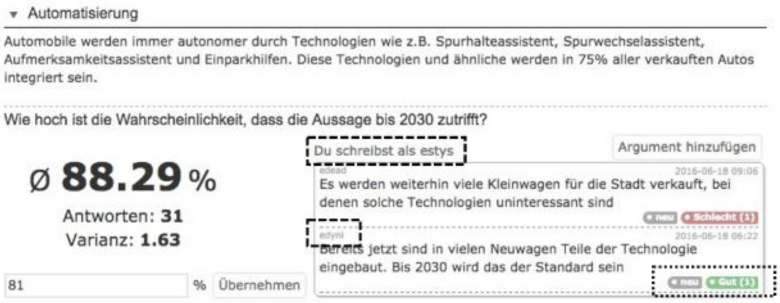

Figure 2: Social elements offered to the treatment group (sRTD) in the argument area of a question. Dashed boxes highlight the generated user names in the illustration. The dotted boxes highlight the labels. These elements were missing in the control group.

Participants: The participants were mainly students of Industrial Engineering, Economics, and Business Information Systems from the Karlsruhe Institute of Technology (KIT), Germany. From a contacted 50 participants, 46 answers returned, 22 (female: 9) in the RTD and 24 (female: 10) in the sRTD[4]. For our analysis we only considered participants that completed both the survey and the attached questionnaire. In the RTD / sRTD participants were asked about their opinion on the future of the automobile industry. Therefore, we expected that this topic would not overly attract or bore the students. Due to their technical background, we expected them to have at least some expertise.

Procedure: The experiment was conducted in the context of a research seminar. Two students spread the invitation to the RTD survey among their acquaintances via email. The recipients were asked to participate in the RTD survey and give their estimations on the future of electronic vehicles. Participants accessed the platform using a personalized link to ensure them ending up seeing their own answers when accessing the survey multiple times. Therefore, aborting and re-accessing the survey was possible at any time. The assignment of participants to the treatments was randomized and participants were not aware of treatments. After finishing the survey, the participants received a questionnaire asking them to self-assess their experience during the survey. The follow-up questionnaire was conducted with the same platform, however, group feedback and the opportunity to provide arguments was suppressed. The questionnaire consisted of 22 items, which were related to one of the following constructs: *(perceived) Addressability*[5] (user names), *(perceived) Social Reputation*[6] (labels), and *Commitment*[7]. Some items were negated to check for consistent answers. The survey was open for one week. After ten days a mail was sent to those participants that participated in the first round, saying "thank you" for participation and offered them the opportunity to re-access the platform to see the results, and optionally change their opinion. Still, the participants were not aware of the treatments. If they accessed the platform again we logged the event.

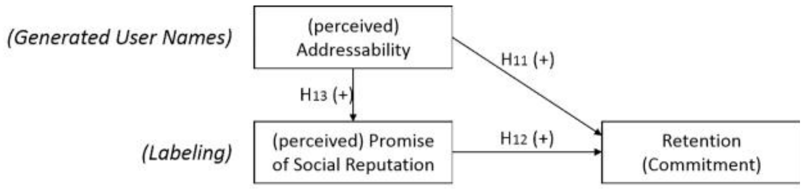

Figure 3: Research model for online experiment (prototype evaluation).

[4] A two-sample test for equality of proportions shows no significant inequality (p=.958)
[5] Items were taken from Gefen & Straub (2004); Lin (2004); Rovai (2001)
[6] Items were taken from Kankanhalli et al. (2005); Wang & Wang (2010); we added five self-formulated items
[7] Self-formulated

5.1.2 Evaluation of the Prototype

Since gender was equally distributed between the control and treatment groups, it is not necessary to control for it in the overall effect. The number of participants is relatively small, so we will only focus on the main trends. The constructs *Addressability, Reputation,* and *Commitment* in our questionnaire received a Cronbach's α of 0.763, 0.786, and 0.548. Only the last value does not exceed the threshold of 0.7, which is usually regarded for constructs to be "acceptable". The construct only consisted of two items and was formulated by our team due to a lack of pre-evaluated items in the literature. Therefore, it (these conclusion / data) needs to be handled with caution.

