
Ayman Alarabiat¹ and Isabel Ramos²
¹Amman College of Banking and Financial Sciences, Al-Balqa’ Applied University, As slat, Jordan
²ALGORITMI Centre, University of Minho, Portugal
ayman.arabiat@bau.edu.jo
iramos@dsi.uminho.pt
DOI: 10.34190/JBRM.17.2.04

Abstract: After almost 30 years of being used in the information system (IS) discipline, only a few studies have focused on how IS scholars apply the method’s guidelines to design Delphi studies. Thus, this paper focuses on the use of the Delphi method in IS research. To do so, articles published between 2004 and 2017 in the Senior IS Scholars’ collection of journals of the Association of Information Systems (AIS), describing Delphi studies, were analysed. Based on analysis of sixteen (16) retrieved IS studies, we concluded that IS researchers have applied the method’s most important phases and the procedural recommendations to promote rigor were considered in the majority of the analysed studies. Nonetheless, IS researchers still need to include detailed information about (1) the steps taken to ensure the validity of the achieved results, (2) better describe the process of selecting and recruiting the experts, and (3) experiment with innovative techniques to keep participants involved in the Delphi process.

Keywords: Delphi Method; research method; information system; literature review; qualitative research.

1. Introduction

In 1963 the Delphi method was released for open use by companies and in research (Dalkey and Helmer, 1963), after about 10 years of being applied exclusively, for security purposes, in military defense project (Dalkey and Helmer, 1951). Since then, the method has had a significant impact in research through its quite interesting contribution in producing and developing thousands of studies in a variety of research fields (Akins, Tolson and Cole, 2005; Gallego and Bueno, 2014; Gallego, Luna and Bueno, 2008; Heiko, 2012; Kaynak, Bloom and Leibold, 1994; Landeta, 2006; Linstone and Turoff, 1975; Okoli and Pawlowski, 2004; Pare et al., 2013; Powell, 2003)

In a broad definition, Delphi method is a structured, organized, and iterative process aiming to distill and to correlate opinions from a compositional group (panel) of individuals (usually experts) concerning a particular problem, topic or task. The Delphi method process typically implies subjecting the panelists to a series of intensive questionnaires, interspersed with controlled feedback, until some reliable consensus is reached (Dalkey and Helmer, 1963; Heiko, 2012; Hsu and Sandford, 2007b; Linstone and Turoff, 1975; Nevo and Chan, 2007; Okoli and Pawlowski, 2004; Schmidt, 1997). The method enables researchers to address issues that require eliciting the experience and intuition of experts (Rowe and Wright, 2011). Nowadays, despite the changing in the method attractiveness during the 1980s and 1990s (Heiko, 2012; Skinner et al., 2015), the method has been gaining progressively a wider recognition as one of the most stable and fruitful research methods used in research, (Gallego and Bueno, 2014; Hasson and Keeney, 2011; Heiko, 2012; Rowe and Wright, 2011). As Heiko (2012) states “The method value has been scientifically and practically proven” (Heiko, 2012, p.1526).

For a long time, researchers have considered the Delphi method as valuable, flexible, and well-suited to most research topics. The method has been widely and effectively adapted to serve various research purposes (Gallego and Bueno, 2014; Heiko, 2012; Okoli and Pawlowski, 2004; Rowe and Wright, 2011). The fact is “Delphi method is a versatile research tool that researchers can employ at various points in their research” (Okoli and Pawlowski, 2004, p.26).

However, the Delphi method, like others research methods, is not without limitations and weaknesses. These limitations must be identified and recognized so then they can be addressed and rectified as possible, in order to mitigate their undesirable effects (Gordon, 1994; Hasson and Keeney, 2011; Heiko, 2012; Hsu and Sandford, 2007b; Landeta, 2006; Rowe and Wright, 2001, 2011).
The limitations of the Delphi method have been reported by several researchers. Those researchers have suggested ways of improving the application of the Delphi method as well (Akins, Tolson and Cole, 2005; Gallego and Bueno, 2014; Gordon, 1994; Hasson and Keeney, 2011; Heiko, 2012; Hsu and Sandford, 2007b; a; Landeta, 2006; Linstone, 1985; Okoli and Pawlowski, 2004; Powell, 2003; Rowe and Wright, 2001, 2011). However, there is little information on how those developments and suggestions have been applied in subsequent studies performed by IS researchers.

Hence, our literature review was aimed at obtaining a general understanding of how Delphi studies have been designed and implemented in IS research and how recommendations for rigor and overcoming of the method’s limitations were integrated.

The current review contributes to a better understanding of how IS researchers are using the Delphi method and at which extent the recommendations to ensure rigor and overcome the method’s limitations are being implemented.

This paper provides an overview of the precautions that researchers should take in the application of the method and describes how these precautions have been implemented by the authors of papers in the Senior Scholars’ Basket of Journals, which includes eight journals considered by the Association for Information Systems (AIS) as journals of high scientific quality. The senior IS scholars’ basket of journals of AIS are: European Journal of Information Systems (EJIS), Information Systems Journal (ISJ), Information Systems Research (ISR), Journal of the Association of Information Systems (JAIS), Journal of Information Technology, Journal of management information systems (MIS), Journal of Strategic Information Systems, and Management Information Systems Quarterly (MISQ). The Delphi’s published papers in the senior IS scholars’ basket of journals of AIS represent only a part of the Delphi studies carried out by the IS community, but they have a high quality seal given by the journals that published them, which is indicative that they are studies of high scientific rigor. Therefore, they are models for the research done by researchers less experienced in applying the Delphi method. The analysis of the method application process is relevant to develop learned lessons that shed light on how to make better decisions along the Delphi process when designing and performing IS studies (Day and Bobeva, 2005; Gallego and Bueno, 2014; Galliers and Huang, 2012; Skinner et al., 2015; Sarker, Xiao and Beaulieu, 2013).

