

AN EXPLORATORY DESIGN SCIENCE STUDY ON THEORY TESTING USING CROWDSOURCING

by

IJEOMA USHAKA

A thesis

submitted to the Victoria University of Wellington

in fulfilment of the requirements for the degree of

Doctor of Philosophy

in Information Systems

Victoria University of Wellington
2019

General Abstract

Theory in Information Systems (IS) is very important to the development of the field. Theory building, and theory testing seeks to accumulate knowledge about the relationships between people and technology. Testing theory can be difficult to accomplish, especially when it involves humans, a diversity of methods and sources, multiple experiments, large data sets, and careful tuning of conditions and instruments. Crowdsourcing is a strategy supporting the distribution of activities to crowd workers, which suggests that it may be used to support theory testing.

This exploratory study seeks to analyse the adoption of crowdsourcing in theory testing, and to develop guidance for researchers to instantiate the strategy in their research projects.

The study adopts the design science research paradigm to explore incorporating the crowdsourcing strategy in theory testing, and to evaluate its viability and utility. According to the principles of design science research, the study is structured around the construction of several interconnected IS artefacts: 1) a conceptual framework articulating the main principles of theory testing; 2) a pattern model of theory testing, which codifies existing research approaches to theory testing; and 3) a decision tool, which codifies guidelines for researchers making decisions on which research activities to crowdsource.

In order to build the conceptual framework and pattern model, the study conducts a systematic review of theory testing in the IS domain. Both the conceptual framework and pattern model are then operationalized in the decision tool. The utility of the various artefacts is then assessed with the participation of research practitioners.

This study is relevant because it synthesizes knowledge about theory testing, builds innovative artefacts supporting the adoption of crowdsourcing in theory testing, helps academic researchers understanding the theory testing process, and enables them to adopt crowdsourcing for theory testing.

Keywords

Theory building; Theory testing; Crowdsourcing; Design Science research.

Acknowledgements

The completion of this dissertation would not have been possible without the incessant support I received from numerous persons, and I would like to express my profound gratitude to them.

Firstly, I bow my knees in acknowledgement and appreciation to my Heavenly Father, of whom the whole family in heaven and earth is named, for His unfailing mercies and love to me throughout this phase of my life.

I would also like to acknowledge and thank my supervisors, Associate Professor Pedro Antunes and Dr David Johnstone for their invaluable support and guidance throughout my PhD journey. For opening your doors to me whenever I want an audience and chat with you, for the immense knowledge and ideas given, contributing to the completion of this work.

A special thanks to the school of Information Management (SIM) academics, staff and administrative team for providing the necessary support needed for researching, tutoring and marking. Special thanks to Usha Varatharaju and Anette Klaassen for your invaluable support, encouragement, and friendship throughout my study period, especially during the birth of my son. My sincere thanks to Jean Grant for your material support to my son. My sincere appreciation to all PhD students who participated in my study, providing valuable insight and feedback to my research topic. Thank you for your support and encouragement, the fun times we had and all cohort meetings.

Finally, I would like to thank my family for their support, prayers, sacrifices and understanding throughout my study. Thanks to my husband, Boniface Ushaka, for your love, assistance, support, encouragement and motivation and my son Daniels, for brightening my world. Thanks to Arise church community and all my friends, for encouragement, adding fun and beauty to my life and showing me what it really means to care.

Table of Contents

General Abstract	3
Acknowledgements	4
List of Tables	9
List of Figures.....	10
The Original papers.....	12
1 Introduction	
1.1 Research Background.....	18
1.2 Research Problem.....	20
1.3 Research Methodology.....	21
1.3.1 Design Science Research	17
1.4 Research Framework	22
1.4.1 Research Questions	25
1.4.2 Research Objectives.....	26
1.4.3 Research Contribution	26
1.5 Research Process in Detail	28
1.5.1 Build Activities.....	28
1.5.2 Evaluation	34
1.5.3 Research Checklist	35
1.6 Outline of the Manuscript.....	36
1.7 References	38
2 On The Adoption of Crowdsourcing for Theory Testing (Paper 1)	
2.1 Introduction.....	44
2.2 Related Work.....	46
2.2.1 Focus on Theory with Testable Propositions	47
2.2.2 Considering Crowdsourcing in Theory Testing	50
2.3. Literature Review	52
2.4 Further Analysis and Discussion	60
2.5 Concluding Remarks	62
2.6 References	63
3 Towards The Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: Conceptual Framework and Model (Paper 2)	
3.1 Introduction.....	70
3.2 Approach	72

3.2.1 Development Method.....	73
3.2.2 Evaluation Method	76
3.3 Development Steps	76
3.3.1 Conceptual Framework.....	76
3.3.2 Model.....	78
3.4 Problem Frame Assessment	80
3.5 Results	82
3.5.1 Familiarity with Theory Testing (E1).....	82
3.5.2 Familiarity with Model Constructs (E2).....	83
3.5.3 Selection of Activities that could be Crowdsourced (E3)....	84
3.5.4 Problems and Features (E4)	84
3.5.5 Feedback from the Interviews	85
3.6 Discussion and Conclusions.....	85
3.7 References	87
4 Patterns of Theory Testing with Human Subjects: A Design Science Perspective (Paper 3)	
4.1 Introduction.....	90
4.2 Research Approach.....	92
4.3 Conceptual Framework	94
4.4 Pattern Model.....	99
4.4.1 Review Procedure.....	99
4.4.2 Pattern Theory	102
4.4.3 Pattern System	105
4.5 Evaluation.....	108
4.5.1 Understanding of Theory Generation and Its Cycles (E1).	110
4.5.2 Understanding of Pattern Model (E2).....	112
4.5.3 Understanding of the Elements Defined by the Pattern Model (E3).....	113
4.5.4 Usefulness of Pattern Model (E4)	113
4.6 Discussion and Conclusions.....	114
4.7 References	116
5 Towards the Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: Analytical Framework Design (Paper 4)	
5.1 Introduction.....	120
5.2 Research Background.....	122

5.2.1 Crowdsourcing Requirements in Respect to Theory Testing	124
5.3 Research Method	125
5.4 Decision Support Crowdsourcing Analytical Framework	127
5.4.1 Crowdsourcing Attributes	128
5.4.2 Analytic Framework Development	131
5.5 Justification by Applying the Procedure: Example Cases.....	133
5.6 Discussion and Recommendations	141
5.7 Conclusion.....	143
5.8 References	144
6 Analyzing the Adoption of Crowdsourcing in the Whole Theory Testing Lifecycle (Paper 5)	
6.1 Introduction.....	148
6.2 Theoretical Background	153
6.3 Research Approach.....	155
6.3.1 Artefact Justification And Evaluation	158
6.3.2 Card Sorting Method.....	158
6.4 Conceptual Framework of Theory Testing.....	160
6.5 Pattern Model.....	163
6.5.1 Literature Review	163
6.5.2 Analytical Lens.....	166
6.5.3 Applying the Analytical Lens	168
6.5.4 Model Construction Using Theory Testing Patterns.....	169
6.6 Method for Assessing what Theory Testing Activities can be Crowdsourced	172
6.7 Examples.....	176
6.8 Evaluation.....	179
6.8.1 Utility of the Conceptual Framework of Theory Testing. ...	179
6.8.2 Utility of the Pattern Model of Theory Testing Activities... ..	180
6.8.3 Utility of the Method for Assessing what Theory Testing Activities can be Crowdsourced.	181
6.9 Discussion	182
6.9.1 Limitations of this Study	185
6.10 Conclusions.....	185
6.11 References	186

7	Design and Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: A Design Science Perspective (Paper 6)	
7.1	Introduction.....	200
7.2	Research Context.....	202
7.2.1	Research Background and Goals	203
7.2.2	Decision Support Systems View	205
7.3	Research Method	206
7.4	Dss Artefact.....	210
7.4.1	Tool's Conceptual Design	210
7.4.2	Tool Development.....	210
7.5	Evaluation.....	217
7.5.1	Usefulness of the DSS in Decision Making (E1)	219
7.5.2	Understanding of the Intent of the DSS (E2).....	220
7.5.3	Applicability of the DSS (E3).....	220
7.6	Discussion and Conclusion	221
7.7	References	223
8	Contributions and Conclusions	
8.1	Discussion	228
8.1.1	Contributed Artefacts	229
8.1.2	Contributions to Theory.....	233
8.1.3	Contributions to Practice.....	236
8.1.4	Limitations Of The Research.....	237
8.1.5	Future Work	238
8.2	Concluding Remarks	238
8.3	References	246
Appendices		260
Appendix A.	Theory Testing Coding Forms.....	260
Appendix B.	Ethics Approval	379

List of Tables

Table 1.1.	Design Science Contribution Types derived from (Gregor & Hevner, 2013)	28
Table 1.2.	Summary of Research Activities Aligning to Papers	32
Table 1.3.	Evaluation Methods	34
Table 1.4	Design science research question map/checklist adapted from (Hevner & Chatterjee, 2010; Hevner et al., 2004)	35
Table 2.1.	Crowdsourcing Concepts	50
Table 2.2.	Crowdsourcing requirements	53
Table 2.3.	Summary of Literature Review	55
Table 2.4.	Crowdsourcing Activities and Its Examples	59
Table 2.5.	Crowdsourcing Pathways	61
Table 2.6.	Crowdsourcing Activities	61
Table 3.1.	Familiarity with Theory Testing	83
Table 3.2.	Top activities identified by the participants as able to CS	84
Table 4.1.	Gregor's classification of theory (, 2006)	95
Table 4.2.	Summary of articles after selection and filtering stages	100
Table 6.1.	Gregor's Classification of Theory (, 2006)	154
Table 6.2.	Selection of Articles for Review	164
Table 8.1.	Summary of research activity and results answering research questions	231
Table 8.3.	Summary of research objectives and contributions	242