A polyserial correlation between *Commitment* and the treatment ("social" as reference level) showed a significant correlation (r=-.06, p=.039)[8]. Yet the data report (data is plural) no significant correlation between the treatment and the revisit after the last email (ϕ-Coefficient = -0.13). The other constructs did show a slight, but not significant rise from the RTD to the sRTD treatment. However, a multiple linear regression showed that *Commitment* is, by a considerable portion, explained by *Social Reputation*. A multiple linear regression was calculated to predict Commitment based on Social Reputation and *Addressability*. A significant regression equation was found (F(2,41)=9.828, p=.000), with a R^2 of .324. Participants' predicted *Commitment* is equal to 0.134 + 0.018(n.s.) *Addressability* + 0.771*** *Social Reputation*, where both independent and dependent constructs were coded or measured on a Likert-Scale from 1 to 5 ("fits not at all" to "fits completely"). Adding gender as an independent variable resulted in no significant effect (0.263(n.s.) "female"). Therefore, H_{11} has to be rejected, while H_{12} is accepted. A weak positive correlation was, in addition, found between *Addressability* and *R*eputation r(43)=0.31* (H_{13}).

As an interim conclusion, we can say that the social treatment lead to a significant rise in self-reported *Commitment* of the participants to the survey and that commitment is by a large portion explained by the self-reported perceived possibility to gain social reputation. A connection between the perceived addressability and commitment, as expected by the design element of the generated user names, could not be shown, though it had a weak correlation with the promise to gain social reputation. Therefore, we dropped this design element in the field test in order to further explore the isolated effect of social reputation on retention and recommend it for future research.

5.2 Cycle 2: IT Artifact Implementation and Evaluation Study

For the second design cycle a sRTD artifact was instantiated on the FAZ.NET-Orakel and a two round sRTD survey was implemented for the German Federal Elections 2017.

5.2.1 Method

For the evaluation of the IT artifact, a field study was conducted. The design element of the generated user names was dropped, due to the only small effects in the prototype evaluation and as it would have confused participants on the FAZ.NET-Orakel, which was not in the sense of our industry partner. Besides, the artifact provided identical functionality than the prototype. The survey was conducted in the context of a prediction market for the vote-share of the parties.

Study Design: Our research model for the naturalistic evaluation is illustrated in Fig. 4. Based on the findings in the prototype evaluation, that *(perceived) Promise of Social Reputation* showed a positive effect (see Subsection 5.1), our considerations regarding Design Principle 9, and Bolger & Wright (2011), it is now hypothesized that positive (actual positive) *Social Reputation* increases *Subsequent Platform Engagement* regarding *Retention* (H_1) and *Activity* (H_2). *Retention* is operationalized as the participation of a user in the second round, if the user participated in the first round. *Activity* is an additional construct that measures the overall interaction with the sRTD in the second round. It allows a

[8] Signif. codes for all analyses: 0 *** 0.005 ** 0.01 * 0.05 . 0.1 n.s. 1

more graduated evaluation of retention. E.g., if all participants from the first round would have participated in the second round, it would still allow to further evaluate the effect of the *Social Reputation*.

As this was a field test, many other influences may interfere with the dependent construct. Mere increased overall activity could also explain both, (quantitatively) higher Social Reputation and *Subsequent Platform Engagement*. Because of restrictions on the platform, no control treatment could be performed and, therefore, we cannot finally distinguish between causality and correlation. If a person is very active on the platform in general, this may also explain higher *Subsequent Platform Engagement*. To control for such influences, we modeled the overall activity of the platform in the proxy construct *Engagement in Prediction Market* as an independent control variable with the hypothesis that the overall increased activity has no effect on the *Subsequent Platform Engagement* (H_3 and H_4) in our sRTD. We assumed no effect of the prediction market as a connected platform on the dependent constructs as neither the incentives, nor the objectives were presented as linked to the participants.

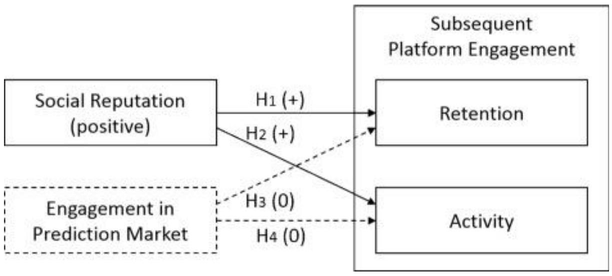

Figure 4: Research model for field test (artifact evaluation).

