Towards Delphi Rigor: An Investigation in the Context of Maturity Model Development

Samuel Reeb
Business Information Systems - Information Management, TU Dresden, Dresden, Germany

samuel.reeb@tu-dresden.de

Abstract: The maturity model (MM) and Delphi research areas are extensive and diverse, leading to numerous approaches. This study addresses the Delphi method regarding its rigor requirements within the IS literature. To this end, the example of maturity model development is investigated. Hence, Delphi studies for MM development are identified and analyzed regarding their rigorous application and design. The examination focuses on the connections between maturity model aspects and the Delphi methodology. Hence, relevant aspects of maturity model and Delphi literature are elaborated, and criteria for methodological rigor are derived. A key challenge is linking the method to the specific design objective (in this study example, the development of a maturity model). After conducting a literature search to identify studies that use the Delphi method for MM development, these criteria are used as a basis for deductive content analysis. The results indicate a lack of clarity regarding the methodology, as different aspects are reported, although the general demands, starting points, and goals were similar. A need for design guidelines for planning and conducting Delphi studies is emphasized and addressed in this paper. Hence, guidelines for designing Delphi studies are developed and presented, considering relevant aspects for ensuring rigorous implementation. The focus lies on the linkage between study design and intended model elements, as this demonstrates the complexity of the study design and the relevance of the design decisions through an example. The guidelines integrate the different methodological aspects of maturity model development and the Delphi methodology, providing an orientation framework for the design process of such research projects. Therefore, this study contributes to existing research by proposing design guidelines for Delphi studies to foster rigor in the specific context of maturity model development. Although the presented guidelines focus on the maturity model context, the general design approach and decisions are transferable and applicable to other domains. Hence, this research contributes to the Delphi literature by providing insights into how relevant elements should be addressed in the designing process of a Delphi Study. Scholars should investigate how the presented guidelines must be adapted for other domains in future research.

Keywords: Maturity model, Rigor, Delphi, Guideline, Model development, Research design

1. Introduction

The Delphi method is a well-established procedure in information systems (IS) research (Gallego and Bueno, 2014; Paré et al., 2013), and multiple variations of the approach exist (Paré et al., 2013; Skinner et al., 2015). The method's strengths lie in identifying current topics, consolidating different experts' opinions (de Bruin et al., 2005), and adaptability to specific research conditions (de Bruin et al., 2005; Gallego and Bueno, 2014). A methodologically rigorous application is essential to realize these benefits adequately. As the method allows adapting to study specifics (Gallego and Bueno, 2014), its design process is rather complex, and it risks losing methodological rigor (McKenna, 1994; Paré et al., 2013; Rauch, 1979; Skinner et al., 2015; Strasser, 2017). Consequently, the method is suitable for investigating complex issues, as comprehension questions and uncertainties are eliminated through its iterative process, and an adaptation to the concrete objective is possible (Skinner et al., 2015). However, it is questionable whether existing discussions about the methodological approach are sufficient and specific enough to ensure a clear understanding and accurate implementation.

The complex maturity model (MM) domain, comprising different development steps, is a possible application area for the Delphi method (Becker et al., 2009; de Bruin et al., 2005). In a narrower sense, model development is carried out in the "design model" step, for which various methods (e.g., case study interview, focus groups) can be applied (Mettler, 2011). Several scholars suggest that using the Delphi method is appropriate and beneficial for the "design model" step, as model deficiencies are already addressed at an early stage (Becker et al., 2009; de Bruin et al., 2005; Lasrado et al., 2015; Pereira and Serrano, 2020). Hence, the Delphi method achieves a more practical result in a shorter time during model development and tackles problems such as "limited exposure to relevant context" (Lasrado et al., 2017) or "competition for attention" (Lasrado et al., 2017) and should arguably be applied more often.

However, despite the Delphi method's advantages, MMs are rarely developed using this approach (Pereira and Serrano, 2020). The development of MMs using the Delphi method is barely addressed in current research, even though the documentation of this research process is essential as it is complex and diversely applied. A rigorous approach is vital because the Delphi method's diversity for MM development allows for various arbitrary options

Reference this paper: Reeb, S., 2023. Towards Delphi Rigor: An Investigation in the Context of Maturity Model Development *The Electronic Journal of Business Research Methods*, 21(1), pp. 01-13, available online at www.ejbrm.com

in the procedure and the associated risk of quality loss. Accordingly, the research design for the MM and the Delphi study must be aligned considering the different methodological aspects to ensure a rigorous approach. To ensure the quality and reliability of such a study and to counteract arbitrariness in the conduction, it is necessary to know how a Delphi-based MM development study can be designed. Consequently, the question arises as to whether researchers using a Delphi study to develop a MM follow a methodologically rigorous approach, which is not addressed in the literature. Therefore, MMs represent a suitable application area to study the Delphi method from a rigorous point of view, as the complexity of both approaches requires a deep understanding of the connections and a rigorous application of the Delphi method.

Therefore, this paper investigates the methodological rigor of Delphi studies in the IS domain using the MM development process as an example. Typical design demands for both methods are elaborated and used to analyze identified studies of maturity model (MM) development with the Delphi method. Accordingly, the following research questions emerge.

- To what extent do existing studies on maturity model development utilizing the Delphi method follow the methodological rigor?
- How can maturity model development studies using the Delphi method be designed rigorously?

2. Theoretical Background on Criteria for Methodological Rigor

The work on MMs in IS research is extensive (Carvalho et al., 2019; Mettler, 2011) and applied in various domains (Lasrado et al., 2015; Mettler, 2011). MMs depict how capabilities for a domain develop through discrete, successive maturity levels (Becker et al., 2009; Stelzl et al., 2020). The most prominent representatives of a methodological framework for MM development in the IS literature are de Bruin et al. (2005), Becker et al. (2009), and Mettler (2010). Comparing the approaches shows that the research designs follow similar steps (Mettler, 2011). Preeminent authors of the literature body repeatedly describe the Delphi method as a suitable and beneficial method for designing a MM (Becker et al., 2009; de Bruin et al., 2005; Lahrmann et al., 2011; Pereira and Serrano, 2020).