List of Figures

Figure 1.1.	Research Framework (Adapted from (Hevner et al., 2004)	23
Figure 1.2.	Design Science Research Knowledge Base (& Hevner, 2013)	27
Figure 1.3.	The Research Process	29
Figure 1.4.	The Research Outline	31
Figure 2.1.	Steps to theory testing, adapted from (Bitektine, 2008; De Vaus, 2013; MacKenzie et al., 2011; Niederman & March, 2015)	49
Figure 2.2.	Sample of Coding Form	56
Figure 2.3.	Activities and pattern systems for theory testing (* indicates the activity can be crowdsourced)	57
Figure 2.4.	Flowchart for Requirement Check	58
Figure 3.1.	The adapted method based on SCOA and DSS DSR	74
Figure 3.2.	Conceptual framework, adapted from (Bitektine, 2007; De Vaus, 2013; MacKenzie et al., 2011; Niederman & March, 2015)	77
Figure 3.3.	Model showing theory testing patterns and activities (* means the activity can be crowdsourced)	79
Figure 3.4.	Participants sorting cards	81
Figure 3.5.	Frequency distribution of problems with theory testing identified by participants	84
Figure 3.6.	Frequency distribution of DSS features prioritized by the participants	85
Figure 4.1.	Research approach	93
Figure 4.2.	Conceptual framework of theory testing	96
Figure 4.3.	Fragment of the data extraction form used in the review	102
Figure 4.4.	Synthesis of the literature review showing theory testing patterns. Patterns are divided between data collection and data analysis. Primary activities are presented with dark background, while secondary activities have a clear background.	106
Figure 4.5.	Card sorting exercise	110
Figure 4.6.	Card sorting exercise result showing Framework Understability	111
Figure 4.7.	Card sorting exercise result showing Model Understability.....	112
Figure 4.8.	Card sorting exercise result showing Model Utility.....	113
Figure 5.1.	Build and evaluate activities within a DSR method (Adapted from (Peffers et al., 2007)	127
Figure 5.2.	Crowdsourcing Attributes	129

Figure 5.3.	Feasibility check that determines if a function can be crowdsourced (A V-branch corresponds to a logical AND while a U-branch corresponds to an OR)	132
Figure 5.4.	Theory testing patterns (Adapted from (Enwereuzo et al., 2017))	133
Figure 6.1.	Overview of the research	156
Figure 6.2.	Card sorting exercises	160
Figure 6.3.	Conceptual framework of theory testing	161
Figure 6.4.	Fragment of the data extraction form used in the literature review	166
Figure 6.5.	Synthesis of the literature review showing theory testing patterns	169
Figure 6.6.	Crowdsourcing Criteria	172
Figure 6.7.	Template for Assessing if a Pattern can be Crowdsourced	176
Figure 6.8.	Card sorting exercise result showing Model Utility	181
Figure 6.9.	Card sorting exercise result showing Template Utility	182
Figure 7.1.	Research framework	203
Figure 7.2.	Research Framework for DSS Tool (Adapted from (Peffers et al., 2007))	207
Figure 7.3.	Build and Evaluation Activities (Adapted from (Sonnenberg & vom Brocke, 2011))	207
Figure 7.4.	Tool's Conceptual Design (adapted from (Holsapple, 2008))	211
Figure 7.5.	Template-based tool on theory testing and CS	212
Figure 7.6.	Crowdsourcing criteria	214
Figure 7.7.	User Interface-Based Prototype	216
Figure 7.8.	Card sorting exercise	218
Figure 8.1.	Interrelated yet different artefact outcomes	229
Figure 8.2.	Research design products (adapted from ((Baskerville et al., 2018; & Hevner, 2013))	230

The Original Papers

Paper 1:

Enwereuzo, I., Antunes, P., & Johnstone, D. (2018). On the adoption of crowdsourcing for theory testing. In *European Conference on Information Systems (ECIS), Portsmouth, UK*.

Abstract: This paper examines the possibilities of using the crowdsourcing strategy for theory testing. We first analyze the relationships between theory building and theory testing activities. Then, based on a systematic review of 248 papers published in MISQ, we characterize the intents and pattern systems of activities that have been used for theory testing. Finally, we ascertain which activities can be crowdsourced or not and pinpoint a set of pathways supporting partial and total crowdsourcing. The obtained results show that a large number of activities related to data gathering can be crowdsourced and that a number of intents have viable pathways supporting partial crowdsourcing.

Paper 2:

Enwereuzo, I., Antunes, P., & Johnstone, D. (2017). Towards the Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: Conceptual Framework and Model. *Paper presented at ACIS*.

Abstract: The information systems (IS) discipline has not accorded the same attention to theory testing as it has to theory building. Further, crowdsourcing presents rich opportunities for the theory testing process that have not been fully explored. This paper builds on previous work, employing a design science research (DSR) paradigm in order to develop a decision support system artefact that will help early career researchers identify viable theory testing approaches, and how crowdsourcing can help facilitate the testing process. As part of the DSR build/evaluate cycle, this paper presents a conceptual framework and model of theory testing in IS, and the problem frame in which they are situated is evaluated using Schön's theory of reflective practice and problem/solution framing. Data collected from PhD students revealed an

incomplete level of knowledge of theory testing, and a lack of awareness of the possibilities provided by adopting a crowdsourcing strategy.

Paper 3:

Enwereuzo, I., Antunes, P., & Johnstone, D. (2019). Patterns of Testing Theory with Human Subjects: A Design Science Perspective. *Paper Presented at AMCIS*.

Abstract: This paper addresses the challenges of theory testing. Adopting the design science paradigm, we address this challenge by delimiting the line that separates theory building from theory testing by conceptually characterizing its essential aspects: the relationships between humans, organizations and technology, which led to the development of a conceptual framework for theory testing. Practically, the methods, processes, instruments, and tools needed to acquire and analyze data for theory testing is considered, leading to the development of a pattern system. Card sorting was used as an evaluation strategy for the conceptual framework and the pattern model.

Paper 4:

Enwereuzo, I., Antunes, P., & Johnstone, D. (2019). Towards the Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: Analytical Framework Design. *Paper Presented at AMCIS*.

Abstract: Crowdsourcing can be an adopted strategy for researchers where tasks are distributed to internet users to harness different forms of data, which adds to the reliability and validity of the research process. As theory testing is an essential part of the research process, involving more of operationalization than conceptualization, of which different forms of activities that most times needs input from different and diverse participants is required, the need to consider what theory testing activities can be crowdsourced is paramount. Adopting a design science paradigm to manage this challenge, we design an analytic framework which comprises of important attributes that need to be considered if crowdsourcing is to be used for any of the theory testing activities. The framework which was justified by the use of sample cases gives us an insight into what attributes make such activity to be easy to crowdsource, difficult to

crowdsource or impossible to crowdsource. The value of this artefact lies in its capacity to help researchers making complex design decisions, in this case regarding how they can utilize crowdsourcing to their advantage.

Paper 5:

Enwereuzo, I., Antunes, P., Johnstone, D., & De Vreede, G. J. (2019). Analyzing the Adoption of Crowdsourcing in the Whole Theory Testing Lifecycle. (*Journal article yet to be submitted*).

Abstract. We explore how researchers may adopt the crowdsourcing model in the whole theory testing lifecycle, which could bring a crowd to execute various research project activities ranging from data collection to management, analysis, and quality control. Testing theory is a fundamental component of theory building, which makes this a relevant research problem, in particular in the information systems field where theory building has a fundamental role in shaping the identity of the field. We adopt a design science research perspective over the problem. In particular, we develop three artefacts: a conceptual framework of theory testing, a pattern model of theory testing activities, and a crowdsourcing template to help decision making on which research activities to crowdsource. This study provides contributions to both research and practice. Regarding the former, we contribute artefacts supporting the development of new approaches to test – and ultimately develop – theory. Considering the latter, the developed artefacts offer opportunities to manage the theory testing lifecycle, which in turn may foster new research methods.

Paper 6:

Enwereuzo, I., Antunes, P., Johnstone, D., & De Vreede, G. J. (2019). Design and Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: A Design Science Perspective. *Paper Presented at PACIS*.

Abstract. The integration of crowdsourcing in behavioral research in the IS field offers several advantages and opportunities. This paper builds on prior study, employing a design science research (DSR) paradigm to design, develop and evaluate a tool that assists researchers in adopting crowdsourcing when testing theory about behavioral phenomena. The proposed tool is based on an extensive review of literature on how

theory has been tested, and a pattern model that standardizes extracted concepts, activities, processes and relationships into patterns. We discuss the architecture of the proposed tool and present two prototypes, one used for knowledge articulation by representing, extracting, organizing and acting on relevant information and the other on decision making and recommendation for the tool users. Evaluation results show the applicability and utility of the tool.

As the first author, Ijeoma Ushaka has taken the lead in writing these papers with editing done by the co-authors.

CHAPTER 1

1 INTRODUCTION

1.1 RESEARCH BACKGROUND

Theory in Information Systems is defined as “abstract entities that aim to describe, explain, and enhance understanding of the world and, in some cases, to provide predictions of what will happen in the future and to give a basis for intervention and action” (Gregor, 2006). The importance of theory is undeniable. Theory acts as an educational device that can raise consciousness about a specific set of concepts. Theory helps scientists to describe and explain a phenomenon, process or sequence of events. Theory also helps scholars to explain the complexities of the empirical world by providing a linguistic tool for organizing it (Colquitt & Zapata-Phelan, 2007).

There are two general forms of validity of theoretical considerations: formative and summative (Lee & Hubona, 2009). The former deals with the process by which a theory is formed or built, e.g. through the interconnection of theory building and theory testing. Summative validity deals with how theory fits the real world i.e. how the data fits the hypothesis. Theory building is an ongoing process of producing (observe), confirming (experience, think about), applying (understand) and adapting (act) theory (Lynham, 2002). It also involves the development of new ideas and concepts, and the construction of conceptual frameworks, new methods and models (Nunamaker Jr et al., 1990). Theoretical propositions are hereafter defined after the development of such concepts.

After theoretical propositions have been defined, the researcher becomes concerned with determining whether those propositions have any connections with the empirical world (Mueller & Urbach, 2013), which can be done by testing such theory. The importance of theory testing has been increasingly emphasized over the years (Campbell, 1990; Colquitt & Zapata-Phelan, 2007). Theory building advances theoretical understanding while theory testing substantiates such advancement. This research is focused on IS researchers who need to do theory testing.

There are many challenges that researchers face when testing theory (Ashton, 2014; Nel et al., 2017). Several researchers have given recommendations and support to

overcome these challenges over the years, but as technology advances, there arise other challenges in different facets. Some of these challenges include physical safety (Munro et al., 2009), maintaining boundaries between researchers and participants, participants eligibility and willingness to participate (Green & Thorogood, 2018; Holden et al., 2015), confidentiality of information (Holden et al., 2015; Nel et al., 2017), analysing large volume of data, ensuring validity (Holden et al., 2015), recruiting participants, logistic challenges like difficulties with transportation, travel time, cost etc. (Israel et al., 2006; Pierce et al., 2012), and data reliability (Adzeh, 2014).

Support is needed to address these challenges. Therefore, we embark on considering the adoption of crowdsourcing in the theory testing process. In particular, **we explore if the nature and advantages of crowdsourcing could be applied for testing theories in IS.**

The term crowdsourcing was coined and made popular by the Wired journalist and Northeastern University journalism professor Jeff Howe. In an article for Wired magazine, Jeff Howe defined crowdsourcing as “taking a task or function once performed traditionally by an employee of a company or institution and outsourcing it to a large group of people in the form of an open call” (Howe, 2006). The open call is sent to a pool of undefined/heterogenous people, often with little or no restrictions on who may respond (Schenk & Guittard, 2011; Whitla, 2009; Yuen et al., 2011; Zhao & Zhu, 2012). This can be done both to exploit the aggregated capacity of labor and to harness the hidden or undiscovered talents that may be buried in the crowd (Parvanta et al., 2013).