From the collected field data we interpreted the absolute number of received positive labels before the second round as *Social Reputation*, the absolute number of trades in the prediction market for the party results before the second round as *Engagement in Prediction Market*, a logical indicator if the participant participated in the second round (either with answers, comments or labels) as *Retention*, and the absolute number of answers, comments and labels in the second round as *Activity*. We only focus on those participants that participated already in the first round.

Participants: Participants were only recruited among the participants of the prediction market. Here we can legitimately assume a certain degree of self-selection towards a knowledgeable group. Participation in the Real-Time Delphi survey was completely optional and not related to any of the incentives of the prediction market.

Procedure: Technically the software artifact in the field study was identical to the sRTD prototype, besides small design adaptions to the corporate design of the FAZ.NET-Orakel platform. For the field test a survey was implemented asking for the relations between considerations about possible political coalitions and the effect on the expected election outcome for different parties. The survey was always accessible by a link, prominently placed at the top of the corresponding prediction market. Two rounds were performed. Between the two rounds, additional questions (suggestions of the first round by the participants) were added and negatively rated comments were deleted (which is usually the facilitator's task (Gnatzy et al., 2011) and supported by the labels as argued in Subsection 2.2). Round 1 (8th of July to 2nd of August 2017) contained five questions and one question where suggestions could be provided. Round 2 (3rd August to 24th September 2017) added five questions (total: 10 questions). The advertisement for the survey was kept at a minimum: One on-site message at the introduction of the survey and one on-site message at the start of the second round. These measures were taken to ensure that no initial "novelty hype" interfered with the "normal" use of the tool. Some participants have been introduced to the tool itself earlier in a preliminary technical test.

5.2.2 Evaluation of the IT Artifact

Overall, 90 participants participated in the survey by providing answers. Demographic data is not available due to limitations by our industry partner. There is, however, no reason to assume a distinct change of demography between the rounds, and we do not expect any effect (as also shown in Subsection 3). 41 participated only in the first round, 7 participants in both rounds, and 42 participants only in the second round (which are, therefore, not considered in the regressions based on first round activity etc.). During the two rounds (first round), 29 (20) comments and 40 (14) labels were provided by the participants. We test the paths from our independent constructs to *Retention* as illustrated in our research model for the field test (see Fig. 4) due to its binary coding with probabilistic regressions, paths to *Activity* due to its continuous coding with linear regressions. We always test the predictive power of the measures collected of the first round on the behavior in the second round.

A simple probabilistic regression was calculated to predict *Retention* based on *Social Reputation* during the first round. A weakly significant regression equation was found (p=.093), with (pseudo[9]) R^2 of 0.345. Participants' predicted likelihoods to retain participation is equal to -1.132 + 1.213. *Social Reputation* is measured in the absolute number of positive labels that were received by a participant's comments. A simple linear regression was calculated to predict Activity based on *Social Reputation* during the first round. A significant regression equation was found (F(1,26)=5.22, p=.031), with R^2 of 0.167. Participants' predicted likelihoods to participate actively in the second round is equal to 0.7232 + 1.0577 "content-creating activities".

A simple probabilistic regression was calculated to predict *Retention* based on *Engagement in Prediction Market* during the first round. A not significant regression equation was found (p=.882), with (pseudo) R^2 of 0.001. *Engagement in Prediction Market* is measured in the absolute number of trades in the corresponding markets. A simple linear regression was calculated to predict *Activity* based on *Engagement in Prediction Market* during the first round. A not significant regression equation was found (F(1,26)=5.22, p=.603), with R^2 of 0.011.

In conclusion, participants that received positive feedback in the form of labels in the first round had a significantly higher probability of retention for the second round survey. This positive feedback also correlates with higher activity in the second round. These effects cannot be explained by a general higher activity, as *Engagement in Prediction Market* had no significant correlations with the dependent constructs. Therefore, these results suggest a causality in favor of our expectations, though this cannot be definitively proven, due to a non-existent control group.