1.1 Process of reviewing the literature

In order to fulfill the aforementioned aims and contributions, we searched within the time frame (2004-2017) the eight top publications referred in the senior IS scholars’ basket of journals of the Association for Information Systems (AIS).

In order to know how the method’s developments and suggestions concerning its application have been integrated in Information Systems (IS) studies, a study of Paré et al (2013) presents the results of the rigor assessment of 42 IS Delphi studies retrieved from several databases that were published between 1984 and mid-2010. From those, 28 of the studies were published before 2005 (67%). Their study focused mainly on the rigor of the implemented process. The authors’ aim was to gather as many studies as possible regardless the ranking of the journals where they were published.

This 2013 literature review guided our study and it was decided to complement it by focusing on scientific journals that have a strong consensus on their relevance for the international IS community. Accordingly, we decided to restrict our assessment to the AIS Senior Scholars’ basket of journals. The eight journals in the AIS Senior Scholars’ basket of journals were searched using the keyword “Delphi”. The search covered the years 2004-2017. Only sixteen (16) IS research papers were retrieved. We considered that an interval of 12 years would cover the most recent studies, those benefiting most from the recommendations of previous scholars that had refined the method.

1.2 Structure of the paper

The remainder of this paper is organized as follows. The next section provides an extensive description of the main methodological steps of the Delphi method, including the selection and composition of the experts’ panel, setting of panel size and number of Delphi rounds, building consensus among the experts, and
assessments of the validity of Delphi results. This section is then followed by the presentation of the quantitative and qualitative results of our literature review. In the final part of the paper, we list several recommendations reflecting on how to apply the method based on the insights gained from the analysis of literature.

2. **Methodological Steps for Conducting a Delphi Study**

Various challenges and limitations concerning the use of the Delphi method have been identified as critical issues to consider (Akins, Tolson and Cole, 2005; Gallego and Bueno, 2014; Gordon, 1994; Hasson and Keeney, 2011; Heiko, 2012; Hsu and Sandford, 2007a; b; Landeta, 2006; Linstone, 1985; Okoli and Pawlowski, 2004; Powell, 2003; Rowe and Wright, 2001, 2011). Therefore, many of prior Delphi contributions have sought to define, discuss, and solve such challenges in order to ensure the validity and credibility of research results. If they are not adequately addressed, the quality of the Delphi results will be questioned.

The aim of this section is to summarize the main concerns, therefore providing the theoretical ground to assess the IS studies gathered.

2.1 **Selection and composition of the panel of experts**

Selecting the appropriate Delphi participants is one of the most important challenges in the entire Delphi method process because it directly relates to the quality of the results generated (Hsu and Sandford, 2007b; Hasson and Keeney, 2011; Rowe and Wright, 1999; Landeta, 2006). The Delphi method is not intended for generalization of results, instead aims at providing in depth insight about a complex problem (Gordon, 1994; Powell, 2003; Linstone and Turoff, 1975). Wheeler (1985) states that the Delphi “is more concerned with exploring minds than setting down precise recommendations” (Jenkins and Smith, 1994, p.413). Therefore, participants in the process should not be selected randomly (Akins, Tolson and Cole, 2005); instead they must be selected carefully (Gordon, 1994; Murphy et al., 1998; Powell, 2003). As Gordon (1994) states, the panel “represent the synthesis of opinion of the particular group, no more, no less” (Gordon, 1994, p.4).

In order to establish a clear set of criteria for deciding who should be considered as an “expert”, several authors recommend that: 1) the participant has first-hand background knowledge and experience on the topic under investigation; 2) the participant can provide useful contributions; 3) the participant is willing to allocate a substantial amount of time to participate, since Delphi method can be time-consuming by requiring the participation in multiple rounds of iteration to revise initial judgments (Adler and Ziglio, 1996; Hsu and Sandford, 2007b; Stitt-Gohdes and Crews, 2005); 4) the participant is respected and well-known in the relevant fields (Ludwig, 1997). Other criteria may also be used. For example, to request the expert’s experience self-rating may be a suitable technique to identify participants with the adequate experience in the problem area of the study. This technique requires that participants consider themselves experts in the topic (Linstone, 1985), (knowing about a topic does not necessary mean to be an expert (Keeney, McKenna and Hasson, 2010). Adding to this procedure, the selection of experts may be informed by an analysis of their scientific publications (Hsu and Sandford, 2007b).

A significant challenge of the Delphi method is the possibility of experts to drop out during the process (Borg and Gall, 1983), or that they may become less willing to participate over the period of the study. Keeping experts engagement during the study is not as easy as it may seem since it is expected that they answer the same questions over and over again, therefore consuming a lot of their time (Landeta, 2006). To avoid the effort, experts may change their opinion to force a quicker [apparent] consensus, which may cause a negative effect on the final result credibility (Rowe and Wright, 1999). It has been suggested that a good strategy to ensure the experts’ engagement is to select experts who are engaged with the topic, and ensure their ability and willingness to participate before starting the study (Skulmoski, Hartman and Krahn, 2007). Another strategy is to design a communication process whereby researchers keep frequent contact with all panelists during the data collection and analysis phases (Day and Bobeva, 2005; Hsu and Sandford, 2007b; Linstone, 1985; Ludwig, 1997), therefore building a sense of involvement and ownership.