The potential advantages of using crowdsourcing in research can be enormous. Crowdsourcing can provide researchers with a huge amount of labor and expertise in a short period of time (Conley & Tosti-Kharas, 2014). For example, a researcher could solicit codes for sampling text using CS. The researcher could analyse the responses and triangulate with their own coding to identify any changes to the coding schema or task itself, whereas these iterations would be costly and time consuming using human coders recruited through traditional means (Morris, 1994). Crowdsourcing may reduce the cost of completing a research task, by providing more agile use of resources and removing administrative work. Using

crowdsourcing for research can also provide access to a large pool of experimental data (Behrend et al., 2011), access to specialized knowledge (Conley & Tosti-Kharas, 2014), such as evaluating the output of existing machine learning or artificial intelligence algorithms (Callison-Burch, 2009), access to organizational context (Brabham, 2008), repeatability and procuring the experimental subjects (members of the crowd) from a large pool of human actors (Brabham, 2008; Ford et al., 2015; Smith et al., 2013). This therefore leads to the research problem to be investigated.

1.2 RESEARCH PROBLEM

Challenges experienced by researchers, especially Ph.D. students, often leads to delays, which is often seen or experienced during the theory testing phase. This is because, it often takes less time to conceptualize a phenomenon than to test it, mainly due to lack/insufficient resources.

These delays are most times detrimental to the research process, especially if results from such study are time-based. Some delays may result in out-of-date data, in some cases, additional funding is needed and delays in the potential benefits of such work. These challenges are problems that needs supports to mitigate them (Sarkar, 2014).

We explore and prioritize these challenges, while considering the theory testing process and how the CS strategy can be adopted in the process. **The relationship between crowdsourcing and the whole theory testing process has not yet been systematically explored and established**, this is the research problem which we seek to explore and tackle in this research.

Establishing this relationship is important from both theoretical and practical perspectives. From a theoretical perspective, finding innovative ways to conduct research is paramount to rigorously test existing theories and advance new theories. The existing crowdsourcing platforms also offer relatively easy-to-implement opportunities to manage the theory testing lifecycle, which in turn may foster the adoption of new methods, as well as the combination of existing methods, which may be explored in different ways. All in all, this suggests the possibility to combine more robust with more flexible approaches to research.

1.3 RESEARCH METHODOLOGY

In this section, the methodology adopted to address these challenges is discussed. We adopt Design Science Research (DSR) as the research paradigm adequate for this study and discuss the research design in detail.

1.3.1 Design Science Research

The Information Systems (IS) discipline is based upon two main paradigms, which are behavioural science and design science (Hevner et al., 2004). Behavioural science is rooted in natural sciences and it seeks to develop and justify theories that explain or predict organizational and human phenomena as relates to IS. Design science is a problem-solving paradigm that seeks to create and evaluate innovative IS artefacts that solve identified organizational problems (Gregor & Hevner, 2013). Design science involves the designing of artefacts, making research contributions, evaluating the designed artefacts, and communicating the results to appropriate audiences (Gregor & Hevner, 2013). By artefacts we mean anything that can be transformed into a material existence as an artificially made object (e.g., model, instantiation) or process (e.g., method, software) (Goldkuhl, 2002; Gregor & Hevner, 2013). Artefacts may also include social innovations (Aken, 2004), or new properties of technical, social, or informational resources (Järvinen, 2007). In essence, any designed object with an embedded solution to a research problem could be seen as design science (Peffers et al., 2007).

In this study, we adopt DSR as our research paradigm because:

1. This study aims to explore the feasibility of adopting crowdsourcing for theory testing. This might not be seen as a problem. As some researchers might say, we have been testing our theories. But it can be viewed as an opportunity to explore an emerging field, using its advantages as a means of creating more ways to theory testing. Design science is adequate for such exploratory purposes (Pries-Heje & Baskerville, 2008).
2. The problem can be considered as a wicked problem. Wicked problems are problems that can only be formulated in terms of a solution and whose solution cannot be denoted as true or false, but only good or bad (Hevner et al., 2004;

Kuechler & Vaishnavi, 2012). They are problems that might not be considered as specific problems and whose solutions are suggested or simplified, and its usefulness is demonstrated or shown through the artefact(s) developed (Goel & Pirolli, 1992; Rittel & Webber, 1973). Therefore, a design science research approach is appropriate in addressing this particular problem.

3. By addressing the objectives of this study, we will be developing artefacts such as a conceptual framework of theory testing, a pattern model of theory testing and a tool which will act as a decision support tool for researchers to determine if and how their theory can be tested using crowdsourcing. These are IT artefacts that can be used to address the problem (Gregor & Hevner, 2013; Kuechler & Vaishnavi, 2012), and therefore they can be developed through DSR.
4. By developing innovative artefacts, we contribute to design theory addressing the operationalization of a theory. These principles of implementation contribute new knowledge to the IS discipline (Gregor, 2006).

1.4 RESEARCH FRAMEWORK

There are two main design activities conducted by DSR: build and evaluate. Build is the process of constructing artefacts with a specific purpose, while to evaluate is to assess the utility of these artefacts (March & Smith, 1995). Based on these premises, we consider a conceptual framework for design science research proposed by Hevner et al. (2004), as seen in figure 1.1. This framework will guide this research.

Design science research integrates three cycles that demonstrate the connection between research activities and outputs: the relevance cycle, the design cycle and the rigor cycle.

The relevance cycle usually triggers design science research and it is also the ultimate instrument to assess its success. The relevance cycle defines the application domain of the research. Our primary goal is establishing the links to research practice. The application domain includes practitioners which our case refers to IS researchers seeking to test theory. We establish close links to research practice by codifying how

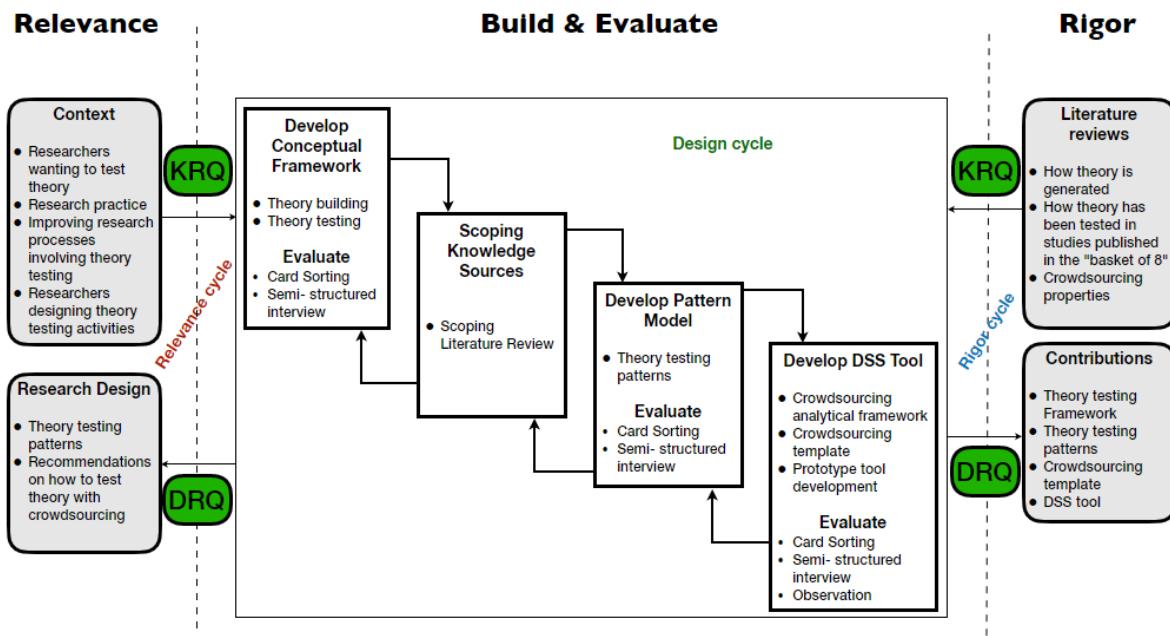


Figure 1.1

Research Framework (Adapted from Hevner et al. (2004))

a relevant set of IS research has approached theory testing in the form of a pattern model. This synopsis can then be used to inform future research endeavours. We also

elaborate a practical set of templates that help making decisions about which theory testing activities can be crowdsourced or not. And finally, we show how the templates can be applied in practice, through the development and evaluation of the decision support tool. Design science research must also be driven by utility, which is the acceptance criteria that validate the research results as designed artefacts must fulfil their intended purpose by improving the environment (Gregor & Hevner, 2013; Hevner & Chatterjee, 2010). Relevance can be achieved by selecting an appropriate evaluation strategy that demonstrates artefact performance in real world contexts (Cleven et al., 2009; Peffers et al., 2012; Venable et al., 2012).

The rigor cycle positions design science research in the scientific foundations that form the knowledge base. These foundations include existing scientific theories and engineering methods that are used by research, and prior knowledge bordered by experience and expertise (Gregor & Hevner, 2013). The examination of this prior knowledge enables the researcher to avoid the pitfalls of producing routine design (existing knowledge applied in familiar problem areas in a routine way, adding no contribution to the knowledge base) instead of design science, and to build on existing

research to ensure innovation. Additions to the knowledge base in the form of research contributions are communicated to the academic community (Gregor & Hevner, 2013; Hevner & Chatterjee, 2010).

Considering rigour, we extract the main concepts of theory testing from general literature on theory development (refer to KRQs), thus providing an adequate frame from which to address the problem (Sonnenberg & vom Brocke, 2011). As there are multiple ways in which a design problem can be framed, an attempt to solve a problem should start with a particular viewpoint (Holmström et al., 2009). In our case, we regard theory testing as the combination of preparation and action stages, where action encompasses data collection and data analysis. We then review the knowledge sources to extract and articulate knowledge on how theory has been tested in the IS field and to also determine the extent to which a body of empirical knowledge supports or reveals any interpretable trends or patterns with respect to pre-existing findings (Paré et al., 2015) (refer to DRQs). This is done through a descriptive literature review by collecting, codifying and analysing data that reflects the frequency of relevant topics found in the literature (Paré et al., 2015). The outcomes of the review establish a clear link between the problem and the state-of-the-art, which is considered important to assert rigour (Gregor & Hevner, 2013; Peffers et al., 2007). A systematic approach to data collection was adopted to increase rigor and transparency (Kitchenham et al., 2009; Paré et al., 2016). We also followed the data collection guidelines suggested by Kitchenham et al. (2009), which delineate a transparent process for synthesising the results.