6 Discussion and Conclusion

This paper presents the result of our DSR project adding a social character to Real-Time Delphi to raise participant retention during the surveys. In the first design cycle, eight design principles for Real-Time Delphi platforms were identified and formulated, based on literature. The examination of existing Real-Time Delphi platforms showed that mainly two central approaches are established in current research and that they differ especially regarding the survey layout and the point in time when the group estimation is presented. In both cases we decided to formulate design principles in a way that appeared to be closer to the traditional Delphi method. Subsequently, a ninth design principle was added: *Enable social interaction to promise the gain of social reputation*. We derived the need for this design principle out a lack of retention (Mullen, 2003; Okoli & Pawlowski, 2004; Reid, 1988) during Real-Time Delphi studies and literature, which is arguing that commitment to online platforms can be raised by social elements (Bolger & Wright, 2011; Kloker et al., 2016; Lampel & Bhalla, 2007).

Thereafter, a prototype of the social Real-Time Delphi was implemented considering all design principles relevant for our preliminary experiment. For the ninth design principle we made the design decisions, as suggested by Kloker et al. (2016), to use generated user names per questions to maintain anonymity and at the same time guarantee addressability. In addition, we provided the possibility to add labels to arguments, in order that the platform promised users to perceive social reputation when others appreciate their contributions. A two-treatment, between-subject online experiment, and a follow-

[9] McFadden

up questionnaire showed, that the assumption (that the promise to gain social reputation by social elements raise the self-reported commitment) is valid. Though theory suggested, an interaction or positive effect of the generated user names could not be demonstrated. This may have been due to several reasons: First, the generated user names did resemble vocals and consonants in order to be readable. However, they were not common names and, therefore, may not have induced “social presence” or were recognized as real names for real persons. Second, we did not explain to the participants, how and for what reasons the user names were generated. Based on the description, they may have assumed that each user gets an individual user name per question. We however, did not examine this and the lack of social presence induced by the user names may be due to a lack of understanding. Nevertheless, as we could not show an effect on our dependent variable, the design element of generated user names was discarded in the field test. However, we would suggest it for further evaluation in future studies.

The prototype was subsequently instantiated as a full IT artifact and evaluated in a field study. A two round Real-Time Delphi survey was conducted in the context of the Federal Elections 2017 in Germany on a prediction market platform. We found a significant positive effect of the received positive labels (*Social Reputation*) in the first round on *Retention* in the second round and *Activity* in the second round. This effect cannot finally be attributed to a causality, due to the lack of a control treatment in the field test based on restrictions of our industry partner. However, we controlled for the effect that both may be explained by the general activity and engagement of the participants in the overall platform, which could be refused. Therefore, both theory and our results indicate that the introduction of social elements to Real-Time Delphi surveys raise *Subsequent Platform Engagement*.

Limitations of the current work regarding our literature review are especially laid in the bad accessibility of descriptions of many Real-Time Delphi platforms. Hence, we could only consider those which were described in literature. Limitations of our preliminary experiment are the modest sample size and the questionable reliability of the self-formulated construct *Commitment*. Hence, strong implications should be made cautiously. Future studies may better be based on previously validated constructs, e.g., “IS continuance intention” by Bhattacherje (2001). We decided to formulate the items our-selves in order to do justice to the round-based character of the study. Nevertheless, as Real-Time Delphi studies also highlight the asynchronous character, pre-validated measures on general IS continuance and retention independently of rounds may be the better fit for future studies. Limitations of our field test are the lack of a control treatment, which was not possible in the given context on the platform FAZ.NET-Orakel. Therefore, causality and correlation cannot be finally distinguished. We also notice that bad comments (labeled by the participants) were deleted between the first and the second rounds by the moderator to be consistent with traditional Delphi method, where this is usually done by the researchers. However, this may also have an effect that we cannot estimate in our study. At last, though we showed a significant effect of positive social reputation on retention, this does not mean that it isjust necessary to provide all participants with positive feedback (potentially by the moderator). There is, probably, a certain trade-off with credibility and ludicrousness, which, however, can also not be derived from the current study and should be subject to future research. In addition, the field test should be replicated in the future with increased observations and at several points in time to create stronger evidence for the hypothesis and to show a stable effect over time.

Nonetheless, the work at hand contributes in many ways to the current research’s question, how to encourage participants to take part in knowledge sharing over a long time. We enrich existing literature by formulating the design principles for Real-Time Delphi platforms. In addition our experiment and field test provided evidence in favor of the expedience of the ninth design principle in an experiment setting as well as with real-world data. We showed that implementing social elements in actual anonymous online settings raises commitment and retention. So, considering our ninth design principle helps researchers and practitioners to conduct Real-Time Delphi surveys more successfully, risk less drop-outs, and at the same time raise commitment. This is related with less cost and a higher probability of success. Further research may prove this claim to be true for other knowledge management systems that require subsequent platform engagement.