The Delphi method represents a structured, iterative procedure of an anonymous expert survey, in which expert knowledge is collected with the help of questionnaires, condensed, and controlled feedback of a statistically processed group response is given to the experts (Gallego and Bueno, 2014; Paré et al., 2013; Rowe and Wright, 1999). Delphi studies must have the following four generic characteristics to count as a Delphi study (Gallego and Bueno, 2014; Rowe and Wright, 1999):

- *Iterative process:* the study represents a round-based process, and in each round, the experts can communicate and adjust their opinions through a questionnaire and have the group position reflected to them (Paré et al., 2013; Rowe and Wright, 1999).
- *Anonymity*: the experts give their opinion independently and individually through questionnaires, without social pressure, and have the opportunity to change their opinion through the iterative process without losing face (Rowe and Wright, 1999; Skinner et al., 2015). Furthermore, it ensures that experts cannot trace the origin of provided information back to a single expert (Rowe and Wright, 1999).
- *Controlled feedback*: feedback on the group's responses is provided to the participating experts between rounds of questioning for comments and/or as a starting point for the follow-up round (Rowe and Wright, 1999; Skinner et al., 2015). The research team eliminates irrelevant information for the study in advance (Strasser, 2017).
- *Statistically processed group response*: The answers can be processed quantitatively and statistically to provide an anonymized group response to the experts (Rowe and Wright, 1999; Strasser, 2017).

The Delphi method is beneficial for designing a MM, as weaknesses can be identified and addressed in the initial development phase (Pereira and Serrano, 2020). An adequate development through a literature analysis is unlikely beyond the structural depth of the component layer (de Bruin et al., 2005). Thereby, a Delphi study is beneficial in the following points, which are relevant in the development of a MM within a complex domain (de Bruin et al., 2005):

- exploring and presenting complex topics (Okoli and Pawlowski, 2004)
- bringing opinions together for an improved result (de Bruin et al., 2005)
- making a scientific contribution to knowledge in a domain (Delbecq et al., 1975)
- lack of empirical evidence (Murphy et al., 1998).

Furthermore, the general objective of concept development of Delphi studies and the intention of the MM development coincide (de Bruin et al., 2005). Another advantage of the Delphi method is its adaptability to the specifics of a study (McKenna, 1994; Strasser, 2017).

Gallego and Bueno state that the majority of scholars use a modified Delphi method, as the versatility and flexibility to change each variable allow one to meet the specific needs of the research (Gallego and Bueno, 2014). Adjustments in the methodology can be helpful but reduce the quality and credibility of the method and thus the research results (Gupta and Clarke, 1996). Accordingly, methodological modifications risk losing methodological rigor, regardless of whether this is done intentionally (McKenna, 1994; Rauch, 1979; Skinner et al., 2015; Strasser, 2017). Several research contributions address the issue of rigor in Delphi studies and emphasize the need to consider general design elements and specifics of Delphi variants (Gallego and Bueno, 2014; Paré et al., 2013; Skinner et al., 2015; Strasser, 2017).

The remainder of this paper examines whether existing studies of MM development applying the Delphi method meet methodological rigor. Hence, this section elaborates on methodological requirements for Delphi studies to develop MMs, focusing on the interaction of both approaches. Table 1 shows an overview of the design elements of a Delphi study that are connected to the design elements of a MM. These requirements are later used to analyze existing studies regarding their methodological rigor.

Table 1: Overview of Design Elements and Their Relations to Delphi and MM

Delphi design element		**Relating MM design elements**
Delphi variant		Central constructs
Expert Panel	Expert Criteria	Model scope (e.g., function, auditorium), especially domain Central constructs
	Panel composition	Model scope (e.g., function, model complexity)
Number of rounds and Statistical Processing		Central constructs Development Basis
First-round design		Development Basis
Modifications		Research objective Model characteristics (model scope)

The Delphi method is used for forecasting, decision making, and concept development (Okoli and Pawlowski, 2004; Paré et al., 2013). Since the introduction of the classical Delphi method by Dalkey and Helmer (1963), multiple variations of the method have been introduced and adapted to different research problems (Okoli and Pawlowski, 2004; Paré et al., 2013; Strasser, 2017). The key characteristic of a modified Delphi study is difficult to define (Hasson and Keeney, 2011), causing no clear and uniform distinction in the literature on what a (genuine) Delphi variant is. Strasser (2017) identified seven genuine Delphi variants: Classic Delphi, Policy Delphi, Decision Delphi, estimate-feedback-talk-estimate (EFTE) Delphi, Ranking-Type Delphi, Argument Delphi, and Disaggregative Policy Delphi, which differ in their focus and objective. Furthermore, Delphi variants representing implementation variations of genuine variants, like Online Delphi, Real-Time Delphi, and Mini-Delphi, exist (Hasson and Keeney, 2011; Strasser, 2017). A central decision scholar must make is the selection of an appropriate Delphi variant according to the research objective, as Delphi variants can vary in their characteristics (e.g., terminology, goals, procedure),.

As this decision depends on the research objective (development of a MM), scholars need to specify this by elaborating on the intended MM´s central construct. MMs are structured into the central constructs of maturity levels, process area (components, sub-components), goals, assessment elements, development path, and improvement actions. A clear understanding of this aspect and decision is crucial for the design process. A short explanation of the different constructs and alternative terms is provided in the following, as no consistent terminology for the different structural elements exists in current literature. Hence, scholars should specify which central construct(s) is/are elaborated on through the Delphi study and define the terminology used regarding the existing literature on MMs.

Maturity levels (also maturity stages) represent the highest level of abstraction of a MM (Lasrado et al., 2015) and describe an archetypal representation of characteristics and conditions of a developmental level (de Bruin

et al., 2005; Object Management Group, 2008; Stelzl et al., 2020). During development, the levels should be given a short name and a short description of the essential elements of the stage (de Bruin et al., 2005).

Each stage is described by a collection of process areas (also capability area, dimension, focus area) (Team CMMI Product, 2006), which can be divided into Components (also process area threads) and Sub-Components (also process area, factor) for more complex domains (de Bruin et al., 2005). A process area represents different facets of a content cohesive area (Hasson and Keeney, 2011). A process area represents a collection of content-related goals (also factor specification (Object Management Group, 2008)). A goal describes the intended condition within a process area that significantly improves this area (Team CMMI Product, 2006). The goal layer is the link between the abstract layers of the MM and the operational assessment (Object Management Group, 2008; Team CMMI Product, 2006). The goals represent a measurable basis and can include practices that exemplify what should be done to achieve the process area and reach the targeted level of maturity (Object Management Group, 2008). Consequently, the assessment elements form the lowest level of a MM (Stelzl et al., 2020).