2.2 **Setting the size of the panel and the number of Delphi rounds**

Many pioneering Delphi studies had used and recommended a small panel size (Linstone and Turoff, 1975). Dalkey et al. (1969) have found that “a suitable minimum panel size is seven” (Linstone, 1985, p.644). However, recommendations on the number of experts to involve vary in the literature. For instance, 15 to 18
experts are usually considered as an acceptable sample size in a homogenous panel (Day and Bobeva, 2005; Adler and Ziglio, 1996; Okoli and Pawlowski, 2004), and less than 10 may impact negatively the quality of results (Okoli and Pawlowski, 2004). According to Witkin and Altschuld (1995), the Delphi panel should never exceed 50 (Hsu and Sandford, 2007b). Others state that "the majority of Delphi studies have used between 15 and 20 respondents" (Ludwig, 1997, p.2), or 35 experts (Gordon, 1994). Some studies have employed a large number of experts, some studies involving several hundred or thousands of experts as reported by (Akins, Tolson and Cole, 2005; Gallego and Bueno, 2014; Hsu and Sandford, 2007b; Linstone and Turoff, 1975)

There are important debates about the expert panel size since it may influence the final results (Hasson and Keeney, 2011; Rowe and Wright, 1999). From one hand, some studies found a positive relationship between the number of experts and the results’ reliability. According to Donohoe and Needham (2009), the Delphi results validity, efficacy, and reliability are dependent on expert group size (Donohoe and Needham, 2009).

Some believe that higher degrees of reliability require more participants; as the panel size increases, the reliability of results increase (Donohoe and Needham, 2009; Murphy et al., 1998). Others observed that the accuracy of results declines with small panels (Dalkey, Brown and Cochran, 1970; Linstone, 1985). Meanwhile, other authors questioned such relationship. That is, for some scholars a small number of experts is sufficient to ensure the reliability of results (Akins, Tolson and Cole, 2005). They stress that a causal relationship between the size of the panel and the rigor of the final consensus (concerning validity and reliability) cannot be established (Gordon, 1994). In other words, a large panel of experts does not necessarily means better results (Keeney, Hasson and McKenna, 2011). The result of this debate has been the recommendation for considering the expert’s knowledge and experience in the topic under investigation. That is, if the identified experts have consistent and extended experience in the topic, then 10 to 15 experts are sufficient (Adler and Ziglio, 1996; Delbecq, Van de Ven and Gustafson, 1975). It has been argued that a group of 7 highly experienced participants attains the highest level of performance (Brochhoff, 1975). When a study requires the involvement of large number of experts, researchers are encouraged to define various homogenous experts’ panels which seems convenient technique to facilitate the process of reaching consensus and decreasing the number of rounds thereafter (Adler and Ziglio, 1996). In conclusion, there is no consensus in the literature on the appropriate panel size for the Delphi study (Akins, Tolson and Cole, 2005; Hsu and Sandford, 2007a).

The Delphi method depends on multiple iteration rounds for obtaining consensus among panel participants (Dalkey and Helmer, 1963; Gordon, 1994; Linstone and Turoff, 1975). An investigation into the basics of the Delphi method shows no consensus on the suitable number of rounds (Akins, Tolson and Cole, 2005; Gordon, 1994; Hsu and Sandford, 2007b; a; Linstone and Turoff, 1975). Generally, it is stated that the method’s rounds vary between 2 and 10 (Woudenberg, 1991).

Theoretically, the Delphi method can be continuously iterated until consensus is achieved. Since some authors have found that more than 3 rounds tend to add very little improvement on insights provided by participants (Linstone and Turoff, 1975), Delphi studies should typically include 2 or 3 rounds (Delbecq, Van de Ven and Gustafson, 1975). Some Delphi studies have shown that more than two rounds do not necessarily consolidate consensus (Nowack, Endrikat and Guenther, 2011; Rowe and Wright, 1999), while others point out that 3 iteration rounds are needed to collect the information for consensus (Daniel and White, 2005; Custer, Scarcella and Stewart, 1999). Some caveats have been raised related to implementing large number of rounds. The higher the number of rounds carried out, the slower is the convergence among the experts (Dalkey and Helmer, 1963; Linstone and Turoff, 1975). Too many rounds may exhaust experts and lead to important consumption of researchers’ resources and time (Schmidt, 1997). As a result, experts may give up such a long Delphi process. Due to the effort required, participants may change their opinions in order to hasten the end of the study, therefore building a fake consensus (Gallego and Bueno, 2014; Hsu and Sandford, 2007b; Landeta, 2006; Rowe and Wright, 1999; Yousuf, 2007).

2.3 Reaching experts consensus

Reaching a high or (an acceptable) level of consensus among experts is as crucial requirement if the Delphi method (Delbecq, Van de Ven and Gustafson, 1975; Linstone and Turoff, 1975). However, as “there is no pressure to arrive at a consensus” (Dalkey and Helmer, 1963, p.4), Delphi studies tend to end when the change
in the experts' opinion between rounds is minimal, even though a strong consensus had not yet been reached (Okoli and Pawlowski, 2004; Pare et al., 2013; Schmidt, 1997; Skulmoski, Hartman and Krahn, 2007). Hence, the recommendation is to stop the Delphi inquiry when theoretical saturation is achieved (Skulmoski, Hartman and Krahn, 2007) or when a satisfactory level of concordance has been reached rather than reaching the full consensus among the experts (Schmidt, 1997; Okoli and Pawlowski, 2004; Pare et al., 2013). In relation to the consensus concept, scholars have commented favorably the reaching of an agreement among experts as the method's goal, instead of striving for consensus (Meijering, Kampen and Tobi, 2013). The findings of a Delphi study will represent the experts' opinion, rather than an indisputable fact, therefore consensus is not an issue (Powell, 2003). Keeney et al. (2011) highlight that the Delphi method is not about getting the right or wrong answer; instead it seeks valid opinions of experts (Keeney, Hasson and McKenna, 2011). Hence, it does not only reveal the points of agreement among experts but also the ones that can generate beneficial debates (McKenna, 1994).

