PAVING THE WAY FOR REAL-TIME DELPHI IN INFORMATION SYSTEMS RESEARCH: A SYNTHESIS OF SURVEY INSTRUMENT DESIGNS AND FEEDBACK MECHANISMS

Research paper

Thiebes, Scott, Karlsruhe Institute of Technology (KIT), Department of Economics and Management, Karlsruhe, Germany, scott.thiebes@kit.edu

Scheidt, Dino, University of Cologne, Cologne, Germany, dino.scheidt@smail.uni-koeln.de

Schmidt-Kraepelin, Manuel, Karlsruhe Institute of Technology (KIT), Department of Economics and Management, Karlsruhe, Germany, manuel.schmidt-kraepelin@kit.edu

Benlian, Alexander, Technical University Darmstadt, Department of Information Systems and E-Services, Germany, benlian@ise.tu-darmstadt.de

Sunyaev, Ali, Karlsruhe Institute of Technology (KIT), Department of Economics and Management, Karlsruhe, Germany, sunyaev@kit.edu

Abstract

Real-time Delphi is an information technology-enabled extension of the popular Delphi method. While Real-time Delphi is acknowledged to decrease study durations and panel-attrition by employing information technology to engage participants in an asynchronous fashion, its broad adoption in the information systems community has yet to be seen. Given the growing wake of ubiquity of the conventional Delphi method in information systems and the technique's lasting implementation uncertainties, now seems to be the right time to revisit Real-time Delphi to give way to calls of a better empirical foundation. This research reviews extant literature on Real-time Delphi and provides a synthesis of employed survey instrument designs and feedback mechanisms (two central aspects of Real-time Delphi). It contributes to the methodological knowledge base of information systems research by serving as an implementation reference and foundation on which researchers can build to establish Real-time Delphi as a novel and promising addition to the information systems research methods repertoire.

Keywords: Delphi, Real-Time Delphi, IS Research Methods, Literature Review.

1 Introduction

Delphi is an expert opinion-based forecasting technique that is often criticized for its long study durations and high panel attrition (Bardecki, 1984). Aiming to overcome the issues of the conventional Delphi method (Delphi), Gordon and Pease (2006) developed the concept of an information technology-enabled real-time extension of Delphi, called real-time Delphi (RT Delphi). Compared to Delphi, RT Delphi is acknowledged to decrease study durations and panel attrition by employing information technology to engage participants in an asynchronous fashion (Gnatzy et al., 2011, Gordon and Pease, 2006), thus enabling researchers to meet their goals more quickly and successfully (Gallego and Bueno, 2014, Gordon, 2007). Adding to this, RT Delphi's technological foundation and asynchronous nature also allow for the collection and analysis of rich and diverse data. Researchers can leverage RT Delphi to collect data beyond self-reported insights, enabling new ways of theorizing, much closer to the aspirations of computational social science (Lazer et al., 2009). An example for is the tracking and analysis of participants' behaviour in real-time to scrutinize how group consensus (or disagreement) is formed over time through interaction and group dynamics. Thereby, RT Delphi may enable scholars to explore a set of diverse or even competing theories based on emergent if-then relationships along a continuum of participant interactions.

Regardless of Delphi's recent momentum in the information systems (IS) field (Skinner et al., 2015), where affinity to information technology can be assumed and where we are constantly striving to employ novel means for querying experts who are chronically short on time, the potential of RT Delphi seems largely untapped. This observation contradicts the Delphi method's continued growth in popularity since its first broad appearance in Linstone and Turoff (1975). A likely reason for this can be found in an increased method complexity and the absence of ready-to-use software tools for researchers seeking to employ RT Delphi. Adding to this, the expanding pressure to overcome Delphi's limitations and the ever-growing number of domain specific applications have led to a pool of hard-to-distinguish 'Delphi techniques' (Hasson and Keeney, 2011, Landeta, 2006, Rowe and Wright, 2011). With the absence of RT Delphi studies in leading IS outlets and the growing number of overlapping Delphi techniques, there is inherent uncertainty in employing RT Delphi in one's research and little to no guidance for researchers interested in conducting RT Delphi studies in our field. We are convinced that this uncertainty and striking absence of guidance for IS researchers as well as the non-existence of ready-to-use software tools that support the RT Delphi process are main reasons for IS researchers not adopting this method. Aiming to address these issues (i.e., method uncertainty, lack of guidance, and absence of tools) for the IS community we therefore ask:

RQ: How can extant research on/employing RT Delphi inform the application and design of RT Delphi in IS research?

To answer our research question, we conduct a three-staged literature review of studies related to RT Delphi by following the approach of Boell and Cecez-Kecmanovic (2014). With this study, we seek to provide guidance to those in our community enticed by an efficient Delphi design. Herein is the main goal to review and cluster extant RT Delphi studies to make them accessible as references for the IS community. Furthermore, this study aggregates and synthesizes identified survey instrument designs and feedback mechanisms to lay the groundwork for a recommendation model of and design-driven research on how to instrument RT Delphi in IS research. In summary, our study's contribution is twofold. First - and to the best of our knowledge - we are among the first to introduce RT Delphi to the IS community. To this end, the provision of systematic references in this research reduces method ambiguity and provides first blueprints for how to conduct RT Delphi studies for those in our community enticed by this variant of Delphi. In doing so, we contribute to the IS method repertoire by instituting a novel and efficient, yet emergent, alternative to the traditional Delphi method. Second, our synthesis of extant RT Delphi research highlights potential challenges for employing this method and how IS scholars interested in exploring and designing new means of querying experts' opinions can help addressing them. Thereby, we call for initiating a process of standardizing RT Delphi to be able to establish it as a valuable addition to the IS research methods portfolio.

The remainder of this paper is structured as follows. In section two we provide a brief outline of the conventional Delphi method and compare it to its real-time extension. We also introduce our understanding of survey instrument designs and feedback mechanisms. Section three provides a detailed description of our research approach, whereas section four presents our main results in form of a synthesis of survey instrument designs and feedback mechanisms in extant RT Delphi research. We discuss our results regarding how our research can inform instrumenting RT Delphi for IS research and limitations of our research in section five. Finally, section six provides a brief conclusion of our paper.

2 Theoretical Background

2.1 The Conventional Delphi Method and the Real-Time Delphi Approach

Delphi is a qualitative research method for querying experts' opinions on future trends and developments (Landeta, 2006). It involves a structured and iterative process of collecting and consolidating expert opinions on a certain topic as well as providing moderated feedback about the consolidated opinions to the panel (Bardecki, 1984). The origins of Delphi trace back to the late 1940s where the method was developed at the RAND Corporation to gather experts' opinions about a specific problem (Dalkey and Helmer, 1963). Key characteristics of Delphi are: (1) *Anonymity* – Participants make responses privately without feeling pressured by dominant individuals; (2) *Iteration* – Allowing participants to change their opinions; (3) *Controlled feedback* – Presenting feedback in the form of actual arguments or simple statistical summaries of the groups response between iterations, thus allowing all members to input into the process and not just the most vocal ones; and (4) *Statistical group response* – At the end of the procedure, group judgement is (often) expressed as a median in which the extent of the participants' opinion spread may be used as an indication of the degree of participants' agreement.