These hierarchical elements are related to each other through the development path (also maturation path), mapping the goals and practices to the maturity levels. Accordingly, this defines what a mature development state, respectively the status perfect, characterizes and how the path is designed (Mettler and Rohner, 2009). Improvement actions are needed to move along the maturation path, indicating how process areas, goals, and maturity can be achieved (Reeb and Pinnecke, 2021).

After selecting a Delphi variant, a vital aspect of the Delphi study is selecting a suitable expert panel. First, it is crucial to define what constitutes an expert status (Hasson and Keeney, 2011; Skinner et al., 2015) and disclose the criteria for replicability (Skinner et al., 2015). The expert criteria depend on the research objective (Hasson and Keeney, 2011). In the context of MM development, these are related to the MM´s scope, and scholars should relate the expert selection to relevant criteria like the domain (e.g., software development, business process management), audience/ user (e.g., higher management, executives), or function (e.g., descriptive, prescriptive) of the intended MM. Although these criteria are examples, this author argues that the domain must be defined and related to the expert selection criteria, as should the intended central constructs. Furthermore, aspects like the function or the complexity of the intended MM can influence the panel's composition. Thus, scholars need to define whether the panel consists of homogeneous or heterogeneous experts (Gallego and Bueno, 2014). Overall, it is relevant that the selection and composition of the expert panel are reported and related to the research´s objective and relevant characteristics of the MM.

Another element is the number of executed iterations/rounds. The fundamental goal is to reach a consensus, which theoretically can lead to an unlimited number of rounds (Hsu and Sandford, 2007; Torrecilla-Salinas et al., 2019). An absolute consensus may not be achievable, for example, due to opinion divergence among experts (Paré et al., 2013). Furthermore, too many rounds can lead to experts dropping out or a not substantively change of opinion towards the group opinion (Dransfeld et al., 2000; Gallego and Bueno, 2014). Accordingly, the researcher needs to know when to stop the data collection (Schmidt, 1997). Scholars need to define rules for reaching an adequate level of consensus and stopping the iterations (Okoli and Pawlowski, 2004; Paré et al., 2013). These rules must be related to the intended central constructs, what specific characteristics are investigated, and how their quality is measured. Measuring the quality is further related to the central Delphi design element of a statistically processed group response (Rowe and Wright, 1999; Strasser, 2017). The answers of each round must be statistically analyzed, and the results reported to the panel to indicate the panel's opinion. Thus, suitable measurements must be defined for the central constructs elaborated through the Delphi study. The development of a MM is connected to related MMs (new development, further development, model combination, structure transfer, content transfer) (Becker et al., 2009). Hence, it is beneficial to investigate relating MMs, as insights for quality criteria can be found in relating models.

As an additional element, the design of the first round represents a distinctive feature within the Delphi study. Following the Classic Delphi design, the first round starts with a qualitative survey using open-ended questions (Hsu and Sandford, 2007; Skinner et al., 2015). The use of structured questionnaires with pre-selected items is an accepted modification of the Delphi method (Hsu and Sandford, 2007; Skinner et al., 2015). In this case, the questionnaire should be based on an extensive literature review and a rigorous grounding of the included items (Hsu and Sandford, 2007) and the relating MMs in the MM context. Hence, scholars should indicate whether they ground their Delphi study on existing work and how it influenced the first round.

An advantage of the Delphi method is its modifiability, which at the same time presents a crucial challenge to its rigor. Hence, it is relevant that researchers report and justify modifications. Thereby, modifications must be

related to the research objective and the corresponding characteristics of the MM, which should be reported in the MM´s scope.

3. Research Design

Following vom Brocke et al. (2009), a systematic literature analysis was conducted to identify relevant literature. First, the review scope was defined following Paré et al. (2015). The literature search and selection process was designed according to Webster and Watson (2002). Finally, the identified literature was analyzed using deductive content analysis, as suggested by Elo and Kyngäs (2008), based on the above-defined criteria for a rigorous Delphi and MM development methodology.

Defining the scope of a review "is a necessary first step of clarification in any literature review, which bears implications for the later search process" (vom Brocke et al., 2009). Following the categorization of Paré et al. (2015), a critical review is conducted. Hence, representative studies for applying the Delphi method in the context of MM development are analyzed and critically held up against the criteria for a rigorous study.

According to Webster and Watson (2002), the literature search was conducted as a commonly applied approach in the IS domain. There was no restriction regarding the period of the investigation. To identify relevant and representative literature, four central academic databases (Business Source Complete, Emerald Insight, Science Direct, SpringerLink) and two high-quality conference bases (AIS Electronic Library, IEEE Xplore), typical for the IS domain, were accessed. The database SpringerLink was limited to the disciplines "Computer Science" and "Business Management" to include only papers relevant to the IS domain. The query string "maturity model*" AND "Delphi" was used, searched in title, abstract, and attributed keywords. The title, abstract, and complete text of the results were screened regarding their relevance to the defined scope. Furthermore, according to Webster and Watson (2002), the identified literature was scanned by a reference backward and forward search with one iteration to identify further relevant articles. During this step, no other literature was identified. The results of the literature search are summarized in Table 2. Overall, 61 papers were scanned regarding relevance, with 14 suitable for this research's objective.

Table 2: Overview of Search and Selection Process

Database/search step	Number of results	Number of relevant results
Business Source Complete	8	5
Emerald Insight	0	0
Science Direct	12	2
SpringerLink	35	2
AIS Electronic Library	6	3
IEEE Xplore	10	4
Duplicates	n/a	2
Forward & Backward search	n/a	0
Total	61	14

A deductive content analysis following Elo and Kyngäs (2008) was conducted to analyze the literature. A structured coding scheme was developed representing the presented characteristics of the previous section. MAXQDA was used to code the literature following the coding scheme. Finally, the codes were analyzed and compared to the rigorous demands, focusing on the relation of Delphi and MM aspects.