The design cycle is at the centre of design science research. In this cycle, artefacts are built and evaluated in successive iterations that refine the design and improve its utility. At each stage of iteration, input is drawn from the rigor cycle in the form of applicable knowledge, and from the relevance cycle in the form of business needs. Conversely, outputs are fed to the rigor cycle as additions to the knowledge base, and to the relevance cycle as contextualized applications in the research environment (Hevner & Chatterjee, 2010).

We note that in design science research, the evaluation/justification of artefacts provides two complementary types of contributions: 1) contribute to the iterative

development of artefacts; and 2) contribute to assess the relevance of the ensemble artefact. Since DSR is still a relatively young paradigm, there has been some ongoing discussion regarding this dual perspective on evaluation/justification (Sonnenberg & vom Brocke, 2011). In particular, should the evaluation be focussed on the last stage of the research, where the primary concern is relevance, or done multiple times during the project, thus contributing to justification? In fact, DSR does not have to be limited to a single, summative evaluation of the ensemble artefact. Evaluations and justifications can be done throughout the project to increase systematicity and transparency. In this research, we adopt the iterative view, which considers how artefacts support each other in the development of the ensemble artefact, and then how the ensemble contributes to generate justificative knowledge (Sonnenberg & vom Brocke, 2011).

1.4.1 Research Questions

Having identified the methodology and framework to be used in this study and explained why, we proceed by highlighting the research questions guiding the study. These questions represent what this study addresses, investigates and ultimately answers (Thuan et al., 2019).

The overarching Research Question (RQ) for this study can be stated as this “To what extent can crowdsourcing be used in the whole theory testing process”? To answer this main question, we further breakdown the RQ to two main aspects – Knowledge RQ and Design RQ, both having sub-question.

Knowledge Research Questions (KRQ)

RQ1. What is the general knowledge about theory testing in IS discipline?

- 1.1 What types of theory can be tested?
- 1.2 What are the conceptual elements of theory testing?
- 1.3 What are the challenges researchers face during theory testing?
- 1.4 How has theory testing been conducted by researchers?

Design Research Questions (DRQ)

- RQ2. How to assess what theory testing activities can be crowdsourced or not?
- 2.1 How to operationalize a theory so that it may be tested using crowdsourcing?
 - 2.2 Can a decision support tool for the incorporation of CS in testing IS theory be developed?

1.4.2 Research Objectives

To address the research questions, we consider an overarching Research Objective (RO) which is to develop an innovative decision tool aimed at assisting researchers testing IS theory to incorporate CS. This RO is further dissected into smaller bits of ROs for clarity.

RO1) Identify, condense and make sense of the conceptual elements of theory testing through literature review.

RO2) Undertake a systematic review of how theories in the IS domain have been tested, whiles analysing what types of theories can be operationalized

RO3) Profile the different ways in which theories have been tested.

RO4) Identify and make sense of the conceptual elements of crowdsourcing and crowdsourcing criteria necessary to make the decision to crowdsource through literature review.

RO5) Help IS researchers making decisions about how to integrate CS in the process of testing IS theory by providing support.

1.4.3 Research Contribution

The DSR knowledge base can be broadly divided into two distinct types: descriptive knowledge, the “what” knowledge about natural phenomena and the laws and regularities of phenomena; and prescriptive knowledge, the “how” knowledge of human-built artefacts (Gregor & Hevner, 2013).

This research will be drawn from both the descriptive knowledge, where existing justificatory theories and knowledge relating to the research objectives is drawn and the prescriptive knowledge base and contributing to design theory (See figure 1.2). A design theory can be about both the principles underlying the form of the design, also about the act of implementing the design in the real world and viewpoints (Gregor & Jones, 2007). They both provide a baseline of knowledge on which to build and evaluate the artefacts.

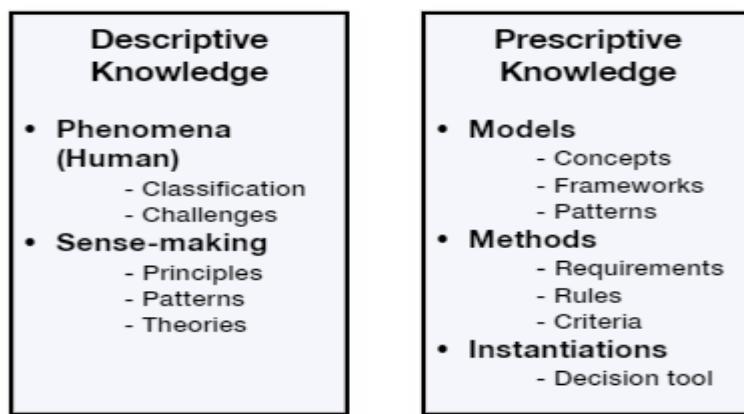


Figure 1.2. Design Science Research Knowledge Base (adapted from (Gregor & Hevner, 2013))

The development of artifacts adds or contributes to the knowledge base, and there are different levels of artifact development and different contribution types. The Table 1.1 shows the different levels and contribution types. Contributions for this study ranges from development of design theory, which involves the review and consolidation of theory testing in a design science perspective from existing literature, then moving to the next level of abstraction is the development of a pattern model, which acts upon the identified building blocks and finally, the more specific abstraction is the development of a tool, which operationalizes the elements of the pattern model into decision support tool for researchers.

Table 1.1 Design Science Contribution Types derived from (Gregor & Hevner, 2013)

Knowledge Base	Contribution Types	Specific Contributions Addressed by Thesis
More abstract  More specific	Conceptualisations	Conceptual Framework
	Models and Operational rules	Pattern Model
	Instantiations	Decision Tool

1.5 RESEARCH PROCESS IN DETAIL

In this section, a more detailed description is given on how the research will be conducted. According to the principles of design science, the developed approach will consist of several IS artefacts (Gregor & Hevner, 2013).

More specifically, the research process follows four main stages: first stage includes *conceptual framework*, second stage includes *scoping knowledge source*, third stage includes a *pattern model*, and the fourth stage includes *the decision tool*, which comprises of crowdsourcing analytical tool, crowdsourcing process criteria template and the decision tool (see figure 1.3).

1.5.1 Build Activities

The research process defines four build activities, which are: build conceptual model, scoping knowledge source, build pattern model and build decision tool. All these activities are based on rigorous research techniques (Hevner & Chatterjee, 2010). We now give a brief specification of the nature of each activity, and the role each activity plays in addressing the research objectives, details of these techniques are discussed in relevant chapters. Considering the design science research tenets, the next section starts with the build activities of each stage, followed by their evaluations.

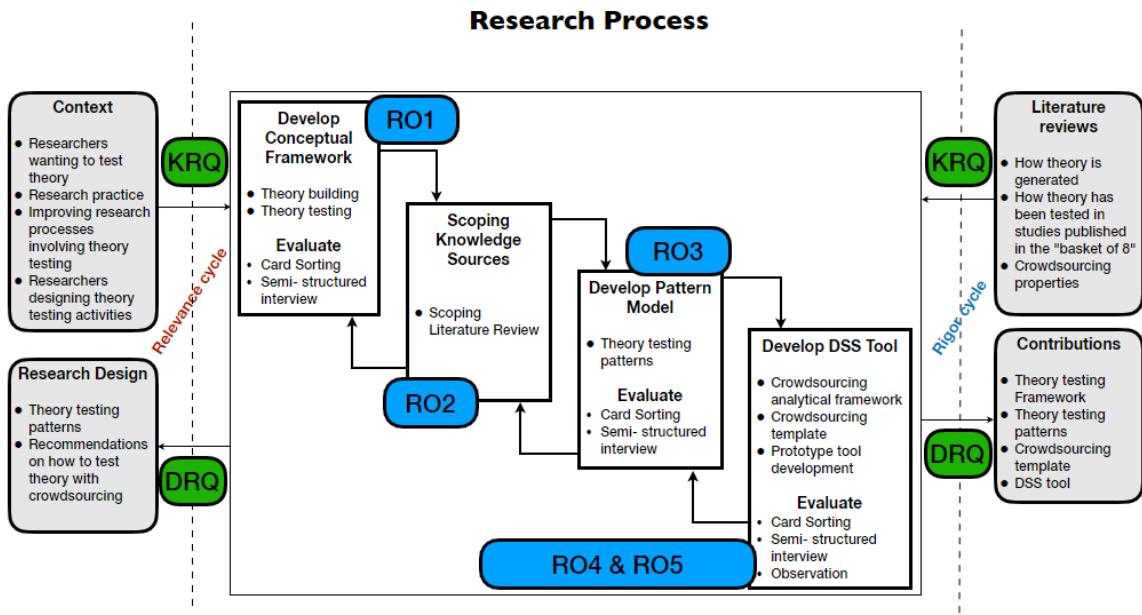


Figure 1.3.

The Research Process

Build Conceptual Model

The conceptual modelling grasps the main concepts and relationships in the domain of interest (Webster & Watson, 2002). This activity helps to identify and frame the conceptual elements of theory testing, thereby addressing RO1. In this activity, the main building blocks of theory testing are structured in a meaningful way, clearly identifying the core elements of theory testing independently from theory building.

The main concepts, processes, relationships and building blocks are scoped from knowledge bases, which in this case are existing literatures. After identifying such concepts and relationships, we then graphically represent them in a conceptual model (Jonker & Pennink, 2010).

Scoping Knowledge Sources

This activity extracts and articulates the existing knowledge related to the DSS that we seek to develop (Thuan et al., 2016), thereby addressing RO2 and RO4. The relationship between the state-of-art and the DSS is established. This is an important aspect for design science because it contributes to rigour (Gregor & Hevner, 2013; Peffers et al., 2007).

The ‘scoping’ classifier refers to a comprehensive sample strategy, in which a vast range of knowledge sources existing in the domain is covered. The descriptive literature review method was adopted to identify the different ways theory has been tested in the IS field. This method was selected because it helps to determine the extent to which a body of empirical knowledge supports or reveals any interpretable trends or patterns with respect to pre-existing findings (Paré et al., 2015). This was done by collecting, codifying and analysing data that reflects the frequency of relevant topics found in the literature (Paré et al., 2015). The systematic approach was adopted to increase rigour and transparency (Kitchenham et al., 2009; Paré et al., 2016). The guidelines suggested by Kitchenham et al. (2009) was followed: 1) select articles; 2) filter articles; 3) data extraction; and 4) data synthesis. These steps are explained in detail in paper 1 and 3.