References

Abadie, F., Friedewald, M. Weber, M. 2010, 'Adaptive foresight in the creative content industries: anticipating value chain transformations and need for policy action', *Science and Public Policy* 37(1), 12. URL: http://spp.oxfordjournals.org/content/37/1/19.full.pdf

Ames, M., Naaman, M. 2007, Why We Tag: Motivations for Annotation in Mobile and Online Media, In:'*Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*', CHI '07, ACM, New York, NY, USA, pp. 971–980. URL: http://doi.acm.org/10.1145/1240624.1240772

Best, R. J. 1974, 'An Experiment in Delphi Estimation in Marketing Decision Making', *Journal of Marketing Research* 11(4), 448–452. URL: http://www.jstor.org/stable/3151295

Bhattacherje, A. 2001, 'Understanding Information Systems Continuance: An Expectation Confirmation Model', *MIS Quarterly* 25(3), 351–370.

Bolger, F., Wright, G. 2011, 'Improving the Delphi process: Lessons from social psychological research', *Technological Forecasting and Social Change* 78(9), 1500–1513. URL: http://www.sciencedirect.com/science/article/pii/S0040162511001430

Boulkedid, R., Abdoul, H., Loustau, M., Sibony, O. Alberti, C. 2011, 'Using and Reporting the Delphi Method for Selecting Healthcare Quality Indicators: A Systematic Review', *PLOS ONE* 6(6), e20476. URL: https://doi.org/10.1371/journal.pone.0020476

Dalkey, N. C., Brown, B. B. Cochran, S. 1969, The Delphi method: An experimental study of group opinion, Vol. 3, Rand Corporation Santa Monica, CA. URL: https://www.rand.org/content/dam/rand/pubs/research_memoranda/2005/RM5888.pdf

Darkow, I.-L., von der Gracht, H.A. 2013, 'Scenarios fort eh future of the European process industry – the case of the chemical industry.' *European Journal of Futures Research* 1(1), 1-12. URL: http://dx.doi.org/10.1007/s40309-013-0010-9

Durand, R. 2003, 'Predicting a firm's forecasting ability: the roles of organizational illusion of control and organizational attention', *Strategic Management Journal* 24(9), 821–838.

Gary, J. E., von der Gracht, H. A. 2015, 'The future of foresight professionals: Results from a global Delphi study', *Futures* 71, 132–145.

Gefen, D., Straub, D. W. 2004, 'Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services', *Omega* 32(6), 407–424. URL: http://www.sciencedirect.com/science/article/pii/S0305048304000131

Gnatzy, T., Warth, J., von der Gracht, H. Darkow, I.-L. 2011, 'Validating an innovative real-time Delphi approach - A methodological comparison between real-time and conventional Delphi studies', *Technological Forecasting and Social Change* 78(9), 1681–1694. URL: http://www.sciencedirect.com/science/article/pii/S0040162511000813

Gordon, T. J. 2007, 'Energy forecasts using a 'Roundless' approach to running a Delphi study', *Foresight* 9(2), 27–35.

Gordon, T. J. 2009, The real-time Delphi method, Technical report, The Millennium Project. URL: http://www.econ.uba.ar/unai/archivos/actividades/30_de_octubre/Real_Time_Delphi_Ted_Gordon.pdf

Gordon, T. J., Pease, A. 2006, 'RT Delphi: An efficient, "round-less" almost real time Delphi method', *Technological Forecasting and Social Change* 73(4), 321–333. URL: http://www.sciencedirect.com/science/article/pii/S0040162505001435

Gordon, T. J., Sharan, Y. Florescu, E. 2015, 'Prospects for Lone Wolf and SIMAD terrorism', *Tecnological Forecasting and Social Change* 95, 234–251.

Hevner, A. R. 2007, 'A three cycle view of design science research', *Scandinavian journal of information systems* 19(2), 4.