4. Findings

In this section, the results of the analysis are discussed. Firstly, each paper's overall research objective and research design were analyzed. Every study stated the MM´s domain, and it is indicated that preliminary content was used as an input for the Delphi study either directly or in the context of the paper. Furthermore, three of the 14 papers evaluate the Delphi results in the form of a subsequent practical demonstration or evaluation.

In the context of preliminary work, 13 papers used literature data, and one paper used project data, and a distinction can be made between the two goals of Delphi studies, "refine" and "extend". Seven Delphi studies aimed to "refine" a central construct, developed or derived in the preliminary work, to improve it. Six Delphi studies aimed to "extend" the MMs by elaborating on a deeper level of central constructs (e.g., elaborate on

what sub-components are relevant for identified components). The study by Labaka et al. (2019) had both goals by first "refining" a central construct and then "extending" it.

The studies were further analyzed for the four defining characteristics of iterative, anonym, controlled feedback, and statistical aggregation of a Delphi study (Strasser, 2017), regardless of the context of MM development. Only five of the 14 studies reported the four requirements sufficiently. An additional three addressed each of the four requirements at least to a limited extent by presenting the criteria as requirements for the rigor of a Delphi study from a theoretical perspective but not eliciting how they were accomplished. Scholars in Delphi research have called for a more rigorous and variant differentiated approach in the last ten years (Gallego and Bueno, 2014; Paré et al., 2013; Skinner et al., 2015; Strasser, 2017). As eleven of the 14 examined papers were published since 2016, they should have acknowledged this requirement for rigor in scientific research and incorporated it into their research. Nevertheless, only two of these studies met the mandatory requirements for a rigorous Delphi study, and another three reported at least all requirements from a theoretical perspective. Based on these findings, researchers need further guidance on applying the scientific methodological requirements and the design process of a rigorous Delphi study. The requirements of anonymity and feedback are most often not considered, although these two are decisive characteristics of a Delphi study, which raises the question of whether a Delphi method or another approach was used. For example, Schriek et al. (2016) can be stated, as they apply a modification by using face-to-face meetings and not distinguishing their methodological design, e.g., focus group interviews.

As stated in section 2, the author argues that the criteria of the intended central construct(s) and the development basis of the MM must be defined for an adequate Delphi design. These two criteria were focused on regarding the MM development process during the coding process. Although nearly all scholars mention the intended central construct(s), the results confirm that no consistent wording exists throughout the research area, as mentioned in previous studies (Lasrado et al., 2015; Stelzl et al., 2020). Hence, information about the study's central construct(s) must be provided. It is crucial to provide this information as it gives researchers and practitioners an overview of the intended MM structure, how the study contributes to this model design and its useability for their context. Ten of the twelve papers that mention the intended central construct(s) provide enough information to do so. Nevertheless, four of the ten studies leave room for improvement, as it is quite complex to collect the relevant information, and they allow interpretation to some extent (e.g., Labaka et al. (2019)). All papers state the development basis of their MM by presenting preliminary work for the Delphi study. A distinction can be made between new development (use of self-developed MM), further development (use one existing MM), model combination (use a fusion of different MM), structural transfer (use structure of a MM), and content transfer (use content related central construct(s) of a MM) for the use of the preliminary work as a basis for the Delphi study. Although the intended central constructs and developmental basis are mostly reported, only Reyes and Giachetti (2010) (developmental basis) and Bruin and Rosemann (2007), as well as Kerpedzhiev et al. (2021) (central construct), relate them to design decisions of the Delphi study. Furthermore, additional aspects of the MM scope, focusing on a relation to the Delphi study design, were coded and analyzed. However, no paper relates further scope elements to the Delphi study design, and they are thus not further discussed.

Lastly, design aspects of a Delphi study, which are interrelated with the MM are analyzed and an overview is presented in Table 3. Although the focus is on how the MM aspects relate to the design decisions for the study, each aspect is first analyzed independently. Ten studies specify the first-round design, with two using a quantitative design, five using a qualitative design, and three a combination of both designs, whereas none describes the relation to the MM characteristics. Twelve of the fourteen studies report the number of iterations, with five studies stating a consensus criterion. Only de Bruin and Rosemann (2007) relate the iterations and consensus to MM characteristics, particularly the central construct. Furthermore, eleven scholars report the expert criteria and panel composition, with four relating it to Delphi´s research goal(s) and domain. Hence, the relevant design aspects are usually described to some extent but are only occasionally related to the MM characteristics. None of the papers directly reported a Delphi variant. The best examples are:

- George et al. (2020) present that a modified Delphi is used, and from the context, the application of the classical Delphi can be concluded.
- Schriek et al. (2016) indirectly report using a ranking-type Delphi.
- Serral et al. (2020) designed their study similarly to referenced other Delphi studies and, drawn from the context, using a classic Delphi design.

Table 3: Literature Analysis Results on Delphi and MM Aspects Relation

Study	Delphi Criteria	MM Aspects that are related to…				
		Delphi Variant	Expert Panel	Consensus Criterion	First-round Design	Modifications
de Bruin and Rosemann (2007)	sufficiently addressed	not reported	domain & research goal	central construct	reported but not related	central construct
Reyes and Giachetti (2010)	not sufficiently addressed	not reported	reported but not related	reported but not related	reported but not related	/
Stojanov et al. (2015)	sufficiently addressed	not reported	domain	reported but not related	reported but not related	/
Schriek et al. (2016)	not sufficiently addressed	not reported	reported but not related	reported but not related	reported but not related	central construct
Mens and Ravesteyn (2016)	addressed to a limited extent	not reported	reported but not related	reported but not related	reported but not related	/
Radosavljevic et al. (2016)	not sufficiently addressed	not reported	reported but not related	not reported	not reported	/
Karabacak et al. (2016)	sufficiently addressed	not reported	reported but not related	reported but not related	not reported	/
Smits and van Hillegersberg (2017)	not sufficiently addressed	not reported	domain & research goal	reported but not related	reported but not related	/
Nurcahyo et al. (2018)	addressed to a limited extent	not reported	domain	not reported	not reported	/
Labaka et al. (2019)	addressed to a limited extent	not reported	reported but not related	reported but not related	reported but not related	/
Özturan et al. (2019)	not sufficiently addressed	not reported	not reported	reported but not related	not reported	/
George et al. (2020)	sufficiently addressed	not reported	reported but not related	reported but not related	reported but not related	reported but not related
Serral et al. (2020)	not sufficiently addressed	not reported	reported but not related	reported but not related	reported but not related	/
Kerpedzhiev et al. (2021)	sufficiently addressed	not reported	domain	reported but not related	reported but not related	central construct