Build Pattern Model

The third activity is the pattern model construction, which serves to consolidate knowledge acquired from scoping of the knowledge sources and addressing RO3. Our interest is not focused on the conceptual parts of theory testing, but more on the activities and patterns involved in the theory testing process, as presented by researchers to the community through published articles.

Series of theory testing activities and patterns were identified, with their associated relationships, which was then graphically consolidated in a pattern model. This model was constructed in a generic format, such that it can serve different research methods.

Build Decision Tool

This activity has different sub-activities attached to it. This includes a crowdsourcing analytic framework, a crowdsourcing process criteria template and a decision support tool, addressing RO4 and RO5. All these activities are aimed to assist in decision making.

Build Crowdsourcing Analytical Framework

This is the initial phase in the development of the decision tool. It involves a graphical scheme that provides the different crowdsourcing features. These features were

identified from existing literature bearing in mind what it aims to achieve, which is how such features relates to theory testing and its activities. The graphical scheme serves as a feasibility check which produces three results: 1) we cannot crowdsource; 2) we can crowdsource, but it may be difficult to accomplish; and 3) it is easy to crowdsource.

Build Crowdsourcing Process Criteria Template

This activity came as a modification of the above framework. We translated the framework, which had some modifications into templates. These templates serve as a preliminary stage in design and development. It provides a checklist of relevant information needed for enacting and automating the process (Lee & Bui, 2000). It aims to systematically analyse how to adopt crowdsourcing for the various theory testing patterns. A template may suggest three types of results: cannot crowdsource the pattern; can partially crowdsource the pattern; and can fully crowdsource the pattern. In the two latter cases, the templates also provide considerations about the difficulties involved in crowdsourcing the pattern.

Build Decision Support Tool

This is the last activity and serves to consolidate all the knowledge acquired into a simple tool that assists researchers in decision making. The DS tool is developed as an improved version of the templates. This tool is implemented using Visual Basic for Applications (VBA) Excel, providing wider access to the knowledge base. VBA is the programming language used on an excel interface. It enables building of user-defined functions (UDFs), automating processes and assessing windows API through dynamic-link libraries (DLLs).

In summary, the build activities of the research process are justified: the conceptual mode, scoping of knowledge sources, pattern model and decision tool. The expected outcomes from this includes scoped knowledge from different sources, and their built artefacts, which in turn addresses the research objectives. We give a summarized version of the various build activities and how it is integrated in various papers in table 1.2. Evaluations needs to be done for each of these stages, this is discussed in the next section.

Table 1.2 Summary of Research Activities aligning to Papers

Chapter	Title of paper	Paper source	Conceptual Model development	Scoping knowledge sources	Pattern model development	Decision tool development	ROs	Publication Strategy
Chapter 2	On the adoption of crowdsourcing for theory testing	ECIS 2018	Theory testing contextualization and conceptualization	Theory testing literature, MISQ journal papers (10-year period) and crowdsourcing literature	Theory testing process and activities		RO1, RO2 and RO3	Merge and extract main ideas to form a journal paper, to be submitted in one of the top IS journals (within the basket of 8)
Chapter 3	Towards the development of a DSS supporting the integration of crowdsourcing in theory testing: Conceptual framework and model	ACIS 2017	Theory testing contextualization and conceptualization	Development method	Theory testing process and activities		RO1 and RO4	
Chapter 4	Patterns of testing theory with human subjects: A design science perspective	AMCIS 2019	Theory testing contextualization and conceptualization	Theory testing literature and basket of 8 IS journals	Theory testing process, activities and patterns		RO1, RO2 and RO3	
Chapter 5	Towards the development of a DSS supporting the integration of crowdsourcing in theory testing: Analytical framework design	AMCIS 2019		Crowdsourcing literature		Crowdsourcing analytical framework development	RO4 and RO5	Merge and extract main idea to be publish in a top-ranked IS journal

Chapter 6	Analysing the adoption of crowdsourcing in the whole theory testing lifecycle.	*	Theory testing contextualization and conceptualization	Theory testing literature, basket of 8 IS journals and Crowdsourcing literature	Theory testing process, activities and patterns	Crowdsourcing process criteria templates' development	RO1, RO2, RO3 and RO4	
Chapter 7	Design and development of a DSS supporting the integration of crowdsourcing in theory testing: A design science perspective.	PACIS 2019		Crowdsourcing literature and decision support literature		Decision tool	RO5	

* Yet to be submitted.

1.5.2 Evaluation

When choosing an evaluation method, it is important as a design researcher to balance the interests of practitioners and researchers (Sonnenberg & vom Brocke, 2011), while the practitioners are interested in the applicability and usefulness of the artefact, the researchers are interested in the validity of the artefact. Bearing this in mind led to the selection of the evaluation methods stated below. Evaluation in DSR should not be limited to a single activity conducted at the conclusion of a design construct-evaluate cycle but can be conducted at two points in time relative to the artefact construction (Sein et al., 2011): (1) ex ante where artefacts are evaluated prior to their implementation or actual construction, and (2) ex post where artefacts are evaluated after they have been designed and constructed (Pries-Heje et al., 2008). Therefore, evaluations will be done for each stage of the research process. Literature review was used for the conceptual framework and scoping knowledge source because its focus is more on validity than utility, while the model and DSS artefact, while we are still interested in their validity, its usefulness is much more important, which led to the use of comparative studies, card sorting, observation and interview as an evaluation method. We therefore consider multiple evaluations, which are summarised in Table 1.3.

Table 1.3 Evaluation Methods

Evaluation Method	Conceptual framework	Model	DSS
Comparative studies		✓	✓
Card sorting	✓	✓	✓
Interview	✓	✓	✓
Observation			✓

1.5.3 Research Checklist

To help us with the use of the design science paradigm and the development of artefacts, we consider the checklist of eight questions – the “questions map” proposed by (Hevner & Chatterjee, 2010).

Table 1.4 Design science research question map/checklist adapted from (Hevner & Chatterjee, 2010; Hevner et al., 2004)

S/N	Checklist/question map	Answers
1	What is the research question – the design requirements?	Feasibility of adopting crowdsourcing for theory testing
2	What is (are) the artefacts and how will they be represented?	a) Conceptual Model, b) Pattern Model, c) Decision Tool
3	What design processes will be used to build the artefact?	Two-phase process – Build and Evaluate
4	How are the artefacts and the design processes grounded by the knowledge base and what are the theories (if any) that supports the artefact design and design process?	Evidence- based approach – Review and consolidation from existing literature in a systematic way.
5	What evaluations are performed during the internal design cycles and what design improvements are identified during each design cycle?	a) Literature review b) Comparative studies c) Card Sorting d) Interview e) Observation
6	How is the artefact introduced into the application environment and how is it field tested? What metrics are used to demonstrate artefact utility and improvement over previous artefacts?	Real users, real systems and real tasks in a real environment.
7	What new knowledge (e.g. peer-reviewed literature, new theory, new method, meta-artefacts) is added to the knowledge base?	Models and design solutions
8	Has the research question/objectives been satisfactorily addressed?	Research results, contributions and impact on practice

The questions/checklist stated in table 1.4 is answered according to the requirements of this research.

1.6 OUTLINE OF THE MANUSCRIPT

This chapter has introduced the major concepts of the study, such as theory testing, crowdsourcing, and has highlighted the research goals and objectives. It has also presented the chosen methodology, research framework and process. The remaining chapters of this thesis is as follows (figure 1.4).

Chapter 2 – Paper 1

This chapter contains the first paper, which has been published. It can be referred to as our pilot study. It contains preliminary research and study on theory testing, the model and CS. It serves to put out our idea and get valuable feedback. Some aspects of RO1 and RO2 were explored, thereby giving preliminary answers to RQ1.

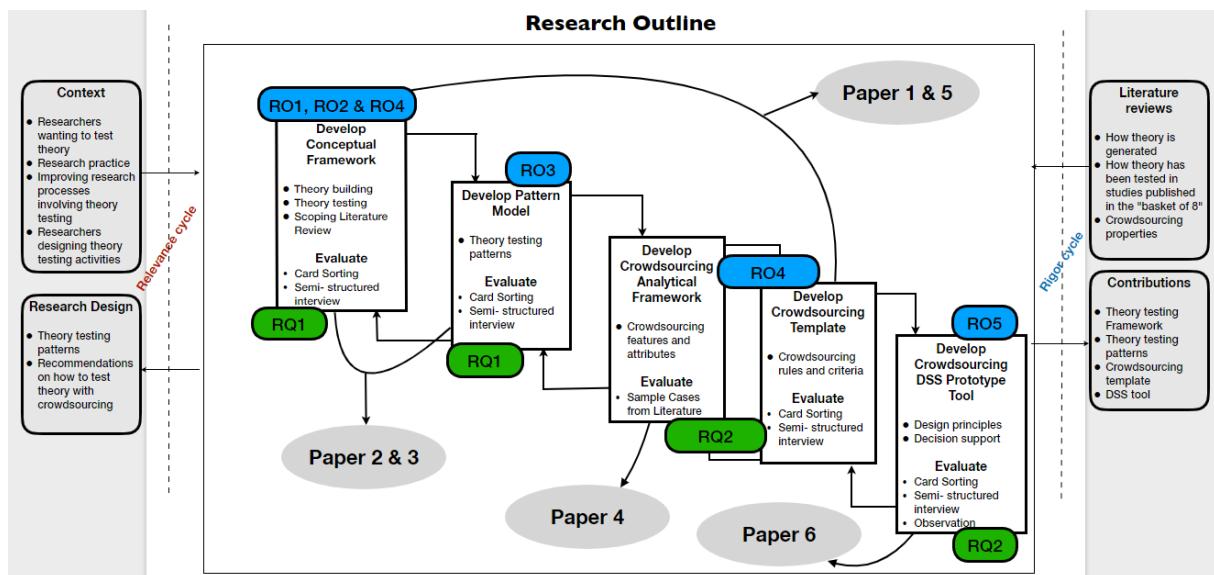


Figure 1.4.

The Research Outline

Chapter 3 – Paper 2

This chapter contains the second paper which has also been published. The main focus of this paper was to highlight the development method selected for all artefacts that we propose to develop and how that is situated in the literature. It also contained the problem frame assessment (PFA) and its results. We also obtained some user requirements for CS and DSS tool.

Chapter 4 – Paper 3

This chapter contains the third paper, which has been published, but has added information, as the submitted version was constrained due to page limit given. This chapter was focused on further research done on theory testing, the different cycles identified, boundary activities between theory building and testing, and the development of the conceptual model. It also contains further work done of the development of the pattern model, and the evaluations done with their results. This chapter further explored RO1, 2 and 3 in a more detailed manner, thereby answering RQ1 in a more detailed manner.