Hevner, A. R., March, S. T., Park, J. Ram, S. 2004, 'Design Science in Information Systems Research', *MIS Quarterly* 28(1), 75–105. URL: http://www.jstor.org/stable/25148625

Ishikawa, A., Amagasa, M., Shiga, T., Tomizawa, G., Tatsuta, R. Mieno, H. 1993, 'The max-min Delphi method and fuzzy Delphi method via fuzzy integration', *Fuzzy Sets and Systems* 55(3), 241–253. URL: http://www.sciencedirect.com/science/article/pii/016501149390251C

Kankanhalli, A., Tan, B. C. Y. Wei, K.-K. 2005, 'Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation', *MIS Quarterly* 29(1), 113–143. URL: http://www.jstor.org/stable/25148670

Kaye, A. 1992, 'Collaborative learning through computer conferencing, In: *Collaborative learning through computer conferencing*' Ed. by A. Kaye, Vol. 90, Springer, chapter Learning t, pp. 1–24. URL: http://link.springer.com/chapter/10.1007/978-3-642-77684-7_1

Kloker, S., Kranz, T. T., Straub, T. Weinhardt, C. 2016, Shouldn't Collaboration be social? - Proposal of a social Real-Time Delphi, In: *'Proceedings of the Second Karlsruhe Service Summit Research Workshop'*.

Kloker, S., Straub, T. Weinhardt, C. 2017, Designing a Crowd Forecasting Tool to Combine Prediction Markets and Real-Time Delphi, In: *'Designing the Digital Transformation. DESRIST 2017. Lecture Notes in Computer Science'* Ed. by A. Maedche, J. vom Brocke A. Hevner, 10243rd Edition, Springer International Publishing, Springer, Cham, pp. 468–473. URL: http://dx.doi.org/10.1007/978-3-319-59144-5_33

Kochtanek, T. R., Hein, K. K. 1999, 'Delphi study of digital libraries', *Information Processing & Management* 35(3), 245–254. URL: http://www.sciencedirect.com/science/article/pii/S0306457398000600

Kuechler, B., Vaishnavi, V. 2008, 'On theory development in design science research: anatomy of a research project', *European Journal of Information Systems* 17(5), 489–504. URL: http://dx.doi.org/10.1057/ejis.2008.40

Kuusi, O., Hiltunen, E. Linturi, H. 2000, 'Heikot signaalit - Delfoi-tutkimus (in Finnish; engl. Weak Signals - a Delphi Study)', Futura 2, 78–92.

Lampe, C., Wash, R., Velasquez, A. Ozkaya, E. 2010, Motivations to Participate in Online Communities, In: *'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems'*, CHI '10, ACM, New York, NY, USA, pp. 1927–1936. URL: http://doi.acm.org/10.1145/1753326.1753616

Lampel, J., Bhalla, A. 2007, 'The Role of Status Seeking in Online Communities: Giving the Gift of Experience', *Journal of Computer-Mediated Communication* 12(2), 434–455. URL: http://dx.doi.org/10.1111/j.1083-6101.2007.00332.x

Landeta, J. 2006, 'Current validity of the Delphi method in social sciences', *Technological Forecasting and Social Change* 73(5), 467–482. URL: http://www.sciencedirect.com/science/article/pii/S0040162505001381

Lin, G.-Y. 2004, Social Presence Questionnaire of Online Collaborative Learning: Development and Validity, In: *'Association for Educational Communications and Technology, 27th, Chicago, IL, October 19-23, 2004'*, pp. 588–591. URL: http://files.eric.ed.gov/fulltext/ED484999.pdf

Linstone, H. A., Turoff, M. 1975, *The Delphi method: Techniques and applications*, Vol. 29, Addison-Wesley Reading, MA.

Linstone, H. A., Turoff, M. 2002, *The Delphi Method: Techniques and applications*, Addison-Wesley Pub. Co.

Linstone, H. A., Turoff, M. 2011, 'Delphi: A brief look backward and forward', *Technological Forecasting and Social Change* 78(9), 1712–1719. URL: http://www.sciencedirect.com/science/article/pii/S0040162510002155

Martensen, M., Ryschka, S. Bick, M. 2016, Studying Social Software Adoption by Management Consultants: Use and Application Categories for Knowledge Management BT, In: *'Innovations in Knowledge Management: The Impact of Social Media, Semantic Web and Cloud Computing'*, Ed. by L. Razmerita, G. Phillips-Wren L. C. Jain, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 101–126. URL: https://doi.org/10.1007/978-3-662-47827-1_5

Mason, W., Suri, S. 2012, 'Conducting behavioral research on amazon's mechanical turk', *Behavior research methods* 44(1), 1–23.