Nevertheless, Delphi scholars stress the importance of ensuring the quality of the final Delphi insights. This requires (1) to identify attempts to manipulate the Delphi process in order to reach its final stage faster (Linstone, 1985); (2) to spot situations of group-pressure towards a specific opinion, which participants may adopt to comply with the group (Grotte, Anderson and Robinson, 1990; Sackman, 1974); and (3) to identify less confident or pro-active participants that simply follow the perceived panel opinion just to bring the process to an end (Hill and Fowles, 1975; Grotte, Anderson and Robinson, 1990).

2.4 Assessing the validity of the Delphi method's results

Validity has been defined as a measure to assess (or test) the theoretical anticipations/ claims/ hypotheses emerging from prior work in the research topic (Zikmund et al., 2013; Sekaran and Bougie, 2010). Validity is essential to ensure that the results are reliable and represent the phenomenon that is being studied (Golafshani, 2003; Zikmund et al., 2013).

Concerning the validity of the Delphi method, Landeta (2006) states that it is a valid instrument for forecasting and supporting decision-making (Landeta, 2006). Validity becomes imperative to ensure the meaningfulness of results and their utility also for research (Yousuf, 2007). Numerous authors have shown that the Delphi method provides evidence for validity (Landeta, 2006), worthiness and effectiveness in conducting both qualitative and quantitative studies (Hsu and Sandford, 2007b; Skulmoski, Hartman and Krahn, 2007). To sum up the literature on the validity issue, we categorized the insights into the following three focal points: First, **validity of consensus**. Validity can be influenced by the number of panel experts, their degree of expertise in the topic and the similarity of their experience (Donohoe and Needham, 2009; Keeney, Hasson and McKenna, 2011; Murphy et al., 1998; Rowe and Wright, 1999; Van Zolingen and Klaassen, 2003). Second, **predictive power of results**. Helmer and Helmer-Hirschberg (1983) pointed out that if relatively stable and verifiable predictions cannot be produced using the Delphi method, the validity of results is hard to determine (Helmer and Helmer-Hirschberg, 1983). Third, **validity of results**. The usual open ended questions used at the first Delphi round may create broad scope statements, if not ambiguous ones, that can translate into biased contributions and biased results (Hsu and Sandford, 2007a). For such reason, some scholars suggest closed questions (modified Delphi). The clarity of the statements directly influences the reliability of the results (Hasson and Keeney, 2011), and so, closed questions are recommended (Hsu and Sandford, 2007a). However, contrary to open ended questions, close ended ones may drastically reduce the richness of the data collected.

3. Delphi Method Studies in the IS Field: A Literature Review

The results of the literature review are presented in two parts. Each of them is designed to achieve a specific goal. In the first part, "Quantitative Results", quantitative information about the analysed studies is presented: the number of papers per journal are presented in Table 1, Table 2 presents details about the application of the Delphi method including the number of implemented rounds, the number of experts included, and the characteristics of the experts' panel. The second part, "Qualitative Results", adopts a different approach; that is, through in-depth discussion, the available information about the implementation of the Delphi method provided in the 16 IS studies is linked to the challenges described in Section 2.

3.1 Quantitative results: the research design of Delphi studies

This section provides an overview of how the Delphi method was implemented in the collected studies. As shown in Table 1, it is evident that the surveyed journals have published Delphi studies almost every year since
2004. It is also clear that the Delphi method does not enjoy of great popularity among scholars who have been publishing in these journals.

Table 1: Distribution of Papers per Journal

<table>
<thead>
<tr>
<th>Journal</th>
<th>Papers</th>
<th>Years of Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Information Systems Research</td>
<td>(Agerfalk, Fitzgerald and Slaughter, 2009)</td>
<td>2009</td>
</tr>
<tr>
<td>Journal of Information Technology</td>
<td>None</td>
<td>--</td>
</tr>
<tr>
<td>Journal of Strategic Information Systems</td>
<td>None</td>
<td>--</td>
</tr>
<tr>
<td>Management Information Systems Quarterly</td>
<td>(Chou, Zahedi and Zhao, 2014; Saunders and Benbasat, 2007)</td>
<td>2007, 2014</td>
</tr>
</tbody>
</table>

The studies covered various IS topics including: evaluation of IT projects (Aloini, Dulmin and Mininno, 2012; Kasi et al., 2008), comparison of IT project’s risks (Liu et al., 2010), evaluation of the impact of a knowledge-based system (KBS) in supporting group decision making (Nah and Benbasat, 2004), comparison of the efficacy of two cognitive based techniques; Delphi and Repertory Grid, for elicitation and capture of knowledge (Chiravuri, Nazareth and Ramamurthy, 2011), identification of challenges in implementing a project management office (Singh, Keil and Kasi, 2009), the improvement of an inter-organizational IT system (Daniel and White, 2005), assessment of natural disaster management websites as a way to identify required improvements (Chou, Zahedi and Zhao, 2014), elicitation of the factors underlying the susceptibility of certain individuals to online “Phishing” attacks (Moody, Galletta and Dunn, 2017), designing of a decision support system (DSS) that aids decision-makers with repurposing electric vehicle batteries (Klör et al., 2017), evaluation of business maturity models (Van Looy, Poels and Snoeck, 2017), elicitation of topics in IS development for future research (Agerfalk, Fitzgerald and Slaughter, 2009), and determination of how to increase the number of top-quality papers published by the IS community (Saunders and Benbasat, 2007).