Chapter 5 – Paper 4

This chapter contains the fourth paper, which has been published. This chapter begins the exploratory work on how to align theory testing and crowdsourcing, and the designing of the first decision tool in form of a graphical scheme. This chapter explored RO4 and 5, thereby answering some aspects of RQ2.

Chapter 6 – Paper 5

This chapter contains the fifth paper, which is yet to be submitted. This chapter majorly contains an improvement to the artefact design in the previous chapter, the CS attributes were modified based on feedback gotten and good constructive criticism done on the previous paper. The CS templates were developed in this paper, which is also one of the decision tools. This chapter explored RO4 and 5, thereby answering RQ2.

Chapter 7 – Paper 6

This chapter contains the sixth paper, which has also been published. This chapter wraps up the thesis work by developing an instantiation. Based on the developed templates and further modifications, a software tool was developed to assist in decision-making. It is a user-friendly tool, and easy to use. This was also evaluated. This chapter explored RO4 and 5, thereby answering RQ2.

Chapter 8 – Conclusion

This is the concluding section, and it summarises the research results and main conclusions from the different papers. Also contains future research work.

1.7 REFERENCES

- Adzeh, K. J. (2014). The challenge of conducting a quantitative business research: Analysis of issues with survey design, sampling, validity, and reliability. *Research Gate*.
- Aken, J. E. v. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of management studies*, 41(2), 219-246.
- Ashton, S. (2014). Researcher or nurse? Difficulties of undertaking semi-structured interviews on sensitive topics. *Nurse Researcher*, 22(1).
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior research methods*, 43(3), 800.
- Brabham, D.C. (2008). Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence: The International Journal of Research into New Media Technologies*, 14(1), 75-90. doi:10.1177/1354856507084420
- Callison-Burch, C. (2009). *Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk*. Paper presented at the Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1.
- Campbell, J. (1990). The role of theory in industrial and organizational psychology. In Dunnette MD, Hough LM (Eds.), *Handbook of industrial and organizational psychology*, Vol. 1(2nd ed.), pp. 39-73.
- Cleven Anne, Gubler, P., & Hüner, K. M. (2009). *Design alternatives for the evaluation of design science research artifacts*. Paper presented at the Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology.
- Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of management journal*, 50(6), 1281-1303.
- Conley, C., & Tost-Kharas, J. (2014). Crowdsourcing content analysis for managerial research. *Management Decision*, 52(4), 675-688.
- Ford, R. C., Richard, B., & Ciuchta, M. P. (2015). Crowdsourcing: A new way of employing non-employees? *Business Horizons*, 58(4), 377-388.
- Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive science*, 16(3), 395-429.
- Goldkuhl, G. (2002). *Anchoring scientific abstractions—ontological and linguistic determination following socio-instrumental pragmatism*. Paper presented at the Proceedings of European Conference on Research Methods in Business, Reading.
- Green, J., & Thorogood, N. (2018). *Qualitative methods for health research*: Sage.
- Gregor, S., & Hevner, A.R. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 337-355.
- Gregor, S. (2006). The nature of theory in information systems. *MIS quarterly*, 611-642.
- Gregor, S., & Jones, D. (2007). The anatomy of a design theory. *Journal of the Association for Information Systems*, 8(5), 312.
- Hevner, A., & Chatterjee, S. (2010). Design science research in information systems. In *Design research in information systems* (pp. 9-22): Springer, Boston, MA.

- Hevner, A. & Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol. 22): Springer Science & Business Media.
- Hevner, Von Alan, March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1), 75-105.
- Holden, R. J., Scott, A. M. M., Hoonakker, P. L., Hundt, A. S., & Carayon, P. J. Q. o. L. R. (2015). Data collection challenges in community settings: Insights from two field studies of patients with chronic disease. 24(5), 1043-1055.
- Holmström, J., Ketokivi, M., & Hameri, A. (2009). Bridging practice and theory: A design science approach. *Decision Sciences*, 40(1), 65-87.
- Howe, J. (2006). The rise of crowdsourcing. *Wired magazine*, 14(6), 1-4.
- Israel, B. A., Krieger, J., Vlahov, D., Ciske, S., Foley, M., Fortin, P., Palermo, A. and Tang, G. (2006). Challenges and facilitating factors in sustaining community-based participatory research partnerships: lessons learned from the Detroit, New York City and Seattle Urban Research Centers. *Journal of Urban Health*, 83(6), 1022-1040.
- Järvinen, P. (2007). Action research is similar to design science. *Quality & Quantity*, 41(1), 37-54.
- Jonker, J., & Pennink, B. (2010). *The essence of research methodology: A concise guide for master and PhD students in management science*: Springer Science & Business Media.
- Kitchenham, B., Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1), 7-15.
- Kuechler, W., & Vaishnavi, V. (2012). A framework for theory development in design science research: multiple perspectives. *Journal of the Association for Information systems*, 13(6), 395.
- Lee, A.S., & Hubona, G. S. (2009). A scientific basis for rigor in information systems research. *MIS quarterly*, 237-262.
- Lee, J., & Bui, T. (2000). *A template-based methodology for disaster management information systems*. Paper presented at the System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on.
- Lynham, S. A. (2002). The general method of theory-building research in applied disciplines. *Advances in developing human resources*, 4(3), 221-241.
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision support systems*, 15(4), 251-266.
- Morris, R. (1994). Computerized content analysis in management research: A demonstration of advantages & limitations. *Journal of Management*, 20(4), 903-931.
- Mueller, B., & Urbach, N. (2013). *The Why, What and How of Theories in Information Systems Research*. Paper presented at the 34th International Conference on Information Systems (ICIS).
- Munro, M., Sogoni, M., Mashiyi, M., Shigadhlala, O., & Tsotsi, N. (2009). Fieldwork experiences and challenges in data collection: The Eastern Cape Province of South Africa. *A paper presented at the 57th session of the International Statistical Institute*.
- Nel, E., Rich, E., Morojele, N., Harker Burnhams, N., Petersen Williams, P., & Parry, C. (2017). Data collection challenges experienced while conducting the international alcohol control study (IAC) in Tshwane, South Africa. *Drugs: Education, Prevention and Policy*, 24(5), 376-383.

- Nunamaker Jr, J. F., Chen, M., & Purdin, T. D. (1990). Systems development in information systems research. *Journal of management information systems*, 7(3), 89-106.
- Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. *European Journal of Information Systems*, 25(6), 1-16.
- Paré, G., Trudel, M., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183-199.
- Parvanta, C., Roth, Y., & Keller, H. (2013). Crowdsourcing 101 a few basics to make you the leader of the pack. *Health promotion practice*, 14(2), 163-167.
- Peffers, K., Rothenberger, M., Tuunanen, T., & Vaezi, R. (2012). *Design science research evaluation*. Paper presented at the International Conference on Design Science Research in Information Systems.
- Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.
- Pierce, C. S., Scherra, E. J. O. J. o. R. N., & Care, H. (2012). The challenges of data collection in rural dwelling samples. 4(2), 25-30.
- Pries-Heje, J., & Baskerville, R. (2008). The design theory nexus. *MIS quarterly*, 731-755.
- Pries-Heje, J., Baskerville, R., & Venable, J. R. (2008). *Strategies for Design Science Research Evaluation*. Paper presented at the ECIS.
- Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, 4(2), 155-169.
- Sarkar, M. (2014). Challenges in Conducting Educational Research: The Case of a Developing Country. *Contemporary approaches to Research in Mathematics, Science, Health and Environmental Education*.
- Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics & Management*(1), 93-107.
- Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS quarterly*, 37-56.
- Smith, D., Manesh, M. M. G., & Alshaikh, A. (2013). How can entrepreneurs motivate crowdsourcing participants? *Technology Innovation Management Review*, 3(2).
- Sonnenberg, C., & vom Brocke, J. (2011). *Evaluation patterns for design science research artefacts*. Paper presented at the European Design Science Symposium.
- Thuan, N. H., Antunes, P., & Johnstone, D. (2016). A Design Science Method for Emerging Decision Support Environments. *arXiv preprint arXiv:1605.04725*.
- Thuan, N. H., Drechsler, A., & Antunes, P. (2019). Construction of Design Science Research Questions. *Communications of the Association for Information Systems*, 44, pp-pp.
- Venable John, Pries-Heje, J., & Baskerville, R. (2012). *A comprehensive framework for evaluation in design science research*. Paper presented at the International Conference on Design Science Research in Information Systems.
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii-xxiii.
- Whitla, P. (2009). Crowdsourcing and its application in marketing activities. *Contemporary Management Research*, 5(1).

- Yuen, M.C., King, I., & Leung, K.S. (2011). *A survey of crowdsourcing systems*. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing* (pp. 766-773). IEEE.
- Zhao, Y., & Zhu, Q. (2012). Exploring the motivation of participants in crowdsourcing contest. Paper presented at the 33rd International Conference on Information Systems, Orlando.

CHAPTER 2

ON THE ADOPTION OF CROWDSOURCING FOR THEORY TESTING

ABSTRACT

This paper examines the possibilities of using the crowdsourcing strategy for theory testing. We first analyse the relationships between theory building and theory testing activities. Then, based on a systematic review of 248 papers published in MISQ, we characterise the intents and pattern systems of activities that have been used for theory testing. Finally, we ascertain which activities can be crowdsourced or not and pinpoint a set of pathways supporting partial and total crowdsourcing. The obtained results show that a large number of activities related to data gathering can be crowdsourced, and that a number of intents have viable pathways supporting partial crowdsourcing.

Keywords: Theory, Theory testing, Crowdsourcing.

2.1 INTRODUCTION

Theory concerns “abstract entities that aim to describe, explain, and enhance understanding of the world and, in some cases, to provide predictions of what will happen in the future and to give a basis for intervention and action” (Gregor, 2006). Theory also provides a linguistic tool for organizing scientific knowledge and scientific research (Colquitt & Zapata-Phelan, 2007). Finally, theory can be used as an educational device to raise consciousness about a set of concepts. Though only validated theory can be considered useful and helpful.

There are two general forms of validating theory: formative and summative (Lee & Hubona, 2009). The former is part of the theory building process while the latter seeks closure with a statement whether there is support or not to a theory. Theory building is an on-going process of observing, confirming, applying, and adapting theory (Lynham, 2002). It also involves the development of new ideas, concepts, conceptual frameworks, methods, and models (Nunamaker Jr et al., 1990). Theory testing is a

natural complement to theory building; just as Bacharach (1989) said, “if it is not testable, no matter how profound or aesthetically pleasing it may be, it is not theory”.