Nevertheless, nine papers provide more details on the structure and procedure of the Delphi, from which the respective Delphi variant can be presumed. Identifying modifications is problematic, as the Delphi variant is usually not presented and must be based on a presumed variant. Still, some modifications were deemed evident and were therefore coded and analyzed. Modifications were identified within four studies, three of which related these to MM aspects. Schriek et al. (2016) and George et al. (2020) apply face-to-face meetings at some point in their Delphi study. While Schriek et al. (2016) relate this decision to the central construct to gain clarity and a shared understanding of the research objectives and constructs, George et al. (2020) do not justify and

relate this modification. Another critical point that must be considered is that neither paper examines this modification in terms of its impact on rigor (e.g., the impact of face-to-face meetings on anonymity). With the same intention as Schriek et al. (2016), justify Kerpedzhiev et al. (2021) their modification of a two-phase approach with a first phase to develop a common language between experts of different domains. De Bruin and Rosemann (2007) use a different modification, who relate their design decision to the complexity of the central construct and split their Delphi study into several ones with the same design to get a deep insight into each process area. Nevertheless, modifications are not enough supported from a scientific perspective. Thus, the danger of rigor loss in Delphi studies and the associated danger of quality loss and credibility exists in the field of MM development, and future research requires a more rigorous approach.

The results indicate a lack of clarity regarding the methodology, as different aspects are reported, although the general demands, starting points, and goals were similar. In summary, it can be said that a Delphi study for MM development should take various aspects into account to meet the requirement of rigor. Since modifications and adaptability are a significant advantage of the Delphi methodology and will remain relevant in future research, guidelines are needed for rigorously designing, conducting, and presenting the corresponding methodology.

5. Research Design Guidelines

According to the need for rigorously designing a Delphi study for MM development, this section introduces a first framework for a rigorous approach. This framework proposes guidelines, based on the preceding considerations, for researchers to ensure rigor in the design and, consequently, ensure the quality and credibility of the study's results. This paper proposes the general steps (illustrated in Figure 1) that scholars should consider when designing and conducting a Delphi study for MM development. The steps are discussed in detail below, focusing on the Delphi design step concerning the scope of the MM.

Figure 1: Design Steps of a Delphi Study for MM Development

5.1 Define Maturity Model Scope

As the overarching research goal is developing a MM or contributing to a specific part of a MM, scholars must first elaborate on the scope of the intended MM. This paper proposes using an overview of MM aspects for designing and developing constructs by Lasrado et al. (2015) to define relevant aspects for the development process. Although not all aspects of this overview may be relevant to every study, and no single recommendation can be made, it provides an overview of aspects researchers should look at and establishes a common vocabulary. Furthermore, this paper is based on the perception that all necessary aspects are included and, accordingly, recommendations are proposed as to which of these aspects should mandatorily be addressed. The recommendations are based on the analysis results and the theoretical work presented in section 2. First, the focus of the model should be defined. It should be determined whether a general or domain-specific model is to be developed, and the corresponding domain should be specified through practice or theory-based problem presentation. Domain and problem-specific requirements for the intended MM should be identified using literature and empirical data to ensure theoretical and practical relevance. Furthermore, the target audience and level of abstraction should be defined. These should be examined and presented, as they set specific knowledge and characteristics of experts that are needed for adequate development. The aspect constructs is considered crucial, as defining the central construct and its relation to the research goal is central for the later decision of a Delphi variant. Although Lasrado et al. (2015) provide a basis for a common vocabulary of the central construct, scholars should define their understanding of the construct and its position in the architecture of a MM. This ensures a clear understanding for scholars and provides a quick overview of the architectural understanding of the intended MM. This paper proposes utilizing the maturity metamodel by Bley et al. (2020) to define the central construct within a MM architecture. Furthermore, it is recommended to present information on every relevant aspect related to the later design process.

5.2 Define Development Base

As Becker et al. (2009) called for, one step in the MM development process is analyzing related MMs. The domain should be examined concerning related MMs, as all reviewed studies develop/derive initial aspects from the literature. Weaknesses in related MMs can be identified, and initial insights into structure, content and development strategy can be provided. Hence, it is recommended to elaborate elements of/for the central construct and define whether these are validated enough to use as a basis to extend the structural depth of the construct/ MM (e.g., develop sub-components for components) or if they should be refined. When developing initial aspects, one should keep in mind that sufficient information can hardly be achieved deeper than the structural layer of components (de Bruin et al., 2005). This aspect defines which content-related considerations and preliminary work from the literature are used as input for the Delphi study.

5.3 Design Delphi Study

The third step addresses the design of the Delphi study. A Delphi study's four generic and mandatory characteristics must be described. Additional elements must be considered, for which using Strasser´s (2017) taxonomy for Delphi studies is suggested. As stated in section 2 and Table 1, five design elements are identified as critical for MM development and are subsequently discussed.

The choice of the *Delphi variant* is seen as a significant decision (Skinner et al., 2015; Strasser, 2017). This paper perceives the epistemological objective of every Delphi study for MM development as concept development. Thus, relevant Delphi variants are the classical and ranking-type Delphi. In Table 4, a mapping of the two variants regarding their suitability for developing the central constructs is proposed, with "X" indicating a fit. A "(X)" indicates a possible but less probable fit, which was added, as the exact research target may vary within a central construct. The constructs, maturity level, (sub-)component, and goal are primarily analyzed regarding what characterizes them and less regarding their relative importance. In contrast, elaborating the development path is about the relative importance of the subcomponents or goals within a maturity level. As assessment elements and improvement actions were of less interest in the analyzed studies, a claim can only be made from a theoretical perspective. Both variants are of equal relevance, as various assessment elements and improvement measures need to be identified and their characteristics and relative importance explained in more detail.