Based in our analysis of the gathered papers, it is possible to conclude that the Delphi method has been used predominantly for qualitative exploratory studies (Agerfalk, Fitzgerald and Slaughter, 2009; Daniel and White, 2005; Kasi et al., 2008; Kasiri, Sharda and Hardgrave, 2012; Klör et al., 2017; Liu et al., 2010; Saunders and Benbasat, 2007; Singh, Keil and Kasi, 2009). It has also been used as a method to elicit the variables that would be used for subsequent quantitative studies (Aloini, Dulmin and Mininno, 2012; Chen et al., 2014; Chiravuri, Nazareth and Ramamurthy, 2011; Chou, Zahedi and Zhao, 2014; Görg and Bergmann, 2015; Liu et al., 2010; Moody, Galletta and Dunn, 2017; Nah and Benbasat, 2004; Saunders and Benbasat, 2007).

Table 2: IS Delphi Rounds and Panels

<table>
<thead>
<tr>
<th>Paper</th>
<th>Number of Rounds</th>
<th>Panel Size</th>
<th>Panel Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Nah and Benbasat, 2004).</td>
<td>2 rounds</td>
<td>7 experts</td>
<td>Single panel of experts Heterogeneous</td>
</tr>
<tr>
<td>(Daniel and White, 2005).</td>
<td>3 rounds</td>
<td>35 experts</td>
<td>Single panel of experts</td>
</tr>
<tr>
<td>(Saunders and Benbasat, 2007)</td>
<td>3 rounds for one panel and 2 rounds for another panel</td>
<td>54 senior scholars 76 junior scholars</td>
<td>Two panels of experts homogeneity in each panel</td>
</tr>
<tr>
<td>(Kasi et al., 2008)</td>
<td>4 rounds</td>
<td>23 experts</td>
<td>Single panel of experts Heterogeneous</td>
</tr>
<tr>
<td>(Agerfalk, Fitzgerald and Slaughter, 2009).</td>
<td>3 rounds</td>
<td>14 experts</td>
<td>Single panel of experts Homogenous</td>
</tr>
<tr>
<td>(Singh, Keil and Kasi, 2009)</td>
<td>2 rounds</td>
<td>22 experts</td>
<td>Single panel of experts Homogenous</td>
</tr>
<tr>
<td>(Liu et al., 2010)</td>
<td>2 rounds for one panel and 3 rounds for another panel</td>
<td>64 experts; 30 IT senior executives, and 34 IT project managers</td>
<td>Two panels of experts homogeneity in each panel</td>
</tr>
<tr>
<td>(Chiravuri, Nazareth and Ramamurthy, 2014)</td>
<td>5 rounds</td>
<td>40 experts; 20 Delphi experts and 20 Rep grid experts</td>
<td>Two panels of experts homogeneity in each panel</td>
</tr>
</tbody>
</table>
Table 2 shows that regarding the panel size, our sample presents a wide variability; the panels vary between 7 and 76 experts. Half the studies included panels of 10-20 experts in size and the other half included 21 experts or more. The table also highlights the use of two panels of experts (Chen et al., 2014; Chiravuri, Nazareth and Ramamurthy, 2011; Liu et al., 2010; Saunders and Benbasat, 2007; Van Looy, Poels and Snoeck, 2017). Additionally, the analysed papers have shown a considerable concern over distributing the experts into homogeneous panels. As Table 2 shows, the majority of studies have formed homogeneous panels.

### 3.1.1 Qualitative results: implementation of recommendations

The purpose of this section is to describe how the analysed studies implemented the theoretical recommendations presented in Section 2 of this paper.

#### Research design of IS Delphi studies: The Delphi method was often used for exploratory purposes in the analysed IS studies. However, some authors apply the Delphi method to test hypotheses ([Aloini, Dulmin and Mininno, 2012; Chen et al., 2014; Chou, Zahedi and Zhao, 2014; Görg and Bergmann, 2015; Moody, Galletta and Dunn, 2017] as well as experimental ([Chiravuri, Nazareth and Ramamurthy, 2011; Klör et al., 2017; Nah and Benbasat, 2004] and descriptive studies ([Liu et al., 2010; Saunders and Benbasat, 2007]. Some studies used the Delphi method to elicit the experts’ opinions and views as the basis for the development of a theoretical model or an assessment instrument that was later tested using the survey method ([Chen et al., 2014; Chou, Zahedi and Zhao, 2014; Görg and Bergmann, 2015]. As an example, in one of the studies the experts’ experience was used to develop an instrument designed to evaluate the content of 50 natural disaster management websites. The instrument was later on validated in a survey study of users of those websites ([Chou, Zahedi and Zhao, 2014]. In another study, the Delphi results were used to develop a theoretical instrument to predict users’ active behaviors on social networking sites ([Chen et al., 2014]. Another Delphi study was designed to identify relevant factors that may significantly predict students’ susceptibility to phishing scams. The study was then followed by a large scale survey to empirically investigate the effect of the identified factors on students’ susceptibility to phishing ([Moody, Galletta and Dunn, 2017]. In summary, IS researchers have been using the Delphi method for a wide variety of purposes, most often for theoretical developments to be tested in subsequent quantitative studies. ([Creswell and Clark, 2017], which is in line with what happens in other scientific fields ([Skinner et al., 2015].