Theory testing involves collecting data either confirming or disconfirming the set of propositions articulated by a theory (Lokke & Sorensen, 2014). We are particularly interested in theory testing that involves human participants producing a wide range of data such as opinions, thoughts, expertise, attitudes, communications, observations, actions, and task logs. Theory testing involving human participation is ethically, conceptually and practically challenging. Especially in social sciences, one has to respect best practices regarding the participants' involvement, the nature of data being collected, diversity of variables influencing human behaviour, and adequacy of data collection processes, methods and tools (Bhattacherjee, 2012). A more practical but nevertheless important challenge is recruiting and engaging study participants (Witschey et al., 2013).

Although many robust methods have already been developed to test theories with human participation, which may include interviews, surveys, and many others, we are still curious to know to what extent the crowdsourcing (CS) strategy, can be used to support the theory testing process, especially in the Information Systems (IS) field, and to understand the advantages and constraints this method brings to theory testing. CS is a strategy in which a single task maybe fragmented into multiple tasks delivered to “a large group of people in the form of an open call” (Howe, 2006). The relationship between the requestor and performers is usually loose, with little or no restrictions on who may perform the task (Schenk & Guittard, 2011; Whitla, 2009; Yuen et al., 2011; Zhao & Zhu, 2012). Besides, the whole process is supported by information technology, which helps reaching the crowd and managing the task execution. CS has been recognised to take advantage of parallelisation and to harness the collective intelligence of the crowd (Parvanta et al., 2013).

From the outset, the potential advantages of adopting CS for theory testing could be significant. CS may help controlling the timing and the scale of data collection in an elastic way, from very short periods but huge scale to long periods and smaller scale. This elasticity also extends to spatial coverage, regarding countries, organisations, contexts, etc. (Brabham, 2008; Ford et al., 2015; Smith et al., 2013). Further, CS may

also reduce the overhead associated to the testing process. This includes removing administrative overhead related to procuring participants, supervision, scheduling, training, managing experimental treatments, selecting samples, distributing research instruments, and collecting results.

Even though the adoption of CS to theory testing seems a compelling endeavour, we must address an important challenge: **how to align theory testing with crowdsourcing**. Research communities have developed theory-testing patterns, which may have to be adjusted to the CS strategy and re-checked again for robustness. And in some cases, we may find that CS may not be feasible.

In this paper, we investigate this alignment problem in the IS field. We analyse various ways in which theories have been tested. Then, we seek to identify **patterns** that could help us discover how theory testing can be aligned with the CS method. To achieve this goal, we undertake a descriptive literature review in the IS field. Since CS may have a particular impact on how theory testing is conducted, we focus our review on the operational dimension of theory testing, rather than philosophical, epistemological or methodological dimensions.

The remainder of this paper is organized as follows. Section two discusses related work on theory types, theory building, theory testing, and how they may be related to CS. Section three describes the method and procedure used for the literature review, as well as the obtained results. Section four contains the discussion of the results and Section five contains some concluding remarks and future work.

2.2 RELATED WORK

An academic discipline will not have much leverage if it does not have useful theory to contribute (Niederman & March, 2015). Theory helps scientists to logically describe and explain a phenomenon, process or sequence of events. Gregor, (2006) suggests a classification into five theory types. This study will focus mainly on theories with testable proposition, therefore, we will consider type III and IV theory.

Type III theory aims to predict “what will be”. These theories are able to predict the outcomes, results or consequences using some explanatory factors but without

necessarily explaining the connections among factors or explaining the independent and dependent variables. Some research approaches falling into this category use statistical techniques such as correlation or regression analysis (Gregor, 2006) to predict but not to explain phenomena.

Type IV theory is concerned with “what is”, “how”, “why” and “what will be”. A theory in this category is capable to provide both the prediction and the scaffolding of the underlying causes of a phenomenon using theoretical constructs. To many, this is the ideal theory as it seeks to define the concepts and propositions underlying a phenomenon, while specifying a set of fundamental relations among different variables that help predicting the phenomenon (Gregor, 2002).

2.2.1 Focus on theory with testable propositions

Type III and type IV theories will be our primary targets for identifying theory-testing patterns, since researchers develop testable propositions with the explicit purpose to test them. The other types of theory do not have to pass the testing challenge, since propositions only need to be either identified (types I and II) or applied (type V).

Goode & Hatt (1952) define testable proposition as an imaginative idea, a guess, a statement of a solvable problem or any thinking that can be put to test to determine its validity. A testable proposition can also be an hypothesis: a tentative generalization which has to be tested to determine its validity (Khan, 2011). A testable proposition should have the following characteristics (Khan, 2011): 1) refer to observable and measurable events, which are pre-conditions for the formulation of testable propositions; 2) be conceptually comprehensible and provide a solution to a defined problem; 3) be formulated in a way that can be tested and verified directly; 4) be related to the existing body of knowledge; and 5) have logical unity and comprehensiveness (Khan, 2011).

The test of a theory is concerned with establishing or refuting the validity of the theory's core propositions. A theory is tested by determining to what degree it provides or not a close fit to empirical data (Colquitt & Zapata-Phelan, 2007). These binary outcomes are the basis for codifying knowledge and creating the opportunity for further research.

Testing and re-testing theory are considered key foundations to progress scientific knowledge (Niederman & March, 2015).

Researchers spend time finding and then explaining fundamental relationships in empirical data (Viswesvaran & Ones, 1995). Researchers often also try to explore the mediators and moderators that define the boundary conditions of a theory. De Vaus (2013) and Bitektine (2008) suggested a stage-gate framework describing both theory building and testing. We adapted that framework to focus primarily on theory testing, noting however a set of pre and post conditions necessary to realise theory testing (Figure 2.1):

Pre-conditions:

1. Have a statement with a brief explanation of the theory and what it intends to achieve. Have a set of conceptual propositions. A conceptual proposition is a statement that shows the relationship between two factors. The more conceptual propositions to be tested, the stronger the test of a theory (Stinchcombe, 1987).

Theory testing steps:

1. If necessary, restate the conceptual propositions as testable propositions. This involves translating abstract concepts into concrete, observable and measurable variables. This step is often called operationalization (De Vaus, 2013).
2. Decide what data are relevant or appropriate to test the testable propositions.
3. Collect relevant data.
4. Analyse the data. Data are analysed to see: a) how much support there is for the testable propositions; b) how much support there is for the conceptual propositions; and c) how much support there is for the whole theory.

Post-conditions:

1. Have a statement assessing the whole theory. The theory may be or may not be supported completely.

Most theories in social sciences are tested quantitatively (Bitektine, 2008; Hyde, 2000; Piper, 2006), while others are qualitatively tested, especially using the case

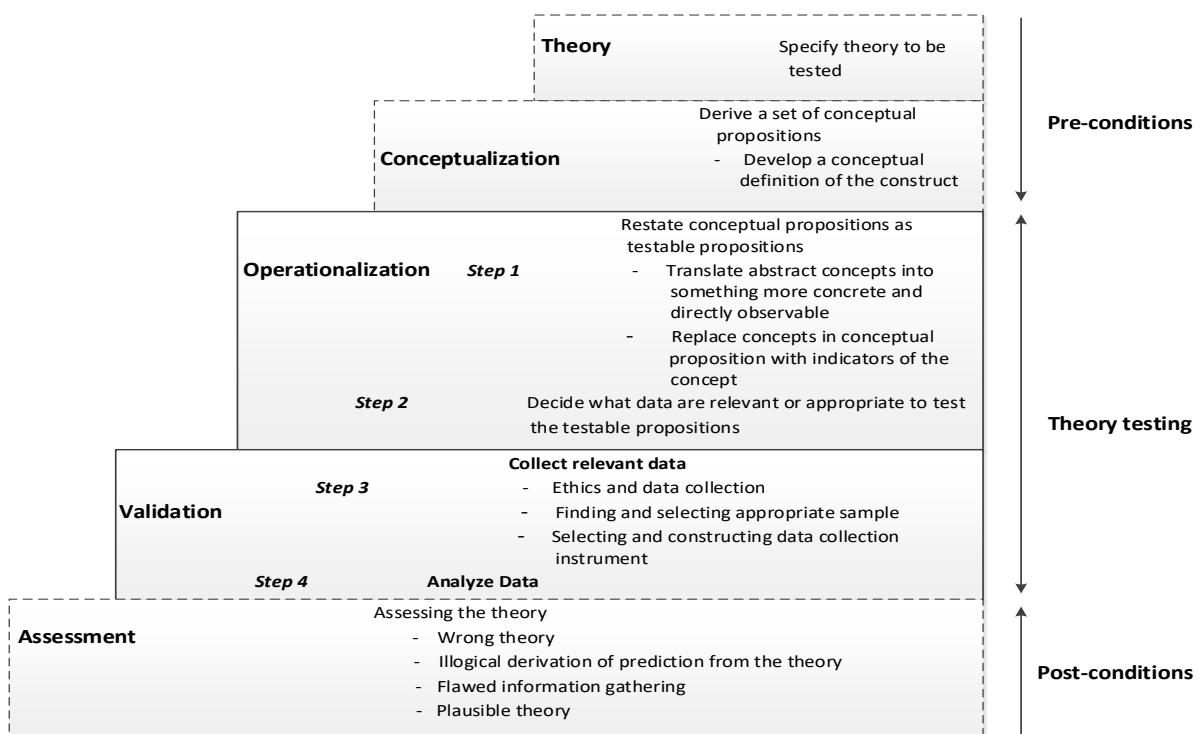


Figure 2.1. Steps to theory testing, adapted from (Bitektine, 2008; De Vaus, 2013; MacKenzie et al., 2011; Niederman & March, 2015c)

study method. Bitektine (2008) identified some important steps to consider for theory testing when adopting the case study method that significantly align to the one mentioned above, with the following amendments:

Pre-conditions:

3. Identify the case(s), unit of analysis and subjects.

Theory testing steps:

4. Analyse the case(s) and identify patterns using the foundations provided by the theory.
5. Analyse internal and/or external validity. Realise that single-case studies tend to generate type-1 errors (accept false hypotheses) while cross-case studies tend to generate type-2 errors (reject true hypotheses) (Gerring, 2006).

2.2.2 Considering crowdsourcing in theory testing

To be able to relate CS to theory testing, it is important to first look at the main concepts pertaining to CS, which we summarise in Table 2.1. The task which includes data collection, could assume multiple forms: questionnaires, social marketing, fact checking, tagging, activity logging, voting, playing, simulation, ideation, commentary, expert opinions, quantified self, etc. This concept of relying on the crowd for data collection is not completely new, since it is also found in citizen's science (Bonney et al., 2009) and living labs (Følstad, 2008). Furthermore, tasks may also involve related

Table 2.1. *Crowdsourcing Concepts*

Concept	Description
Crowdsourcer	The entity that seeks to carry out a specific task by harnessing the crowd. In our context, the crowdsourcer is the researcher (De Vaus, 2013; Parvanta et al., 2013).
Task, Crowd task	The task is what fulfils the crowdsourcer's goals, while the crowd task is what is actually assigned to the crowd. Usually, the crowd task is a decomposition of the task.
Crowd	The people assigned to the task. They tend to be unknown to each other and often unknown to the crowdsourcer.
Incentive	The remuneration, motivation or compensation the crowd receives as a result of fulfilling the task.
Open call	The general invitation to become a member of the crowd, which will fulfil the designated task.
Platform	The technological platform that mediates the crowdsourcer and the crowd.

activities such as data analysis and quality checking, which in essence collect meta-data (data about data).