Table 4: Allocation of Relevant Delphi Variants Regarding Their Suitability for the Development of Different Central Constructs

Delphi Variant	Maturity Level	(Sub-) Component	Goal	Development Path	Assessment Elements	Improvement Actions
Classic Delphi	X	X	X	-	X	X
Ranking-Type Delphi	(X)	(X)	(X)	X	X	X

Additionally, the expert panel must be defined. The author proposes using a knowledge resource nomination worksheet (Delbecq et al., 1975; Okoli and Pawlowski, 2004) to identify and select adequate experts for the panel. Although this process may seem resource-intensive, it is beneficial because selecting appropriate experts is crucial for successful study implementation and reliable and valid results. For complex and diverse domains, a heterogeneous disciplinary panel and different groups of experts for specific components can be helpful (e.g., for multi-dimensional MMs). This may result in several simultaneous Delphi Studies for the intended constructs as modification, as in the study of de Bruin and Rosemann (2007). Independently of the multidisciplinary aspect, it is recommended to use a heterogeneous panel of academics and practitioners. It is vital to define the experts based on criteria relevant to the field and present these criteria and the differences between each group of experts. For defining the expert criteria, domain specifics should be included in this process and the intended auditorium of the MM. Hence, the panel should be related to the domain and complexity of the intended central construct(s). When selecting the experts for the panel, the number of participating experts must be defined (Gallego and Bueno, 2014). No precise information exists in the literature on how many experts are adequate (Paré et al., 2013). Regarding the panel size, various suggestions can be found in the literature, such as not less than seven (Dalkey and Helmer, 1963), between ten and 15 (Linstone et al., 1975), or around 30 (Delbecq et al., 1975). We follow the reommandation by Gallego and Bueno (2014) of ten to 15 experts for homogeneous panels and the recommendation by Linstone et al. (1975) of four to five for each expert group for heterogeneous panels are recommended.

Furthermore, the number of iterations is vital for a Delphi study. It is appropriate to establish rules for reaching an adequate level of consensus and thus end the data collection (Okoli and Pawlowski, 2004; Paré et al., 2013). Defining a consensus criterion as a stopping rule should be related to the central construct and depends on the Delphi type (e.g., typical criteria for ranking type Delphi is Kendell`s W). Additionally, as the consensus criterion is usually defined as a threshold of a statistical value, it can be beneficial to relate this decision to findings of the development base. Furthermore, as the number of rounds increases, the speed of observed convergence between experts decreases, and proceeding over more than three rounds is not appropriate (Linstone et al., 1975). Nevertheless, various recommendations for the number of iterations exist in the literature, ranging from two to ten rounds (Lang, 1995). Recommendations of three rounds (Brooks, 1979; Mulligan, 2002), more than two rounds (Dransfeld et al., 2000; Murphy et al., 1998; Rowe and Wright, 1999) or up to four rounds (Erffmeyer et al., 1986; Hsu and Sandford, 2007) occur most frequently in the literature. Accordingly, a general recommendation of two to four rounds seems to represent the consensus among scholars in the field of Delphi methodology. This discussion leads to further design decisions as researchers must explain which and how aspects can influence the number of iterations. Relating to the research objective (for which construct is consensus intended), scholars must address the issue of what information can be developed and provided in advance to reach consensus with a high degree of probability without providing too much information to overwhelm the experts or bias the results. For example, a greenfield approach may lead to a high number of rounds until a satisfactory consensus is reached, whereas showing all sub-components with elaborated descriptions may influence opinions too much or discourage experts from further participation. Thus, studies need to be set on a reasoned compromise between these two design decisions, as it allows to minimize the Delphi rounds without constraining the experts too much. Another aspect that significantly influences the number of iterations is the overall extent of elements that need to be developed. As previously mentioned, this can be tackled by modifying the study into several ones with the same design.

As mentioned before, the research aim of the Delphi study is related to the development basis, and it can be distinguished between extending and refining. This primarily affects the design of the first round, as the development basis is primarily used to evaluate and refine the theoretical findings (resulting in a predominantly quantitative first round) or obtain the experts' opinions and then synthesize them to compare or expand the theory-based model. As the findings of the analyzed studies show, both approaches can be combined in the first round.

An advantage of the Delphi method is its modifiability, which at the same time presents a crucial challenge to its rigor. Hence, it is of relevance that researchers report and justify modifications. Thereby, modifications must be related to the research objective and the corresponding characteristics of the MM, which should be reported in the MM´s scope. Due to the diversity of applications, further adjustments in the method may be necessary. Paré et al. (2013) advise following at first the recommendations from corresponding literature on the Delphi variant, followed by checking whether corresponding modifications have already been validated in the literature under comparable conditions and if the modification is novel, the impact on the results, the rigor and the study design as well as possibilities to increase validity need to be explored and presented (Paré et al., 2013).

5.4 Conduct Delphi and Define Round Adjustments

Step four defines conducting the Delphi study. First, the entire study design should be reviewed regarding the relations between the research goal(s), the intended MM, the Delphi methodology, and its correct application. The Delphi study can be conducted according to the predetermined design if no adjustment is required. During this step, detailed documentation of feedback or recommendations from experts and taken actions are required. Only the first round can be one hundred percent predesigned. In each subsequent round, the exact design and layout of the questionnaire will depend on the previous round(s) results. Each iteration can focus on a different goal concerning the central construct, distinguishing brainstorming, narrowing-down, validation, and ranking (Paré et al., 2013). Hence, scholars should report what specific goal was followed each round, what results were achieved, and how they relate to the design of the next iteration. Besides designing the specific rounds, it is essential to provide the panel with feedback on the results in every iteration, resulting in the presented process in Figure 2 for each iteration.

Figure 2: Process of a Delphi Round

5.5 Report Results

The last step comprises presenting results to researchers and practitioners for dissemination. Scholars should indicate how and what should be done in the further development process of the MM. As the proceeding steps of the model development are about the practical transfer and application (Becker et al., 2009; de Bruin et al., 2005; Mettler, 2010), action recommendations and possible hurdles concerning the study's central construct should be indicated. Furthermore, the used Delphi procedure, including modifications, should be reflected and insights for future applications given.