#### Identification and Selection of Experts: In order to build credibility, it has been recommended that details are provided about the process of experts’ identification and selection ([Hsu and Sandford, 2007b; Schmidt, 1997]. 10 out of the 16 studies analysed provide a detailed description of how the experts were selected and recruited. The selection process was based on the assessment of the expert’s knowledge and experience in the
topic being researched, as well as her/his willingness to participate and serve as an expert (Ågerfalk, Fitzgerald and Slaughter, 2009; Chou, Zahedi and Zhao, 2014; Daniel and White, 2005; Görg and Bergmann, 2015; Kasiri, Sharda and Hardgrave, 2012; Liu et al., 2010; Nah and Benbasat, 2004; Saunders and Benbasat, 2007; Singh, Keil and Kasi, 2009; Van Looy, Poels and Snoeck, 2017).

4 of the remaining 6 studies provided only minimal information on how experts were identified and selected (Aloini, Dulmin and Mininno, 2012; Görg and Bergmann, 2015; Moody, Galletta and Dunn, 2017; Nah and Benbasat, 2004).

Therefore, 1/3 of the papers published in top journals in the IS field include minimal to no information about the process implemented to select and recruit experts; this process is considered of utmost importance to ensure the quality of the results generated.

**The number of Delphi rounds:** The majority of the studies in our sample implemented 2 to 3 rounds, sometimes applying them to 2 panels (Liu et al., 2010; Saunders and Benbasat, 2007). This is in line with the recommendations.

On the other hand, 4 of the studies that implemented more than 3 rounds used the Delphi method as the departure point for a subsequent quantitative study (Chen et al., 2014; Chiravuri, Nazareth and Ramamurthy, 2011; Moody, Galletta and Dunn, 2017).

The studies that employed 4 or 5 rounds provided a justification for using more than the recommended 2/3 rounds (Chen et al., 2014; Chiravuri, Nazareth and Ramamurthy, 2011; Kasi et al., 2008; Van Looy, Poels and Snoeck, 2017), therefore acknowledging the recommendation of 2/3 rounds. For example, Kasi et al. (2008) used a fourth round so experts were given the opportunity to clarify and to comment on their responses (Kasi et al., 2008). Also, Chiravuri et al. (2011) conducted a fifth round to observe consensus formation not only at an overall level but also for each one of the cases or scenarios that were investigated (Chiravuri, Nazareth and Ramamurthy, 2011). Van Looy et al. (2017) conducted the fourth and fifth Delphi rounds to rank and weight the importance of several criteria to evaluate the maturity of BPMs; these criteria were set in the first three Delphi rounds. The use of more than 3 rounds may be useful in exploratory studies on a little known phenomenon. In the future, it may be interesting to study whether exploratory Delphi studies lead to more rounds until a satisfactory agreement / consensus is reached.

**Panel size:** The panel size may depend on the context and the aims of a particular study. According to (Nowack, Endrikat and Guenther, 2011; Yousuf, 2007), the minimum number of experts to be involved, to some extent, depends on the study design, scope, and context. Our sample seems to confirm this trend. In our sample, descriptive and explanatory research tend to recruit large panel sizes including 73 experts in single panel (Chou, Zahedi and Zhao, 2014), 54 and 76 participants in two panels (Saunders and Benbasat, 2007), and 64 participants in two panels (Liu et al., 2010). Exploratory studies, however, have implemented small panel sizes, which range between 10 to 23 experts as founded in the (Ågerfalk, Fitzgerald and Slaughter, 2009; Aloini, Dulmin and Mininno, 2012; Görg and Bergmann, 2015; Kasi et al., 2008; Kasiri, Sharda and Hardgrave, 2012; Klör et al., 2017) studies. The only exception is (Daniel and White, 2005), that while being an exploratory study it involved a panel of 35 experts. Therefore, our analysis confirms previous understanding that the type of the study will determine the panel size.

From the analysis of the 16 studies also emerged that often the size of the panel is not defined *a priori*, instead resulting from the initial responses to the first call for participation For instance, in the study (Chou, Zahedi and Zhao, 2014), researchers sent out 718 invitations to participate and the 73 experts integrating the panel were those who accepted that initial invitation. Thus, the panel size seems to depend on the researchers’ ability to reach out for experts on the topic being studied and then to convince them to participate. In one paper, authors detailed the difficulties they encountered in recruiting a sufficient number of experts for their study (Nah and Benbasat, 2004).

Another paper detailed difficulties in finding experts with adequate experiences on the topic that was investigated (Singh, Keil and Kasi, 2009).
From the analysis of the collected papers emerges that the recruitment of experts is the most challenging phase of the method as well as the one that has a more uncertain result. The acceptance rate of calls for participation is very low and it is difficult to ensure that the experts most experienced on the topic studied will be available to participate. As this is an important aspect to ensure the quality of the results, it is not possible to avoid having some doubt about the insights obtained in the studies analysed, reinforcing the importance of combining the Delphi method with another research method that validates the results of the Delphi study.