Following Delphi's first application in a social context (Dalkey and Helmer, 1963) and Linstone and Turoff's (1975) broadly acknowledged proposition to use Delphi within the social sciences, Delphi is presumably today one of the best-known qualitative research methods by citation and name (Landeta, 2006). Moreover, Delphi has recently gained momentum in the IS community (Skinner et al., 2015) and is increasingly being utilized for qualitative research in IS (e.g., Chiravuri et al., 2011, Kasiri et al., 2012, Piccinini et al., 2015, Schmidt et al., 2001, Thiebes et al., 2017). Especially so-called ranking-type Delphi studies, which involve the ranking of collected opinions (Schmidt, 1997), are frequently employed in IS research. Schmidt et al. (2001), for instance, employ a ranking-type Delphi approach to elicit and rank software project risks with experts from Hong Kong, Finland, and the United States, whereas Thiebes et al. (2017) also employ a ranking type Delphi study to identify and rank motivating and discouraging factors in sharing individual genomic data. Regardless of some technical variations such as online questionnaire-based Delphi, panel attrition and dropout rates are typical drawbacks of Delphi, increasing with each additional iteration in the opinion and justification cycle (Bardecki, 1984). Delbecq et al. (1975) once estimated that the average Delphi study could take 45 days to 5 months, given that all participants reside in the same country with the use of the postal system. Notwithstanding that digital messaging systems have decreased those numbers, oftentimes researchers still limit studies to no more than two rounds to minimize fatigue or the incentive to adjust opinion evaluations towards the mean positions just to complete the study (Dransfeld et al., 2000).

Confronted with the above limitations of Delphi, Gordon and Pease (2006) introduced the concept of RT Delphi by reforming the conventional iteration cycle, as depicted in Figure 1. To this end, although RT Delphi inherits all characteristics and applications of the conventional method, it exhibits two main differences compared to Delphi and its technical variations. First, RT Delphi addresses inefficiencies of Delphi by establishing an asynchronous, almost real-time, information technology-aided iteration paradigm. This allows for participants to simultaneously engage in the opinion and justification cycle (i.e., giving their own opinions and reacting to other experts' opinions; right) without having to wait for a human moderator to evaluate and redistribute experts' assessments for each iteration (left). Second, the moderator, conventionally coupled to the time-consuming task of evaluating and redistributing experts' assessments, is freed via a platform that requires modern information technologies and

innovative automated or semi-automated means to evaluate experts' assessments and provide feedback to them (e.g., natural language processing).

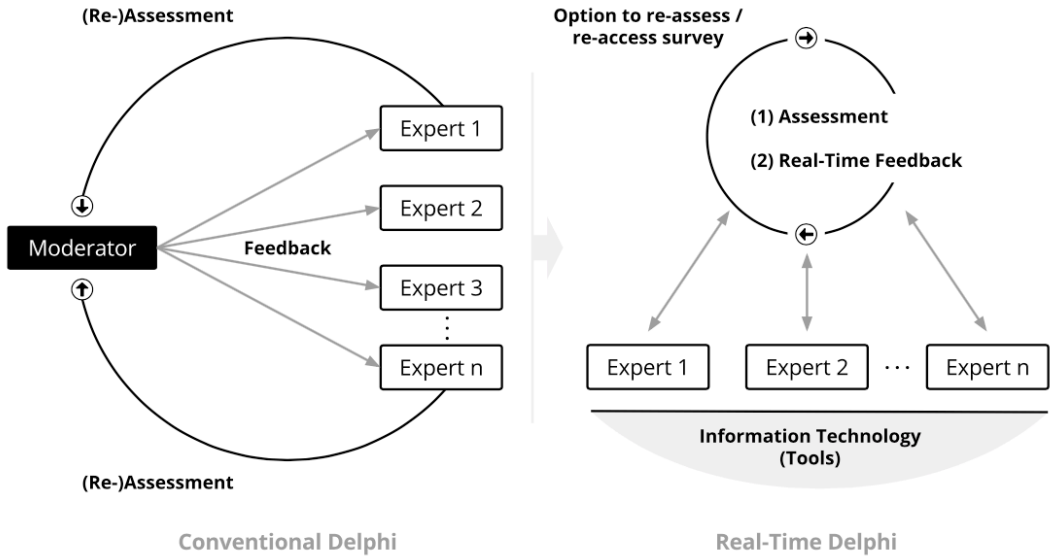

Figure 1. Processes of Conventional and RT Delphi (adapted from Gnatzy et al. 2011).

Together, these two novel characteristics of RT Delphi increase convenience as well as efficiency (Gordon and Pease, 2006) and are hence especially enticing for the IS community, where surveyed experts such as CIOs are often chronically short on time. Although RT Delphi neither touches nor diverges from Delphi in a study's overall goal and validity but merely requires real-time enabled survey instruments and feedback mechanisms for adoption (Gnatzy et al., 2011), its utilization within the IS community still seems to be a distant prospect right now.

2.2 Survey Instrument Designs and Feedback Mechanisms

While the Delphi method allows for the adoption of any form of qualitative and quantitative inquiry, usually along with a written justification (Hasson et al., 2000), Delphi studies commonly employ some form of survey questionnaires. With the use of RT Delphi, researchers implicitly commit to the use of an information technology platform (although the use of information technology does not automatically make a Delphi study a RT Delphi study) that extends the means to display the survey, its questions, and the gathering of input using multimedia possibilities (i.e., the survey instrument). In the context of RT Delphi and the implementation of the asynchronous iteration paradigm, the survey instrument is the initial starting point for a first-time evaluation by a study's participants. It consists of various elements that together can be viewed as *static assets that guide* participants' intents along each iteration of a question item. This working definition assumes that every RT Delphi study must incorporate four survey instrument design decisions: (1) *Survey structure* - The decision of how to display the survey (e.g., are questions displayed separately or grouped); (2) *Question type* - How to state questions (e.g., directly or as a scenario); (3) *Opinion format* - Participants' inputs to questions (e.g., numerical or natural language); and (4) *Justification method* - Means for inquiring participants for opinion justification.