6. Conclusion

This paper examines the rigor of the MM development process using a Delphi approach in the IS domain. First, both methodological approaches were examined in their approach and rigorous demands focusing on their interaction. A systematic literature analysis was conducted and analyzed with a deductive content analysis based on these elaborations. The findings show no clarity on the design aspects, and their methodological interactions exist, and a clarification is needed. Hence, guidelines to develop a Delphi study for MM development were elaborated and presented. This is the first attempt towards a rigorous approach but does not provide a step-by-step approach, as Delphi studies are characterized by their modifiability, and different Delphi types and a variety of MM constructs exist that can be a goal of the development process.

The developed guidelines integrate the different methodological aspects of the MM development and the Delphi methodology. It represents an orientation framework for researchers to design a Delphi study for MM development from a rigorous perspective. To this end, it provides scholars with an overview of aspects that must be considered and in which order they should be addressed. Furthermore, it is explained how the aspects are linked to one another and what decisions regarding opposing aspects in the design process can occur. Hence, scholars can use the guidelines to design a rigorous Delphi study for MM development. Although this paper focuses on the MM context, the general design approach and decisions are transferable and applicable to other domains. Hence, this research contributes to the Delphi literature by providing insights into how relevant elements should be addressed in the designing process of a Delphi Study. Thus, scholars should investigate how the presented guidelines must be adapted for other domains.

The research areas of MM and Delphi are extensive and diverse, leading to numerous existing and possible approaches. It should be noted that this paper only analyzes literature that indicates that a Delphi method to develop MM aspects was applied. Scholars that might have applied this method and did not indicate the methodology could not be included and might have applied it rigorously. This paper does not intend to give a solution that fits every possible scenario of combining these two methods but instead suggests aspects and guidelines to maneuver through the design process. This research gives a foundation and orientation for researchers that want to use the Delphi method for MM development. Furthermore, researchers are encouraged to expand this work on missing aspects and more concrete application scenarios, such as the sole application of a ranking-type Delphi to develop a maturation path.

Acknowledgements

The Article Processing Charges (APC) were funded by the joint publication fund of the TU Dresden, the Medical Faculty Carl Gustav Carus, and the SLUB Dresden.

References

Becker, J., Knackstedt, R. and Pöppelbuß, J. (2009) 'Developing Maturity Models for IT Management', Business & Information Systems Engineering, vol. 1, no. 3, pp. 213–222.

Brooks, K. W. (1979) 'Delphi technique: Expanding applications', North Central Association Quarterly, vol. 53, no. 3, pp. 377–385.

Carvalho, J. V., Rocha, Á., van de Wetering, R. and Abreu, A. (2019) 'A Maturity model for hospital information systems', Journal of Business Research, vol. 94, pp. 388–399.

Dalkey, N. and Helmer, O. (1963) 'An experimental application of the Delphi method to the use of experts', Management science, vol. 9, no. 3, pp. 458–467.

de Bruin, T. and Rosemann, M. (2007) 'Using the Delphi technique to identify BPM capability areas', ACIS 2007 Proceedings, 42 [Online]. Available at http://aisel.aisnet.org/acis2007/42.

de Bruin, T., Rosemann, M., Freeze, R. and Kaulkarni, U. (2005) 'Understanding the main phases of developing a maturity assessment model', Australasian Conference on Information Systems (ACIS), pp. 8–19.

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, Scott, Foresman.

Dransfeld, H., Pemberton, J. and Jacobs, G. (2000) 'Quantifying Weighted Expert Opinion', Technological Forecasting and Social Change, vol. 63, no. 1, pp. 81–90.

Elo, S. and Kyngäs, H. (2008) 'The qualitative content analysis process', Journal of advanced nursing, vol. 62, no. 1, pp. 107–115.

Erffmeyer, R. C., Erffmeyer, E. S. and Lane, I. M. (1986) 'The Delphi Technique: An Empirical Evaluation of the Optimal Number of Rounds', Group & Organization Studies, vol. 11, 1-2, pp. 120–128.

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, vol. 26, no. 9, pp. 987–999.

George, A., Schmitz, K. and Storey, V. C. (2020) 'A Framework for Building Mature Business Intelligence and Analytics in Organizations', Journal of Database Management, vol. 31, no. 3, pp. 14–39.

Gupta, U. G. and Clarke, R. E. (1996) 'Theory and applications of the Delphi technique: A bibliography (1975–1994)', Technological Forecasting and Social Change, vol. 53, no. 2, pp. 185–211.

Hasson, F. and Keeney, S. (2011) 'Enhancing rigour in the Delphi technique research', Technological Forecasting and Social Change, vol. 78, no. 9, pp. 1695–1704.

Hsu, C.-C. and Sandford, B. A. (2007) 'The Delphi Technique: Making Sense of Consensus', Practical Assessment, Research, and Evaluation, no. 12.

Karabacak, B., Yildirim, S. O. and Baykal, N. (2016) 'A vulnerability-driven cyber security maturity model for measuring national critical infrastructure protection preparedness', International Journal of Critical Infrastructure Protection, vol. 15, pp. 47–59.

Kerpedzhiev, G. D., König, U. M., Röglinger, M. and Rosemann, M. (2021) 'An Exploration into Future Business Process Management Capabilities in View of Digitalization', Business & Information Systems Engineering, vol. 63, no. 2, pp. 83–96.

Labaka, L., Maraña, P., Giménez, R. and Hernantes, J. (2019) 'Defining the roadmap towards city resilience', Technological Forecasting and Social Change, vol. 146, pp. 281–296.

Lahrmann, G., Marx, F., Mettler, T., Winter, R. and Wortmann, F. (2011) 'Inductive Design of Maturity Models: Applying the Rasch Algorithm for Design Science Research', in Jain, H., Sinha, A. P. and Vitharana, P. (eds) Service-Oriented Perspectives in Design Science Research, Berlin, Heidelberg, Springer Berlin Heidelberg, pp. 176–191.

Lang, T. (1995) 'An overview of four futures methodologies', Manoa Journal of Fried and Half-Fried Ideas, vol. 7, pp. 1–43.