**Reaching panel consensus and the number of Delphi rounds:** According to the methodological guidelines, the Delphi rounds end when a consensus, or at least a strong agreement, has been reached. For most of the studies analysed this was achieved in 2 or 3 Delphi rounds (Ågerfalk, Fitzgerald and Slaughter, 2009; Chou, Zahedi and Zhao, 2014; Daniel and White, 2005; Görg and Bergmann, 2015; Kasiri, Sharda and Hardgrave, 2012; Liu et al., 2010; Nah and Benbasat, 2004; Saunders and Benbasat, 2007; Singh, Keil and Kasi, 2009). Four examples deserve to be noticed; while in the studies of (Chen et al., 2014; Klör et al., 2017) researchers stopped their Delphi rounds due to the high stable responses they were receiving from all experts. (Singh, Keil and Kasi, 2009) stopped their study after the second round due to some degree of panel fatigue and to avoid potential panel drop-off from the study. Similarly, in the study of (Van Looy, Poels and Snoeck, 2017) the rounds stopped because 75 percent of the experts indicated they were no longer willing to continue the Delphi process. It should be noticed that the (Van Looy, Poels and Snoeck, 2017) study is the only Delphi study in our sample that defined stopping conditions before starting the study. Our results show that while researchers may aim for a certain degree of consensus, it may not be reachable due to the demotivation of experts. Therefore, as recommended by experts in the application of the Delphi method, it is necessary to plan engaging participation tasks.

**Reaching consensus and panel size:** It has been recommended that the panel does not exceed 50 experts (Adler and Ziglio, 1996; Gordon, 1994; Hsu and Sandfor, 2007a; Linstone, 1985; Okoli and Pawlowski, 2004; Linstone and Turoff, 1975). This is considered the limit for an effective management of the participation and to ensure high quality results.

All the studies we analysed provide information about the panel size, which is considered a desirable procedure. The majority of them avoided recruiting more than the recommended 50 experts.

Distributing the experts into homogenous panels seems to facilitate the process of reaching consensus and, in consequence, of reducing the number of rounds required (Chou, Zahedi and Zhao, 2014; Kasiri, Sharda and Hardgrave, 2012; Liu et al., 2010; Nah and Benbasat, 2004; Saunders and Benbasat, 2007). For instance, reaching a consensus after only 2 rounds in (Kasiri, Sharda and Hardgrave, 2012) was possible because all the experts were from the retail industry or were RFID consultants (homogenous panel).

Our analysis of the Delphi studies indicates that reaching consensus (or strong agreement) among participants depends on the panel size and its homogeneity. However, despite the homogeneity of the panel, 3 studies (Chen et al., 2014; Chiravuri, Nazareth and Ramamurthy, 2011; Kasi et al., 2008), required 4 or 5 rounds to reach a clearer consensus about complex/emerging topics.

**Reported Response Rate:** Obtaining a high response rate from participants of Delphi studies seems crucial since it significantly enhance research rigor and validity (Hasson and Keeney, 2011). The studies (Ågerfalk et al., 2009; Chiravuri et al., 2011; Chou et al., 2014; Daniel and White, 2005; Klör et al., 2017) in our sample show a high response rate of experts reflecting a successful procedure implemented to engage the experts in the Delphi process. In (Daniel and White, 2005), the authors link this success to the involvement of experts that were directly recommended to them. In (Chiravuri et al., 2011), previous email and phone contacts ensured that the involved experts were interested in the researched topic and willing to participate in a process that would require some degree of effort; this was considered as the key factor for the high response rate.

The use of websites supports the Delphi process and makes the participation possible without the need for face-to-face meetings, thus contributing to increase the response rate (Daniel and White, 2005; Saunders and Benbasat, 2007; Singh et al., 2009). These websites also allow for the respondent’s anonymity (Cole, Donohoe and Stellefson, 2013; Witkin and Altschuld, 1995).
To engage experts in the Delphi study, in (Kasiri et al., 2012) researchers interviewed experts in the first round (including face-to-face and over the phone interviews), and in the second round a questionnaire was administered via email. In (Liu et al., 2010; Singh et al., 2009) online brainstorming was used to facilitate the first round of their Delphi study. In studies (Ägerfalk et al., 2009; Chiravuri et al., 2011) training was provided to the involved experts. Therefore, our analysis show that in some studies researchers implemented actions aimed at engaging experts in the collection of their views. However, provided information does not allow for the assessment of the effectiveness of their actions beyond the specific context of their studies.

**Assessing Validity and Reliability:** Validity and reliability of results are explained in some of the studies we analysed (Chen et al., 2014; Chiravuri et al., 2011; Chou et al., 2014; Liu et al., 2010). The study of (Klör et al., 2017) present sufficient information regarding the level of agreement among the experts achieved in the two Delphi rounds of the study; such information adds to the rigor and validity of the study. Other studies not only used statistical validation of results, but also provided many details about the applied Delphi process. In (Chen et al., 2014), authors described various stages of the Delphi process they implemented; they also presented the statistics of the achieved consensus, including the percentage of agreement and the coefficient of variation among experts. In (Liu et al., 2010; Singh et al., 2009) the authors detailed the Delphi phases they used together with the achieved results. Later, they calculated the degree of consensus by using Kendall’s coefficient of concordance, as suggested by (Schmidt, 1997). In (Van Looy, Poels and Snoeck, 2017), the authors measured stability of results with Spearman’s rho, Kendall’s tau-b, and the Cohen’s Kappa level of agreement. They also describe the adopted approach to minimize any potential bias either in gathering of information or producing the results. In (Chou et al., 2014), the authors removed experts who gave the same rating to all questionnaire elements since it might indicate a lack of attention and commitment to the study.