Next to the survey instrument design, feedback plays a vital role in Delphi and is crucial for the survey's success (Rowe and Wright, 1999, Rowe et al., 2005). Consequently, feedback mechanisms take effect after an initial opinion has been given. We thus define them as *dynamic assets that stimulate* participants' intents along each iteration of a question item. Previous research has analysed the importance of feedback and iteration for changes in expert estimates (Best, 1974). Falling back to RT Delphi, effective feedback needs to be dispatched in real-time through the inherent information tech-

nology platform to ensure that participants' individual opinion assessments converge towards an objectified group result or cluster around alternative viewpoints, eventually signalling opinion stability to the participants (Gnatzy et al., 2011). Analog to the survey instrument design, feedback mechanisms are also subject to a wide array of multimedia possibilities, potentially capable of varying a survey's effectiveness (Couper et al., 2001, Gnatzy et al., 2011). Hence, each feedback mechanism is required to trigger an actionable response by including any of the following four features: (1) *Signal* - Display or use of information to urge a participant to re-assess an opinion (e.g., a red or green background); (2) *Convergence measure* - The way of estimating a participant's input convergence to the group (e.g., the statistical median); (3) *Convergence display* - The way of displaying convergence to the participant's understanding (e.g., numerical or graphical boxplot); and (4) *Justification access* - The prominence and access of opinion justifications by other participants (e.g., access via another window).

Due to the growing reference and implementation uncertainties that result from the growing pool of hard-to-distinguish and often overlapping 'Delphi techniques' that employ a wide variety of survey instrument designs and feedback mechanisms, this paper synthesizes survey instrument designs and feedback mechanisms for IS researchers seeking to employ RT Delphi in their research.

3 Research Approach

3.1 Literature Review

To identify applications of RT Delphi in extant literature that might serve as reference for our community, we conducted a three-staged literature review that was oriented towards Boell and Cecez-Kecmanovic (2014). An overview of the literature review is given in Figure 2.

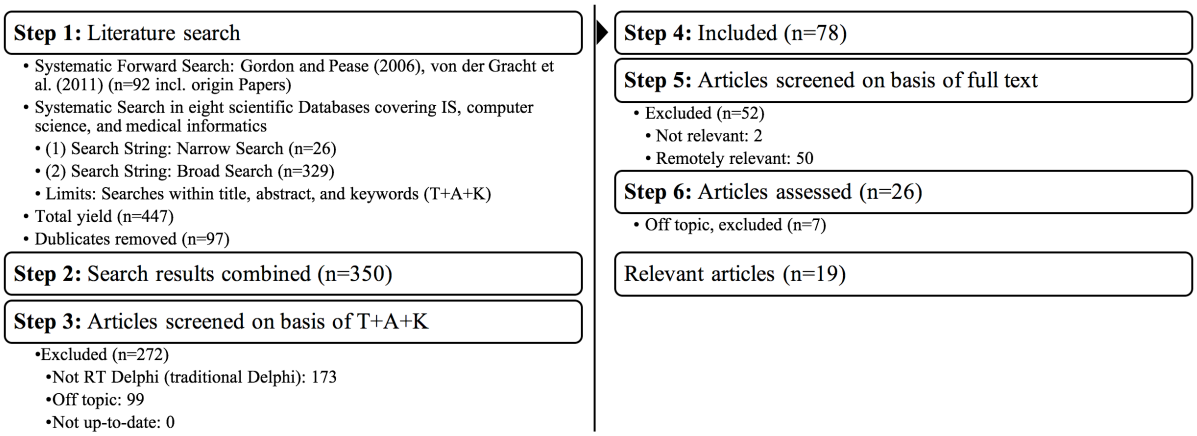

Figure 2. Flow Diagram of Inclusion/Exclusion and Literature Analysis.

Following the hermeneutic literature review approach (Boell and Cecez-Kecmanovic, 2014), the search and acquisition circle was first initiated with a forward search via the Web of Science, beginning from the RT Delphi research origin established by Gordon and Pease (2006), and the first substantial validation of the technique by Gnatzy et al. (2011). The forward search yielded 90 results of which 31 unique publications could be retrieved. After a first analysis and interpretation by screening publications based on title, abstract, keywords, and full texts (if necessary), the research scope was refined to identify additional research articles addressing or implementing RT Delphi. Scientific literature databases covering a wide range of journals and conferences in the domains of information systems, social sciences, and healthcare were searched. Included databases were ACM Digital Library, AIS Electronic Library, EBSCO (Business Source and Academic Search Complete), IEEE Xplore Digital Library, Proquest, ScienceDirect, and Web of Science. We specifically included databases covering literature from the healthcare domain since our forward search indicated a wide diffusion of (RT) Delphi in this domain. Moreover, several researchers referred to RT Delphi as "roundless",

which is why we also included this term in our search strings. Building on the gained terminology, a (1) narrow search string was constructed for a search in title, keywords, and abstracts:

> TIKEAB:("*real-time delphi*" OR "*real time delphi*" OR "*realtime delphi*" OR "*rt delphi*" OR "*round-less delphi*" OR "*round less delphi*" OR "*roundless delphi*")

The narrow search yielded 26 results of which 17 were unique publications. After concluding a second analysis and interpretation with the extended literature set, the research scope was refined one more time to identify further research articles that were not entirely limited to RT Delphi. Incorporating varying terminology but identical search scope, a (2) broad search string was constructed and applied:

> TIKEAB: (*delphi*) AND (TIKEAB: (*real-time* OR "*real time*" OR *realtime* OR *round-less* OR "*round less*" OR *roundless*) AND ABSTR:(*method* OR *stud** OR *surve** OR *panel* OR *expert** OR *online* OR *interactive*))

The broad search yielded 329 results of which 300 were unique publications. Where possible, we limited our search to peer-reviewed publications published in 2006 or later in each iteration, since the concept of RT Delphi was first introduced in Gordon and Pease (2006). After three stages, (1) the forward search (n=92 incl. the origin publication), the (2) narrow search (n=26), and the (3) broad search (n=329) resulted in 447 publications of which, after the removal of duplicates (n=97), 350 combined results remained for further analysis.

3.2 Analysis

Publications were evaluated for their use as a RT Delphi implementation reference in a two-step screening process. First, by referring to title, abstract, and keywords, each publication was matched to one of eight article clusters (cf. Table 1). Clusters reflect an article's value as implementation reference from 1 (highest) to 8 (lowest) and were developed iteratively by two researchers while coding the literature. Matching was also performed by two researchers, independent of each other. To evaluate the matching outcomes, we employed Janson's and Olsson's ι, a multivariate extension of Cohen's κ for multiple judges on the same (nominal) scale (Janson and Olsson, 2001). With a score of ι=0.569 the agreement between the two raters was moderate (Landis and Koch, 1977). Differences in the matching outcomes were resolved by discussion with a third researcher.