Mullen, P. M. 2003, 'Delphi: myths and reality', *Journal of Health Organization and Management* 17(1), 37–52. URL: http://dx.doi.org/10.1108/14777260310469319

Murphy, E. 2004, 'Recognising and promoting collaboration in an online asynchronous discussion', *British Journal of Educational Technology* 35(4), 421–431. URL: http://www.ucs.mun.ca/ emurphy/bjet_401.pdf

Okoli, C., Pawlowski, S. D. 2004, 'The Delphi method as a research tool: an example, design considerations and applications', *Information & Management* 42(1), 15–29. URL: http://www.sciencedirect.com/science/article/pii/S0378720603001794

Rainie, L. 2007, 28% of online americans have used the internet to tag content, Technical report. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2007/PIP_Tagging.pdf.pdf

Reid, N. 1988, The Delphi technique, In: *'Professional Competence and Quality Assurance in the Caring Professions'*, Ed. by R. Ellis, Chapman & Hall, London, chapter 9, pp. 230–254. URL: http://dx.doi.org/10.1016/0020-7489(90)90106-S

Remidez Jr, H., Joslin, C. 2007, Using Prediction Markets to Support IT Project Management-Research in Progress, In: *'International Research Workshop on IT Project Management 2007'*, p. 10.

Rodrigues, E. M., Milic-Frayling, N. Fortuna, B. 2008, Social Tagging Behaviour in Community-Driven Question Answering, In: *'Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology - Volume 01'*, WI-IAT '08, IEEE Computer Society, Washington, DC, USA, pp. 112–119. URL: http://dx.doi.org/10.1109/WIIAT.2008.138

Roschelle, J., Teasley, S. D. 1995, The Construction of Shared Knowledge in Collaborative Problem Solving, In: *Computer Supported Collaborative Learning*, Ed. by C. O'Malley, Vol. 128 of NATO ASI Series, Springer Berlin Heidelberg, chapter 1,5, pp. 69–97. URL: http://link.springer.com/chapter/10.1007/978-3-642-85098-1_5

Roth, A. E. et al. 1986, 'Laboratory experimentation in economics', *Economics and Philosophy* 2(02), 245–273.

Rovai, A. P. 2001, 'Building classroom community at a distance: A case study', *Educational Technology Research and Development* 49(4), 33. URL: https://doi.org/10.1007/BF02504946

Rowe, D. A., C. Y. Alverson, D. K. Unruh, C. H. Fowler, R. Kellems, and D. W. Test (2015). 'A Delphi study to operationalize evidence-based predictors in secondary transition.' *Career Development and Transition for Exceptional Individuals* 38(2), 113-126.

Rowe, G., Wright, G. 1999, 'The Delphi technique as a forecasting tool: issues and analyses', *International Journal of Forecasting* 15(4), 353–375. URL: http://www.sciencedirect.com/science/article/pii/S0169207099000187

Rowe, G., Wright, G. Bolger, F. 1991, 'Delphi: A reevaluation of research and theory', *Technological Forecasting and Social Change* 39(3), 235–251.

Rowe, G., Wright, G. McColl, A. 2005, 'Judgment change during Delphi-like procedures: The role of majority influence, expertise, and confidence', *Technological Forecasting and Social Change* 72(4), 377–399.

Schuckmann, S. W., Gnatzy, T., Darkow, I.-L., von der Gracht, H. A. 2012, 'Analysis of factors influencing the development of transport infrastructure until the year 2030 - A Delphi based scenario study', *Technological Forecasting and Social Change* 79(8), 1373–1387.

Steinert, M. 2009, 'A dissensus based online Delphi approach: An explorative research tool', *Technological Forecasting and Social Change* 76(3), 291–300.

Strauss, H. J. Zeigler, L. 1975, 'Delphi, political philosophy and the future', Futures 7(3), 184–196.