Lasrado, L., Vatrapu, R., Karsgaard, H. B. and Kjaer, J. F. (2017) 'Towards sustainable design for maturity measurement marketplace', International SERIES on Information Systems and Management in Creative eMedia (CreMedia), 2016/2, pp. 25–33.

Lasrado, L. A., Vatrapu, R. and Andersen, K. N. (2015) 'Maturity Models Development in IS Research: A Literature Review', Selected Papers of the IRIS, no. 6 [Online]. DOI: 10.13140/RG.2.1.3046.3209.

Linstone, H. A., Turoff, M. and others (1975) The delphi method, Addison-Wesley Reading, MA.

McKenna, H. P. (1994) 'The Delphi technique: a worthwhile research approach for nursing?', Journal of advanced nursing, vol. 19, no. 6, pp. 1221–1225.

Mens, J. and Ravesteyn, P. (2016) 'Using the Delphi Method to Identify Hospital-Specific Business Process Management Capabilities in The Netherlands', Bled eConference, 22.

Mettler, T. (2010) Supply management im Krankenhaus: Konstruktion und Evaluation eines konfigurierbaren Reifegradmodells zur zielgerichteten Gestaltung, Sierke Göttingen.

Mettler, T. (2011) 'Maturity assessment models: a design science research approach', International Journal of Society Systems Science, vol. 3, 1-2, pp. 81–98.

Mettler, T. and Rohner, P. (2009) 'Situational maturity models as instrumental artifacts for organizational design', Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology - DESRIST '09. Philadelphia, Pennsylvania, 07.05.2009 - 08.05.2009. New York, New York, USA, ACM Press, pp. 1–9.

Mulligan, P. (2002) 'Specification of a capability-based IT classification framework', Information & Management, vol. 39, no. 8, pp. 647–658.

Murphy, M. K., Black, N. A., Lamping, D. L., McKee, C. M., Sanderson, C. F., Askham, J. and Marteau, T. (1998) 'Consensus development methods, and their use in clinical guideline development', Health technology assessment (Winchester, England), vol. 2, no. 3, i-iv, 1-88.

Nurcahyo, R., Darmawan, D., Jannis, Y., Kurniati, A. and Habiburrahman, M. (2018) 'Maintenance Planning Key Process Area: Case Study at Oil Gas Industry in Indonesia', 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). Bangkok, 16.12.2018 - 19.12.2018, IEEE, pp. 1704–1708.

Object Management Group (2008) Business Process Modeling Notation (BPMN): Version 1.0 [Online]. Available at https://www.omg.org/spec/BPMM/1.0/PDF (Accessed 18 September 2021).

Okoli, C. and Pawlowski, S. D. (2004) 'The Delphi method as a research tool: an example, design considerations and applications', Information & Management, vol. 42, no. 1, pp. 15–29.

Özturan, M., Mutlutürk, M., Çeken, B. and Sarı, B. (2019) 'Evaluating the information systems integration maturity level of travel agencies', Information Technology & Tourism, vol. 21, no. 2, pp. 237–257.

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, vol. 50, no. 5, pp. 207–217.

Pereira, R. and Serrano, J. (2020) 'A review of methods used on IT maturity models development: A systematic literature review and a critical analysis', Journal of Information Technology, vol. 35, no. 2, pp. 161–178.

Radosavljevic, M., Barac, N., Jankovic-Milic, V. and Andjelkovic, A. (2016) 'Supply chain management maturity assessment: challenges of the enterprises in Serbia', Journal of Business Economics and Management, vol. 17, no. 6, pp. 848–864.

Rauch, W. (1979) 'The decision Delphi', Technological Forecasting and Social Change, vol. 15, no. 3, pp. 159–169.

Reeb, S. and Pinnecke, M. (2021) 'Continuous Improvement with Maturity Models – A Procedural Model for Action Derivation', 25th Pacific Asia Conference on Information Systems, PACIS 2020, Dubai, UAE, July 12-14, 2021.

Reyes, H. G. and Giachetti, R. (2010) 'Using experts to develop a supply chain maturity model in Mexico', Supply Chain Management: An International Journal, vol. 15, no. 6, pp. 415–424.

Rowe, G. and Wright, G. (1999) 'The Delphi technique as a forecasting tool: issues and analysis', International Journal of Forecasting, vol. 15, no. 4, pp. 353–375.

Schmidt, R. C. (1997) 'Managing Delphi Surveys Using Nonparametric Statistical Techniques', Decision Sciences, vol. 28, no. 3, pp. 763–774.

Schriek, M., Türetken, O. and Kaymak, U. (2016) 'A Maturity Model for Care pathways', 24th European Conference on Information Systems (ECIS). Istanbul, Turkey, June 12-15, 127.

Serral, E., Stede, C. V. and Hasic, F. (2020) 'Leveraging IoT in Retail Industry: A Maturity Model', 2020 IEEE 22nd Conference on Business Informatics (CBI). Antwerp, Belgium, 22.06.2020 - 24.06.2020, IEEE, pp. 114–123.

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 Association for Information Systems, vol. 37, no. 1, pp. 31–63.

Smits, D. and van Hillegersberg, J. (2017) 'The development of a hard and soft IT governance assessment instrument', Procedia Computer Science, vol. 121, pp. 47–54.

Stelzl, K., Röglinger, M. and Wyrtki, K. (2020) 'Building an ambidextrous organization: a maturity model for organizational ambidexterity', Business Research, vol. 13, no. 3, pp. 1203–1230.

Stojanov, I., Turetken, O. and Trienekens, J. J. (2015) 'A Maturity Model for Scaling Agile Development', in 41st Euromicro Conference, pp. 446–453.

Strasser, A. (2017) 'Delphi method variants in information systems research: Taxonomy development and application', Electronic Journal of Business Research Methods, vol. 15, no. 2, pp120-133.

Team CMMI Product (2006) 'CMMI for Development, version 1.2'.

Torrecilla-Salinas, C. J., Troyer, O. de, Escalona, M. J. and Mejías, M. (2019) 'A Delphi-based expert judgment method applied to the validation of a mature Agile framework for Web development projects', Information Technology and Management, vol. 20, no. 1, pp. 9–40.

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R. and Cleven, A. (2009) 'Reconstructing the giant: On the importance of rigour in documenting the literature search process', ECIS.