Cluster	Description
(1) Contribution to RT-Delphi	Articles that directly contribute to the methodological advancement of real-time Delphi.
(2) Validation of RT-Delphi	Articles that compare, assess or validate real-time Delphi stand-alone or compare it to other / conventional methods.
(3) Study using RT-Delphi	Studies that apply real-time Delphi in research.
(4) Discussion of study results	Articles that (remotely) discuss results of real-time Delphi studies and potentially lead to further articles that apply the real-time method.
(5) Related to RT-Delphi	Studies that use some variation of the traditional Delphi method but have some similarities to real-time Delphi. Such as immediate feedback or the way participants are surveyed.
(6) Not sure if real-time was used	Studies where it is not clear if a real-time concept was applied.
(7) Traditional Delphi	Studies that apply the conventional Delphi method.
(8) Off Topic	Discusses results of non-real-time studies; Not related to Delphi; Other

Table 1. Literature Review Article Cluster.

Ineligible articles were excluded by applying predefined exclusion criteria. Accordingly, all articles that were published before 2006 (not up-to-date), off topic (cluster 8) or applied or discussed the conventional Delphi method (cluster 7) were excluded and not further assessed. After the literature screening, 173 publications concerning the conventional Delphi method and 99 unrelated publications were excluded (n=272), leaving 78 publications, corresponding to article clusters 1 to 6, for further assessment. In the second screening, all eligible articles were read in the order of title, keywords, abstract, and full text to assess their final relevance for this study. Here, the relevance of each article was determined by one of three labels, corresponding to its semantic trigger: *(1) Not Relevant* – Mentioning of rounds without RT Delphi; No information technology platform used or use of email; RT Delphi only discussed as further research; Arbitrary Delphi technique unrelated to RT Delphi; *(2) Remotely Relevant* – Mentioning of rounds in the context of RT Delphi (not true real-time); Participants cannot revisit their opinion at any time; RT Delphi was mixed or combined with other methods; A RT Delphi study was conducted, but not methodologically mentioned/described; RT Delphi was solely assessed for a specific purpose (e.g., future studies); and *(3) Relevant* – Real-time application directly related to RT Delphi; Comparison of RT Delphi to other methods; Methodological evaluations; First time experiences with RT Delphi; Use of information technology to evaluate panel assessments in real-time; Novel introduction of a specific survey or feedback element.

Articles were conservatively matched into lower labels where uncertainty, from the perspective of appropriate implementation reference, prevailed. As a result, two publications were deemed not relevant, 50 remotely relevant and 26 relevant. While remotely relevant publications could also lead to some general insights about RT Delphi, this work continues to focus on the 26 articles estimated to be relevant; leading to the exclusion of remotely and not relevant publications (n=52). In a final full text assessment, the remaining articles were analysed in detail for information about real-time enabled survey instrument design decisions or features of feedback mechanisms. During this process, seven additional publications that did not have sufficient information about either aspect, or did not relate enough to the RT Delphi technique, which was only eminent through a detailed investigation of the works methodology, were excluded. Hence, 19 articles provided the foundation for further analysis in this work.

3.3 Descriptive Results

Table 2 illustrates the multi-stage process, along the literature research and analysis, that yielded 19 publications from an original literature sample of 350 publications.

Assessed Cluster		Count (All)	Not Relevant	Remotely Relevant	Relevant	Incl.
1	Contribution to RT Delphi	4	0	0	4	4
2	Validation of RT Delphi	7	0	0	7	7
3	Study using RT Delphi	18	0	11	7	4
4	Discussion of RT Delphi study results	1	0	1	0	0
5	Related to RT Delphi	31	2	23	6	4
6	Not sure if real-time was used	17	0	15	2	0
Sub-Total		**78**	**2**	**50**	**26**	**19**
7	Traditional Delphi	173	173	Not included in second step analysis and assessment.		
8	Off Topic	99	99			
Total		**350**	**274**	**50**	**26**	**19**

Table 2. Overview of Relevance Clustering Results.

As indicated, a large part of the final literature set was derived from research directly concerned with RT Delphi (cluster 1 and 2). Among those, the prominent origin research of RT Delphi by Gordon and Pease (2006), including an additional work by Gordon (2007), and the validation work by Gnatzy et al. (2011). With a combined eight publications (cluster 3 and 5) the pool of academic works that pure-

ly applied RT Delphi is relative small. Notwithstanding that the literature set could be extended by 50 remotely relevant publications, this can be viewed as a confirmation of our assertion of a low penetration for RT Delphi at the beginning of our paper. Additionally, 12 of 19 articles referenced feedback mechanisms, while all 19 articles included some aspect of the survey instrument design.

4 Real-Time Delphi Survey Instrument Design and Feedback Mechanism Aspects

4.1 Survey Instrument Design Characteristics

Analysis of included publications regarding implemented survey instrument designs revealed a total of 11 characteristics for the four-survey instrument design decision dimensions (i.e., Survey Structure, Question Type, Opinion Format, and Justification Method). An overview of each dimension, its characteristics, and related publications is given in Table 3 in Appendix A.

Survey Structure. The survey structure of RT Delphi studies can be grouped into matrix structures and thematic blocks. Along Gordon and Pease's (2006) first introduction of RT Delphi, they also differentiated between a 1D RT Delphi and a 2D RT Delphi. The difference between the two being a matrix like alignment of questions and scenarios in a cross impact or input/output like structure. This structure is especially useful for decision models or input/output validations by participants. To facilitate the navigation of participants within a survey, thematic blocks can be used to group questions of similar kinds together into one display (Abadie et al., 2010). Those questions often follow an underlying, open stated thesis that guides the context of all the questions within that group. For example, questions regarding drug use and drug certifications might be split into two separate thematic blocks but with comparatively similar questions.

Question Type. Question types in RT Delphi studies can be divided into direct estimates, open questions, spatial locations, and statements. Direct estimates require participants to provide an estimate for a currently unknown fact or circumstance based on their expertise in the area of interest (Gordon, 2007). For example, experts might be asked to provide an estimated date for when the world will be completely independent from fossil fuels. Engaging participants in open questions, on the other hand, could be a way to find other, supplemental domains for a given problem (Sunderji and Waddell, 2015). Another potential question type for RT Delphi studies is spatial locations (Di Zio and Pacinelli, 2011). For this type, participants are asked for a location assessment on a two-dimensional map (graphical input). Statement type questions provide participants with specific statements, or projections, about the future or a pre-specified scenario (Abadie et al., 2010). Participants are then required to assess the importance of each statement (or scenario) under different conditions (e.g., economical, scientific, societal, quality of life). Furthermore, this question type also includes the assessment of possible actions that could speed up or strengthen the realization of the presented statement.