Styhre, A. 2002, 'The knowledge-intensive company and the economy of sharing: rethinking utility and knowledge management', *Knowledge and Process Management* 9(4), 228–236. URL: http:https://doi.org/10.1002/kpm.155

Turoff, M., Hiltz, S. R., Yao, X., Li, Z., Wang, Y. Cho, H. K. 2004, Online Collaborative Learning Enhancement Through the Delphi Method, In: *'Proceedings of the OZCHI 2004 Conference, University of Wollongong, Australia, November 22-24, 2004'*. URL: https://web.njit.edu/ turoff/Papers/ozchi2004.htm

Venable, J., Pries-Heje, J. Baskerville, R. 2016, 'FEDS: a Framework for Evaluation in Design Science Research', *European Journal of Information Systems* 25(1), 77–89. URL: https://doi.org/10.1057/ejis.2014.36

Wagner, C., Back, A. 2008, 'Group Wisdom Support Systems: Aggregating the Insights of many trough Information Technology', *Issues in Information Systems* (IIS) 9(2), 343–350. URL: http://iacis.org/iis/2008/S2008_992.pdf

Walker, A. M., Selfe, J. 1996, 'The Delphi method: a useful tool for the allied health researcher', British Journal of Therapy and Rehabilitation 3(12), 677–681.

Wang, H.-Y., Wang, S.-H. 2010, 'Predicting mobile hotel reservation adoption: Insight from a perceived value standpoint', *International Journal of Hospitality Management* 29(4), 598–608. URL: http://www.sciencedirect.com/science/article/pii/S0278431909001510

Zipfinger, S. 2007, '*Computer-aided Delphi: an experimental Study of comparing round-based with real-time implementation of the method*'. Trauner.

Properties (↓) / Publications (→)		Gordon et al.			EBS				Other approaches		
		2007; 2006	2009	2015	Gnatzy et al., 2011	Schuckmann et al., 2012	Darkow and Gracht, 2013	Gary and Gracht, 2015	Steinert, 2009	eDelfoi	Zipfinger, 2007
Anonymity		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Statistical Group Response	provided	✓	✓	✓	✓	✓	✓	✓	✗	n/a	✓
	average	✓	✓	✓	✓	✓	n/a	✓	✗	n/a	n/a
	other statistics	n/a	median	median, minmax	median, IQR	median, IQR	n/a	IQR, SD	✗	n/a	n/a
	# of responses so far	✓	✓	✓	✗	✓	n/a	✗	✗	n/a	n/a
Controlled Feedback	dissension indicator	flag	flag	n/a	color-code	n/a	n/a	color-code	✗	n/a	n/a
	consensus portal	✗	✗	✗	✓	n/a	n/a	✓	dissension portal	n/a	n/a
	group response ... individual response	before	before	before	after	after	after	after	before	n/a	n/a
	others arguments	✓, extra window	✓, extra window	✓	✓, extra window	✓	✓	✓	✓	n/a	✓
	review of arguments	✗	✗	✗	✓	n/a	n/a	n/a	n/a	n/a	n/a
	visual RT feedback	✗	✗	✗	✓	✓	n/a	✓	✗	n/a	n/a
Iterative Process	access anytime	✓	✓	✓	✓	✓	✓	✓	✓	n/a	✓
	asynchr. process	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	"hot periods"	✗	✗	✗	✗	✗	✗	✗	✗	✓	n/a
General Properties	objective	consensus	consensus	consensus	consensus	consensus	consensus	both	dissension	consensus	consensus
	initial values	beta test panel	n/a	small pilot group	small expert group	small expert group	developed research framework	literature and experts	project was prototype	n/a	n/a
	layout	2D (matrix)	2D (matrix)	2D (matrix)	1D (1-question-1-screen)	1D (1-question-1-screen)	1D (1-question-1-screen)	1D (1-screen-1-factor)	1D (1-screen-1-factor)	n/a	n/a
	depiction of progress	✗	✗	✗	in percent	n/a	n/a	n/a	✗	n/a	n/a
	estimation input	4 categories (not linear)	100 and 10 point Likert scale	numeric, multiple choice	metric scale 0-100%, 5 point Likert scale	metric scale 0-100%, 5 point Likert scale	metric scale 0-100%, 5 point Likert scale	metric scale 0-100%, 5 point Likert scale	multiple choice	n/a	n/a
	argument input	✓, extra window	✓, extra window	✓	✓, directly	✓	✓	✓	✓	n/a	n/a
	internet-based	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
	access to survey	user & password	n/a	n/a	hyperlink (email)	hyperlink (email)	n/a	hyperlink (email)	hyperlink (email)	n/a	n/a
	integrated tutorial	✗	✗	✗	✓	n/a	n/a	n/a	✓	n/a	n/a

Table 1. Summarized properties of selected Real-time Delphi studies.