Also, they asserted the validity of their study by using reliability statistical measures like Cronbach’s alpha and Intra class correlation coefficient (ICC). The authors of the study described in (Chiravuri et al., 2011) took several steps to ensure external and internal validity; they applied statistical analysis for reconfirming reliability, namely the Cronbach’s alpha of the measures used in their study. Nevertheless, these are only 6 studies out of the 16 in our sample. This implies that the majority of the collected studies (63%) don’t provide detailed information about the validity of the study and results achieved.

### 3.2 Discussion

The Delphi method became known some 60 years ago when the RAND Corporation used it for military defence (Dalkey and Helmer, 1951, 1963). Since then, Delphi method has been widely proven as a suitable research tool both for prediction proposes and for developing scientific and technical knowledge (Linstone and Turoff, 1975; Skulmoski et al., 2007). For almost three decades (since 1980s), Delphi method has been used by IS scholars with different degrees of popularity (Skinner et al., 2015). Specifically, the method received an increasing attention, particularly from 2006 onwards (Gallego and Bueno, 2014; Pare et al., 2013).

The Delphi method has proven suitable in a variety of fields (Skulmoski et al., 2007), and has evolved and effectively modified to meet the specific needs and objectives of a number of research topics (Gallego and Bueno, 2014; Kasi et al., 2008; Okoli and Pawlowski, 2004; Pare et al., 2013; Skulmoski et al., 2007). The Delphi method has been used in the IS discipline with success. Scholars have often used and adapted it beyond its initial purpose (long-term prediction) to effectively meet the needs of IS research (Okoli and Pawlowski, 2004). While some authors highlighted the method’s flexibility for a wide variety of IS research topics, caution was advised regarding the use of the method to study some of the IS research topics and situations (Skinner et al., 2015; Skulmoski et al., 2007).

The review of Delphi studies presented in this paper provides three main results. First, the Delphi method isn’t a popular research method among the authors publishing in the Senior IS scholars’ basket of journals. In the fairly large time interval of our review (2004-2017) we could find only 16 Delphi studies. Moreover, half of the collected Delphi studies were published after 2011.

Second, in our sample, we found that the majority of papers address in detail the Delphi method’s main steps, challenges, and weaknesses, which creates confidence that the studies were carefully designed and implemented so the method’s strengths were maximized and weaknesses were handled satisfactorily. In fact,
most studies provide abundant information about the implementation process and the precautions taken to ensure reliability and the validity of results.

Third, the Delphi method’s critical issues (e.g., experts’ identification and selection, experts’ allocation into homogenous panels, number of rounds, reaching consensus among experts, and validity and reliability of Delphi’s results) have been addressed in most studies. That is, the majority of analysed papers include a careful description of (i) the implemented procedure to ensure the selection of highly “motivated” experts, (ii) the rationale supporting the researchers’ definition of an adequate panel size for their study, and (iii) the care taken to distribute the experts into homogeneous panels. These approaches to address the challenges of the method follow previous recommendations and seem to confirm their relevance to reach strong consensus among experts and ensure the validity and reliability of results.

Since the surveyed set of journals are considered a reference for high quality research and, therefore, the papers published there are taken as models to replicate by other IS researchers, it is relevant to highlight the areas that must be further explored according to our analysis of the papers.

The studies that tested or validated the results obtained by applying the Delphi method did so by implementing quantitative research. The combination of the Delphi method with another qualitative method was not found in our sample and our research can be complemented with a literature search in other top-ranked journals of studies that combined the Delphi method with another qualitative method. This verification will allow to know other applications of the Delphi method in addition to theoretical development, namely as an aid in in-depth studies of IS phenomena.

Our sample of papers presents some ambiguities that may cast doubt on the validity of the results obtained, namely:

- 1/3 of the papers include almost no details about the process to select and recruit experts;
- The size of the panel seems to be more associated with the ability of researchers to recruit and retain the experts involved in the Delphi process than with a research decision on the quantity and profile of the participants relevant to obtaining an understanding of the topic investigated;
- Despite being referred to various approaches to keep participants involved in the Delphi process, none of the articles offers a profound reflection on the usefulness of the used technique may have in other contexts. Neither are innovations introduced in this dimension, such as the implementation of gamification or other techniques. It emerges from this analysis that in this area, and in this sample of papers, there is no innovation that can be replicated in other studies.
- The majority of analysed papers don’t provide detailed information about the steps taken to ensure the validity of the study and achieved results.

4. Conclusions

This paper describes how the Delphi method has been used in IS research over that last 12 years, as reported in papers published by the AIS Senior Scholars’ basket of journals during 2004-2017. The number of published Delphi studies have focused various IS topics and grew in number after 2011. The analysed studies show that the method has been regularly extended and adapted to serve various IS research needs.

Our analysis shows that researchers interested in learning from the experience of colleagues researching in the IS field need to study the method well and know the existing recommendations to overcome difficulties in implementing the method. In this way, they will be able to critically analyse the experience reported in the articles published in top journals in order to choose those who will inform their research. In addition, it will be interesting to increase innovative practices in the dynamization of the interactions recommended by the method. In particular, IS researchers can use strategies and technologies developed to support creativity and discussion of ideas to increase the likelihood of active participation of the experts in the panel.

The literature review in this paper focused only a specific set of top-ranked journals in the IS field. The insights must be complemented with others that emerge from reviewing other leading IS journals. This extension of the literature review is necessary to gain a comprehensive insight into how the Delphi method is being used in Information Systems field.
References


www.ejbrm.com