Opinion Format. Opinion formats were found to be either of type category, expected probability, or Likert scale. The category opinion format presents participants a list of predefined options they can choose from to reflect their opinions (e.g., highly likely, likely, unlikely, and highly unlikely) (Lintonen et al., 2014). Expected probability collects a participant's belief of some state occurrence in a direct, numerical fashion (e.g., on a scale from 1 to 100) (Gary and von der Gracht, 2015). Likert scales are probably the most frequently used opinion format among the surveyed literature. They are used to aid participants in evaluating believe, occurrence, impact, desirability, usefulness, and other, similar constructs (Keller et al., 2012). However, depending on the number of participants it can be challenging to apply proper convergence metrics.

Justification Method. Approaches to provide justification for one's assessments can be broadly divided into free justifications and guided justifications. Free justifications allow for participants to freely comment on any opinion without any restriction or further guidance (Gordon and Pease, 2006). Instead of having participants justify their opinion freely, they can also be guided by justifying a cer-

tain direction of their opinion (e.g., along a question for a numeric probability, justifications can only be given for why the value should be low or high) (Geist, 2010).

4.2 Feedback Mechanisms

Adding to the 11 potential survey instrument design characteristics, we also identified 10 potential characteristics of feedback mechanisms for RT Delphi studies. An overview of the four feature dimensions (i.e., Signal, Convergence Measure, Convergence Display, and Justification Access), their characteristics, and related publications is given in Table 4 in Appendix A.

Signal. Signals display some form of information to urge participants to reassess their opinions. Among reviewed publications, signals were either of type colour coded or of type flagged. Colour code signals show different background colours for question items to highlight a divergent condition (e.g., red), a convergence achieved condition (e.g., grey), or an unacceptable convergence (e.g., white) (Gnatzy et al., 2011). Flagged signals, on the other hand, mark questions (e.g., with icons or a message) when a participant's opinion is beyond a pre-set distance from the group response (Gordon, 2007).

Convergence Measure. Similar to the conventional Delphi method, many proposals exist for determining the convergence of a group. Convergence measures can be algorithmic, basic statistics, or distances. Algorithmic convergence measures use some form of algorithm to determine group convergence or divergence (Diamond et al., 2014). Popular algorithmic convergence measures include, for example, Cronbach's alpha, Kendall's W, or Cohen's κ. Other commonly used ways to measure convergence in RT Delphi studies are rudimentary, descriptive statistics such as mean, median, distance to the group average, or distribution (quantiles) (Gnatzy et al., 2011). Lastly, convergence can also be measured as a distance on a two-dimensional map (e.g., Euclidean distance) (MacEachren et al., 2006). This type of convergence could be particularly useful in conjunction with spatial location question types or to cluster groups of participants.

Convergence Display. Adding to the measurement of convergence, displaying convergence might differ from study to study and be based on the chosen convergence measure. In our examined literature, applied convergence displays were boxplots, numeric representations, or spatial relocation. For boxplots, convergence measurements of the group are graphically illustrated as a boxplot with interquartile ranges, marking the participant for visual orientation (Keller et al., 2012), whereas for numeric representations statistical convergence measures, such group averages, are directly provided to the participant, potentially requiring statistical knowledge (Gnatzy et al., 2011). Given input and convergence measurements relate to a two-dimensional space on a map, it is necessary to further zoom in for participants to be able to give more precise estimates. To do so, spatial relocation on a map can be used (Di Zio and Pacinelli, 2011). One can, for example, chose an opinion point close to the distance average of the group and enlarge the map such that 50% of participants' opinions are still visible. As a result, relocated participants feel urged to reassess for a new position on the map.

Justification Access. Finally, the prominence and access of opinion justifications by other participants has to be considered as well. We can distinguish between two basic methods for providing access to given justifications, namely related and threaded. When justifications are provided in a related fashion, each question item allows to review associated opinion justifications through other participants (Gordon, 2007). Threaded justification provisioning, on the other hand, allows participants to engage into other justifications inside of a forum-like structure (Dalal et al., 2011). Such a structure might also be disconnected from specific question items.

5 Discussion

In the following, we discuss how our review of survey instrument designs and feedback mechanisms may contribute to promoting the utilization and design of RT Delphi in the IS community. We also highlight two major insights from our review of literature on RT Delphi, namely the weak exploitation

of RT Delphi's multimedia capabilities in extant research and what real-time could mean for conducting RT Delphi studies in IS research.

5.1 Instrumenting Real-Time Delphi for IS Research

Despite RT Delphi promising to deliver improved efficiency and participant convenience by establishing an asynchronous, round-less, almost real-time iteration cycle (Gordon and Pease, 2006), it still seems to be largely ignored by IS researchers. This is highlighted by the fact that no papers in cluster 3 (i.e., studies using RT Delphi) were published in leading IS journals or conference proceedings. A potential reason for this can be found in the striking absence of best-practices and ready-to-use software tools for conducting RT Delphi studies. Thus, researchers in IS wishing to employ RT Delphi are constantly forced to reinvent the wheel by developing their own best-practices and software tools. Not only may this result in arbitrary modifications of the method, something conventional Delphi has been criticized for in the past (Paré et al., 2013), but also lead to high uncertainty among those interested in applying RT Delphi in their research as our community has a long history of debating rigour in our field (e.g., Benbasat and Weber, 1996).

An objective of this descriptive review was to address these uncertainty issues by reviewing and indexing extant research related to RT Delphi. Thereby, this research provides guidance to those in our community enticed by RT Delphi on several levels. First, for those IS researchers seeking to employ RT Delphi, our clusters of studies related to RT Delphi serve as an index of references. Interested researchers may reference papers indexed in this article as success stories of employing RT Delphi and build on the methodological knowledge they capture to justify their choice of method. Second, the synthesis that we present in this paper provides a structured catalogue of common survey instrument designs and feedback mechanisms in RT Delphi studies as well as concrete examples for how to design RT Delphi studies. Those seeking to employ RT Delphi in their own research may therefore use the overview of survey instrument designs and feedback mechanisms as a foundational toolbox to carefully design their own RT Delphi studies, while developers of software tools that support conducting this kind of study may use it as a source of software requirements. Consequently, our paper may serve as a starting point for standardization and the development of best practices for instrumenting RT Delphi in IS research. Third, we present RT Delphi as a promising object of study for scholars interested in design-oriented research by highlighting two insights from our literature review below.

5.2 Weak Multimedia Use in Past RT Delphi Studies

We defined survey instruments as a collection of static assets to guide and feedback mechanisms as dynamic assets to stimulate participants' intents. Despite the diverse possibilities to guide and stimulate participants' intents, RT Delphi did not yet reach its full potential due to the weak use of multimedia so far. As such, our analysis and presented results highlight that to date, conventional Delphi surveys are often simply transferred to a RT Delphi environment. Notable exceptions to this being the studies by Di Zio and Pacinelli (2011) and MacEachren et al. (2006), which make use of interactive maps. As a research discipline concerned with socio-technical systems, we have to acknowledge the vast possibilities of RT Delphi despite simple questionnaire-type surveys. Moreover, as a discipline that has the design of information systems at its heart (Hevner et al., 2004), we must acknowledge the potential that well-designed RT Delphi tools may have for collecting expert judgements, not only for the IS community but also for other communities utilizing the Delphi method.

For researchers and practitioners in the IS domain, multimedia oriented survey instruments and feedback mechanisms bear applications that radically facilitate the complex interface between real situational data and projections at hand, and the inclusion of intrinsic expert knowledge in research and business environments. With the goal in mind to guide participants' intents in the best possible way, any multimedia-supported form of inquiry should be measured by the degree of its resemblances to participant-known interaction modes. Giving way to survey instruments that resemble or simulate known environments (e.g., health reporting dashboards, financial status indicators, or augmented video material of a sales processes) likely yields an increase in upfront willingness to join expert panels (or

simulations). Not only does this improve participants' convenience, but also strengthen survey impact in form of more robust results through higher participant engagement. Consequently, RT Delphi's value for IS research and practice increases dramatically and beyond simply transferring conventional Delphi surveys to a real-time environment. RT Delphi also allows for the introduction of a new breed of survey instruments. Video, interactive graphs and also virtual reality assessments are possible. Airfield control agents could, for instance, witness a plane manoeuvre in a virtual environment and be asked about risk assessments along different points in time. Due to the nature of virtual environments, the experts' opinions could be queried by drawing in a three-dimensional space to indicate high, medium or low risk areas. Those inputs could be further analysed within a feedback mechanism that highlights disagreements in the virtual environment, again guiding further iterations. This, however, also emphasizes one of the identified issues of RT Delphi's feedback mechanisms. The decision of when to query participants for further input in a real-time environment is a fundamental aspect of RT Delphi and essential for protecting participants from an overwhelming experience. Unlike human moderators with a potentially strong understanding of key aspects of the study, machines still have trouble to understand when to push on certain disagreements and when to let others go. Preventing an inflow of countless 'further input required' notifications, the missing moderator must be compensated for by clever use of information technology.

In consequence of the above considerations, we call for more design-driven research on RT Delphi. IS researchers interested in design should investigate how to best leverage the asynchronous and automated opinion-feedback cycle of RT Delphi, enabled by novel technologies such as artificial intelligence or natural language processing, and rich multimedia environments.

5.3 Real-Time and Isomorphism of Survey Instruments and Feedback Mechanisms

Real-time capabilities are the key ingredient for RT Delphi's efficiency benefit over Delphi. Throughout our study, we were confronted with contradicting declarations of real-time and round-less. To become real-time, RT Delphi decouples the moderator usually responsible for redistribution of participants' assessments from the iteration cycle. However, herein also lies a danger for the overall validity of the technique, which, except for Gnatzy et al. (2011), was not discussed by the selected literature. Given that participants can simultaneously engage in the opinion and justification cycle, a danger exists that two interrelated Delphi characteristics are inherently violated. Namely (1) controlled feedback and (2) anonymity. The primary reason for controlled feedback is to bridge knowledge gaps between participants of different expertise (Dalkey and Helmer, 1963, Rowe and Wright, 1999). However, it also enables moderators to remove or obfuscate obvious relations between a participant and her assessments or justifications. In the current real-time environment, there seems to be a large vector of influence on other participants by way of using excessive prose or by giving away leads about one's identity.

In the current RT Delphi literature, real-time is too weakly defined to propose an immediate mitigation. Under the assumption that real-time should be participant subjective (i.e., crafting a feeling of immediacy), the investigation of natural language algorithms used to check and relay justifications by participants seems to be a worthwhile pursuit that so far has not yet been done. This is especially true for the inherently sequential ranking-type Delphi studies, which is the most popular type of Delphi in IS research. Anticipating further downstream extension of real-time concepts from the participant onward to concrete survey elements, an isomorphic structure of survey instruments and feedback mechanisms can be expected. Due to the vast possibilities on the digital scope, it is likely that classical survey constructs will no longer hold for RT Delphi and that future applications of RT Delphi will use the same assets to both guide and stimulate intent. For example, if in the distant future, the (digital) moderator would talk directly to a participant, both, asking for opinions while signalling disagreement, that would invalidate the distinct relationship between survey instruments and feedback mechanisms for stating an opinion.

Considering the ambiguity of the real-time concept, one implication derived from our research is the notion of perceived real-time. Although we assume that real-time systems will increase in significance, the real-time transformation is a gradual process that involves the implementation of asynchronous aspects on several stages of the process hierarchy. Being wary of terminology such as 'roundless' can be a fruitful insight into evaluating the benefit of real-time and complexity in study designs. As a result, we should strive for a clear conceptualization of real-time for our community.

5.4 Limitations and Future Research

While focusing our research on guiding fellow IS researchers in conducting RT Delphi studies and designing novel RT Delphi tools, our research is not without limitations. First, over time, numerous variations to the conventional Delphi method have been proposed, and for sake of novelty, often been renamed. We limited our research to 'Delphi' methods explicitly, potentially ignoring interesting, but remotely related designs. To mitigate this effect, we decided to pre-group our raw results into eight clusters to maximize the discriminatory power in relevance for each publication segment. Despite this approach, our pair-review yielded a moderate divergence (ι=0.569) that can indicate both, a high complexity for the task at hand or a lack of discrete clusters. Our results are based on a subset of relevant literature (n=19), which could be extended to a broader set of remotely relevant literature (n=50). Although we feel confident that a more in depth analysis of additional remotely related literature would lead to little gains in our list of survey instrument designs and feedback mechanisms, we cannot be sure that more interesting insights are yet to be discovered in the often, just briefly described, methodology sections of other publications. Adding to this, information on employed survey instrument designs and feedback mechanisms in individual publications was often scarce. Especially for those studies that reported on the use of RT Delphi (cluster 3).

Our research on RT Delphi serves as a starting point for understanding the underlying concepts of an RT Delphi study by the IS community. From here, various routes of research can be pursued, of which the following three seem most promising. To aid in the diffusion of RT Delphi in our community, a domain specific application of the method can be both, a useful validation and an example reference of the concept. Furthermore, the weak use of multimedia within the RT Delphi studies calls for the investigation of new, engaging survey instrument designs and feedback mechanisms by way of developing novel design propositions to inquiry expert responses. Lastly, the concept of real-time brings methodological challenges only solvable via appropriate automation. With the elimination of the moderator in RT Delphi, a key aspect of control has been sacrificed to an asynchronous workflow and more convenience. Further investigation of how intelligent systems could fill in the role of a moderator to protect the Delphi principles in real-time is therefore needed.

6 Conclusion

Based on the provided analysis, researchers and practitioners need to recognize that RT Delphi in IS is still in a state of emergence despite its introduction more than a decade ago. While RT Delphi carries known, obvious benefits compared to Delphi, more than prototypical applications will not be adopted in IS without leading IS journals' displaying empathy for this kind of study and more importantly adequate guidance for interested researchers. Therefore, our ultimate goal should be to shift our focus from the mechanics of conventional Delphi to investigating the possibilities of RT Delphi by designing, developing, and evaluating tools and guidelines capable of fully embracing the real-time nature of the technique. Considering this situation and the limited diffusion of RT Delphi within the IS community, the main contribution of this research is a rigorous index of reference material concerned with RT Delphi, serving as a starting point to further facilitate RT Delphi adoption in IS. In addition, various elements of real-time enabled survey instrument designs and feedback mechanisms, which serve as a foundation for any RT Delphi study so far, have been defined for the first time. More in-depth research should be conducted beyond this first outline to attain a first state of standardization. Our work is a first step towards this direction by establishing a foundation for design-driven research on RT Delphi on which future research can build to establish RT Delphi as a valuable IS research method.

Appendix A - Survey Instrument Design Aspects and Feedback Mechanism Aspects

Survey Instrument Design Aspect / **Source**	**Survey Structure**		**Question Type**				**Opinion Format**			**Justification Method**	
	MS	TB	DE	OQ	SL	S	C	EP	LS	FJ	GJ
Abadie et al. (2010)		✓				✓					
Dalal et al. (2011)		✓				✓			✓	✓	✓
Di Zio and Pacinelli (2011)			✓		✓					✓	
Diamond et al. (2014)	✓					✓				✓	
Gallego and Bueno (2014)						✓			✓		✓
Gary and von der Gracht (2015)		✓				✓		✓	✓	✓	
Geist (2010)						✓			✓	✓	✓
Gnatzy et al. (2011)		✓									
Gordon (2007)	✓		✓			✓	✓	✓		✓	
Gordon and Pease (2006)	✓					✓	✓			✓	
Hsieh et al. (2011)	✓										
Keller and von der Gracht (2014)						✓		✓	✓	✓	✓
Keller et al. (2012)								✓	✓	✓	✓
Kuklinski et al. (2014)	✓					✓		✓	✓		✓
Lintonen et al. (2014)			✓			✓	✓			✓	
MacEachren et al. (2006)			✓		✓					✓	
Spickermann et al. (2014)	✓					✓		✓	✓	✓	✓
Sunderji and Waddell (2015)				✓					✓	✓	✓
Warth et al. (2013)		✓				✓		✓	✓	✓	

MS = Matrix Structure; TB = Thematic Blocks; DE = Direct Estimates; OQ = Open Questions; SL = Spatial Locations; S = Statements; C = Category; EP = Expected Probability; LS = Likert Scale; FJ = Free Justification; GJ = Guided Justification

Table 3. Compiled List of Survey Instrument Design Aspects.

Feedback Mechanism Aspect / **Source**	**Signal**		**Convergence Measure**			**Convergence Display**			**Justification Access**	
	CC	F	A	BS	D	B	N	SR	R	T
Dalal et al. (2011)										✓
Di Zio and Pacinelli (2011)					✓			✓		
Diamond et al. (2014)			✓							
Gary and von der Gracht (2015)				✓		✓	✓		✓	
Gnatzy et al. (2011)	✓			✓		✓	✓		✓	
Gordon (2007)		✓		✓			✓		✓	
Gordon and Pease (2006)				✓			✓		✓	
Hsieh et al. (2011)	✓			✓			✓		✓	
Keller and von der Gracht (2014)						✓			✓	
Keller et al. (2012)						✓				
MacEachren et al. (2006)					✓					✓
Spickermann et al. (2014)			✓							

CC = Colour Code; F = Flagged (red icon or message); A = Algorithmic; BS = Basic Statistic; D = Distance; B = Boxplot; N = Numeric; SR = Spatial Relocation; R = Related; T = Threaded

Table 4. Compiled List of Feedback Mechanism Aspects.

References

Abadie, F., Friedewald, M. and Weber, K. M. (2010). "Adaptive foresight in the creative content industries: anticipating value chain transformations and need for policy action." *Science and Public Policy* 37 (1)**,** 19-30.

Bardecki, M. J. (1984). "Participants' response to the Delphi method: An attitudinal perspective." *Technological Forecasting and Social Change* 25 (3)**,** 281-292.

Benbasat, I. and Weber, R. (1996). "Research commentary: Rethinking "diversity" in information systems research." *Information systems research* 7 (4)**,** 389-399.

Best, R. J. (1974). "An experiment in Delphi estimation in marketing decision making." *Journal of Marketing Research* 11 (4)**,** 448-452.

Boell, S. K. and Cecez-Kecmanovic, D. (2014). "A hermeneutic approach for conducting literature reviews and literature searches." *Communications of the AIS* 34 (12)**,** 257-286.

Chiravuri, A., Nazareth, D. and Ramamurthy, K. (2011). "Cognitive conflict and consensus generation in virtual teams during knowledge capture: Comparative effectiveness of techniques." *Journal of Management Information Systems* 28 (1)**,** 311-350.

Couper, M. P., Traugott, M. W. and Lamias, M. J. (2001). "Web survey design and administration." *Public Opinion Quarterly* 65 (2)**,** 230-253.

Dalal, S., Khodyakov, D., Srinivasan, R., Straus, S. and Adams, J. (2011). "ExpertLens: A system for eliciting opinions from a large pool of non-collocated experts with diverse knowledge." *Technological Forecasting and Social Change* 78 (8)**,** 1426-1444.

Dalkey, N. and Helmer, O. (1963). "An experimental application of the Delphi method to the use of experts." *Management Science* 9 (3)**,** 458-467.

Delbecq, A. L., Van De Ven, A. H. and Gustafson, D. H. (1975). *Group techniques for program planning: A guide to nominal group and Delphi processes.*

Di Zio, S. and Pacinelli, A. (2011). "Opinion convergence in location: A spatial version of the Delphi method." *Technological Forecasting and Social Change* 78 (9)**,** 1565-1578.

Diamond, I. R., Grant, R. C., Feldman, B. M., Pencharz, P. B., Ling, S. C., Moore, A. M. and Wales, P. W. (2014). "Defining consensus: a systematic review recommends methodologic criteria for reporting of Delphi studies." *Journal of Clinical Epidemiology* 67 (4)**,** 401-409.

Dransfeld, H., Pemberton, J. and Jacobs, G. (2000). "Quantifying weighted expert opinion: the future of interactive television and retailing." *Technological Forecasting and Social Change* 63 (1)**,** 81-90.

Gallego, D. and Bueno, S. (2014). "Exploring the application of the Delphi method as a forecasting tool in Information Systems and Technologies research." *Technology Analysis & Strategic Management* 26 (9)**,** 987-999.

Gary, J. E. and Von Der Gracht, H. A. (2015). "The future of foresight professionals: Results from a global Delphi study." *Futures* 71**,** 132-145.

Geist, M. R. (2010). "Using the Delphi method to engage stakeholders: A comparison of two studies." *Evaluation and Program Planning* 33 (2)**,** 147-154.

Gnatzy, T., Warth, J., Von Der Gracht, H. and 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.

Gordon, T. and Pease, A. (2006). "RT Delphi: An efficient,"round-less" almost real time Delphi method." *Technological Forecasting and Social Change* 73 (4)**,** 321-333.

Gordon, T. J. (2007). "Energy forecasts using a "Roundless" approach to running a Delphi study." *Foresight* 9 (2)**,** 27-35.

Hasson, F. and Keeney, S. (2011). "Enhancing rigour in the Delphi technique research." *Technological Forecasting and Social Change* 78 (9)**,** 1695-1704.

Hasson, F., Keeney, S. and Mckenna, H. (2000). "Research guidelines for the Delphi survey technique." *Journal of Advanced Nursing* 32 (4)**,** 1008-1015.

Hevner, A. R., March, S. T., Park, J. and Ram, S. (2004). "Design science in information systems research." *MIS Quarterly* 28 (1)**,** 75-105.

Hsieh, C.-H., Tzeng, F.-M., Wu, C.-G., Kao, J.-S. and Lai, Y.-Y. (2011). "The comparison of online Delphi and real-time Delphi" In: *Proceedings of Technology Management in the Energy Smart World (PICMET), 2011.* Ed. by IEEE. pp. 1-3.

Janson, H. and Olsson, U. (2001). "A measure of agreement for interval or nominal multivariate observations." *Educational and Psychological Measurement* 61 (2)**,** 277-289.

Kasiri, N., Sharda, R. and Hardgrave, B. (2012). "A balanced scorecard for item-level RFID in the retail sector: a Delphi study." *European Journal of Information Systems* 21 (3)**,** 255-267.

Keller, J. and Von Der Gracht, H. A. (2014). "The influence of information and communication technology (ICT) on future foresight processes—Results from a Delphi survey." *Technological Forecasting and Social Change* 85**,** 81-92.

Keller, J., Von Der Gracht, H. A., Kroehl, R. and Markmann, C. (2012). "ICT-Tools in Foresight-A Delphi Study about Future Developments" In: *ISPIM Conference Proceedings.* Ed. by The International Society for Professional Innovation Management (ISPIM). pp. 1.

Kuklinski, C., Moser, R. and Callarman, T. (2014). "Managing dynamic business environments: India's future automotive industry." *Journal of Indian Business Research* 6 (4)**,** 309-331.

Landeta, J. (2006). "Current validity of the Delphi method in social sciences." *Technological Forecasting and Social Change* 73 (5)**,** 467-482.

Landis, J. R. and Koch, G. G. (1977). "The measurement of observer agreement for categorical data." *Biometrics* 33 (1)**,** 159-174.

Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J. and Gutmann, M. (2009). "Life in the network: the coming age of computational social science." *Science* 323 (5915)**,** 721.

Linstone, H. A. and Turoff, M. (1975). *The Delphi method: Techniques and applications.*

Linstone, H. A. and Turoff, M. (2011). "Delphi: A brief look backward and forward." *Technological Forecasting and Social Change* 78 (9)**,** 1712-1719.

Lintonen, T., Konu, A., Rönkä, S. and Kotovirta, E. (2014). "Drugs foresight 2020: a Delphi expert panel study." *Substance Abuse Treatment, Prevention, and Policy* 9 (1)**,** 18.

Maceachren, A. M., Pike, W., Yu, C., Brewer, I., Gahegan, M., Weaver, S. D. and Yarnal, B. (2006). "Building a geocollaboratory: supporting Human–Environment Regional Observatory (HERO) collaborative science activities." *Computers, Environment and Urban Systems* 30 (2)**,** 201-225.

Paré, G., Cameron, A.-F., Poba-Nzaou, P. and Templier, M. (2013). "A systematic assessment of rigor in information systems ranking-type Delphi studies." *Information & Management* 50 (5)**,** 207-217.

Piccinini, E., Hanelt, A., Gregory, R. and Kolbe, L. (2015). "Transforming industrial business: the impact of digital transformation on automotive organizations" In: *Proceedings of the 36th International Conference on Information Systems.* Ed. by AIS. pp. 1-20.

Rowe, G. and Wright, G. (1999). "The Delphi technique as a forecasting tool: issues and analysis." *International Journal of Forecasting* 15 (4)**,** 353-375.

Rowe, G. and Wright, G. (2011). "The Delphi technique: Past, present, and future prospects—Introduction to the special issue." *Technological Forecasting and Social Change* 78 (9)**,** 1487-1490.

Rowe, G., Wright, G. and 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.

Schmidt, R. C. (1997). "Managing Delphi surveys using nonparametric statistical techniques." *Decision Sciences* 28 (3)**,** 763-774.

Schmidt, R., Lyytinen, K. and Mark Keil, P. C. (2001). "Identifying software project risks: An international Delphi study." *Journal of Management Information Systems* 17 (4)**,** 5-36.

Skinner, R., Nelson, R. R., Chin, W. W. and Land, L. (2015). "The Delphi Method Research Strategy in Studies of Information Systems." *Communications of the AIS* 37 (2)**,** 31-63.

Spickermann, A., Zimmermann, M. and Von Der Gracht, H. A. (2014). "Surface-and deep-level diversity in panel selection—Exploring diversity effects on response behaviour in foresight." *Technological Forecasting and Social Change* 85**,** 105-120.

Sunderji, N. and Waddell, A. (2015). "Using real-time Delphi to develop a consensus on competencies." *Medical Education* 49 (11)**,** 1151-1152.

Thiebes, S., Lyytinen, K. and Sunyaev, A. (2017). "Sharing is About Caring? Motivating and Discouraging Factors in Sharing Individual Genomic Data" In: *Proceedings of the 38th International Conference on Information Systems*. Ed. by AIS. pp. 1-17.

Warth, J., Heiko, A. and Darkow, I.-L. (2013). "A dissent-based approach for multi-stakeholder scenario development—The future of electric drive vehicles." *Technological Forecasting and Social Change* 80 (4)**,** 566-583.