Vukovic, et al., (2010) discussed the kind of contract the crowd is involved in when doing the task, which could be internal, external or mixed. In this study, we assume most theory testing utilizes an external contract.

The crowd can also be classified according to the required domain knowledge. By domain knowledge we mean specialization in an area of knowledge necessary to complete a task (De Boer et al., 2012; Ross et al., 2009). The selection of the crowd members may depend on the task requiring domain knowledge or not, and how CS platforms support this feature. In this study, we do not consider the specific characteristics of CS platforms.

This study contributes to previous research on theory generation and testing by investigating how crowdsourcing can alleviate the challenges of data collection and analysis in theory testing. Using crowdsourcing as a research instrument has been around for some time. In particular, crowdsourcing is becoming a common way to deliver questionnaires in behavioural research (Bates & Lanza, 2013; Behrend et al., 2011; Jarmolowicz et al., 2012). In user studies, it has also started to be used as an instrument to collect user data (Kittur et al., 2008; Stewart et al., 2017), and it is also regarded as an important component in citizen science, where it supports the distributed collection of research data (Bonney et al., 2009; Gura, 2013). However, the relationship between crowdsourcing and theory testing has not yet been systematically explored and established.

By adopting CS, we see that the researcher has to fulfil certain expectations and obligations, which can be expressed as a set of requirements. To align theory testing and CS, it is therefore important to be explicit about such requirements, which serve as a checklist for judging the viability of using CS. Furthermore, the list of requirements also ensures uniformity when judging viability across many different studies. In Table 2.2 we elaborate a list of requirements we consider relevant for theory testing.

We have mandatory requirements – these are requirements that if one of the requirements is not met, then crowdsourcing will not be a viable option, they reflect conceptual and practical limitations that cannot be circumvented without breaking the whole notion of crowdsourcing (e.g., crowdsourcing without the Internet is not crowdsourcing). While the desirable requirements are requirement that even if one is not met, CS is possible, but to get a good result, it is better to meet such requirements. The mandatory requirements are mandatory because some of those requirements are what CS is made up of, and platform limitations and conditions while the desirable

requirements are also platform conditions but are flexible depending on what the researcher wants.

Our selection of mandatory requirements reflects a relatively consensual view in the related literature, but the list of desirable requirements can be seen as more controversial. For instance, many crowdsourcing tasks have been implemented using project websites and social media. However, we suggest that the costs of not using a crowdsourcing platform are just too high for theory testing purposes. Even though many crowdsourcing tasks have been implemented without any skills requirements (e.g. photo classification), in our context we see matching skills as very important to avoid uncontrolled setting and biased results, which could reduce data quality.

In our context, training can be a problem. In many research settings, especially in experimental research, some training is necessary, e.g. training participants to gather environmental data using specific instruments. However, it may be difficult to assess if training was or not successful, especially because tasks are one-off (so, no repeated training sessions). For that reason, we suggest that simple or no training should be the target.

2.3. LITERATURE REVIEW

A descriptive literature review is adopted as the research method for analysing different ways in which theory has been tested in the IS field. This method of literature review is used because it helps to determine the extent to which a body of empirical studies supports or reveals any interpretable trends or patterns with respect to pre-existing findings (Paré et.al., 2015). This is done by collecting, codifying and analysing data that reflects the frequency of the topics found in the literature (Paré et.al., 2015). The use of literature review is important because it grounds the research in the existing knowledge base. A similar approach was used by Thuan (2016) to identify factors influencing the decision to CS. A systematic approach is used instead of the narrative because it increases rigor and transparency and it is a well-structured and defined process (Kitchenham et al., 2009; Okoli & Schabram, 2010).

Table 2.2. *Crowdsourcing requirements*

Requirements	Description	Justification	Examples
Mandatory	Horizontal decomposability	The task must be decomposable into a number of simple, independent tasks that can be executed in parallel by the crowd	Task decomposition leads to faster turn-around time, which is one of the main advantages of CS (Heer & Bostock, 2010; Vukovic et al., 2010)
	One-off	The crowd task must be executed only once for each element of the crowd.	Repeated tasks are usually not allowed (platform limitation) (Mason & Suri, 2012; Yang et al., 2008)
	Limited interaction	The crowd task must be executed by the crowd using a minimal communication channel with the crowdsourcer	Communication loops are usually not possible (platform limitation) (Wexler, 2011; Wu et al., 2013)
	Computer mediation	The crowd task must be remotely executed, using the Internet for communication and coordination	Most CS platforms are web-based (Andriole, 2010; Kleemann et al., 2008)
	Simplicity	The crowd task must be easy to delineate and have clear inputs and outputs	Simplicity is considered one of the most important conditions for the success of CS (Geerts, 2009)
Desirable	Time-boundness	It should be possible to complete the crowd task in a bounded period of time	Timing is crucial to any CS task (Bernstein et al., 2011; Kittur et al., 2008)
	Skills matching	It should be possible to match the domain knowledge required by the crowd task with the skills of the selected members of the crowd	Mismatch of skills and tasks leads to low quality results (Geiger et al., 2011; Schenk & Guitard, 2011)
	Training	Training, if required, should not be complex or prolonged	In current CS platforms, training is still considered challenging (Park et al., 2014)

Our method of systematic literature review is based on (Kitchenham et al., 2009; Okoli & Schabram, 2010), who recommend searching the literature using the following stages: 1) select articles; 2) filter articles; 3) data extraction; and 4) data synthesis. These stages are described below.

Select articles

We are interested in articles that test theory to show the plausibility of such theory. This led to the selection of MIS Quarterly as the focus of the review. We selected this journal because it is known to have a strong theoretical basis and has a strong reputation. Furthermore, this journal follows a positivist tradition of research emphasising theory testing (Mingers, 2003). We selected a 10-year period for the review because this frame seems enough to develop a comprehensive view over the topic. The total number of articles published in the period was 307.

Articles were searched using a set of keywords, a procedure that other published reviews have adopted (Estellés-Arolas & González-Ladrón-De-Guevara, 2012). We searched for papers using a combination of keywords such as ‘theory’, ‘test’, ‘theoretical model’, ‘theoretical framework’, ‘conceptual model’, and ‘conceptual framework’. The search was done on the full body. Articles with any of these keywords were selected. The search identified 298 papers, which underlines the strong theoretical contents of MISQ and supports our selection.

Filter articles

This stage eliminated articles unrelated to this study. As mentioned earlier, we were interested in articles whose theory was tested one way or the other, and which theories belonged to type III and IV categories. Two intermediate steps were followed in this stage. Firstly, after keyword search, an article was taken from the pool and a check was done to determine if it contained a theory with testable propositions (Khan, 2011; Muntermann et al., 2015). We read through the methodology section to determine this. Secondly, we checked if the article was tested using human participants, since our research is restricted to theory testing using human participants. After this filtering stage, 248 articles were selected for coding. In Table 2.3 we summarise the results from the filtering process.

Table 2.3. Summary of Literature Review

Stages	Excluded	Selected
Articles published in MISQ between 2007 and 2017	-	307
Applying list of keywords	9	298
Checking for testable propositions and human participants	50	248
Selected articles	-	248

Data extraction

This step involved a detailed reading of specific sections of each article and coding of relevant data for posterior analysis. According to Paré et.al., (2016), there are three main aspects to consider during coding. The first one is deciding what parts of an article provide most relevant data. We focused primarily on the methodology, discussion and concluding sections. Coding was centred on the activities done by the researchers and related to the testing process.

We were not so much interested in the conceptual parts of theory testing, but more on the activities and patterns involved in the theory testing process. We were concerned with questions like what types of participants were used, what type of resources or instruments were used, what types of research methods and techniques were used, data collection methods, procedures followed during testing, and what were the outcomes of theory testing.

The second aspect considered by Paré et.al., (2016) involved operationalizing the coding procedure. We adopted a combination of open coding with predefined data extraction forms. Open coding was used to ensure that no relevant information would be left out, and to ensure the search was not confined to a small set of questions and assumptions. Data extraction forms were also used to ensure that pre-defined relevant elements could be extracted efficiently. Figure 2.2 shows a fragment of the data extraction form.

Detail of Coding on Theory testing using Crowdsourcing																			
General information				Research Methodology			Theory Type		Theory testing context										
No	Reference	Year	Date	Main findings (according to the data in the abstract)	Method of Data collection	Method of Data Analysis	Others	Type III and IV	Theory Tested	Others	Application	Task	Inputs	Agents	Function	Outputs	Participants	Complexity	Duplicated via the Internet
1	Web and wireless	2006	16-Jun	Hypothesis was tested using humans using the Microsoft Usability Guideline model, to test how participants rate 5 specific categories: content, ease of use, promotion, made-for-use and motivation.	Online and Physica	Statistical analysis		Yes	TAM /MUG Microsoft Usability guideline	Instrumen	Guideline: Data gath	Internet and cel	Simple	yes					

Figure 2.2. Sample of Coding Form

The third aspect considered by Paré et.al., (2016) concerned organizing the extracted data in preparation for analysis. NVivo was adopted to manage the extracted data. Coding was done by the first author. To increase the reliability of the analysis (Krefting, 1991), an independent researcher coded a sample selection of articles in parallel and the codes were compared and adjusted for clearer understanding.

Data synthesis

Data synthesis focussed on identifying patterned activities in theory testing. A pattern is a regular and intelligible form or sequence in which something is done. It can also be seen as a generalized solution to recurrent problems (Penker, 2000).

Some vital elements associated with patterns are *intent*, *structure* and *implementation* (Kafura, 1998; Penker, 2000). As implementation is usually very specific to a problem context, in this study we only consider the first two elements. Intent summarizes the general purpose of a study taking a theory testing perspective, which does not concern the specific problem under investigation (Penker, 2000). In our synthesis, we identify five different intents (Figure 2.3).

Structure regards the organisation of activities, which builds up pattern systems with linked activities. In our synthesis, we identified a large number of pattern systems involving 52 different activities (Figure 2.3). To designate the activities, we used generic names commonly recognized by scholars, e.g. interviews, site selection, cluster analysis, etc. The activities were related in patterned systems using directed arrows.

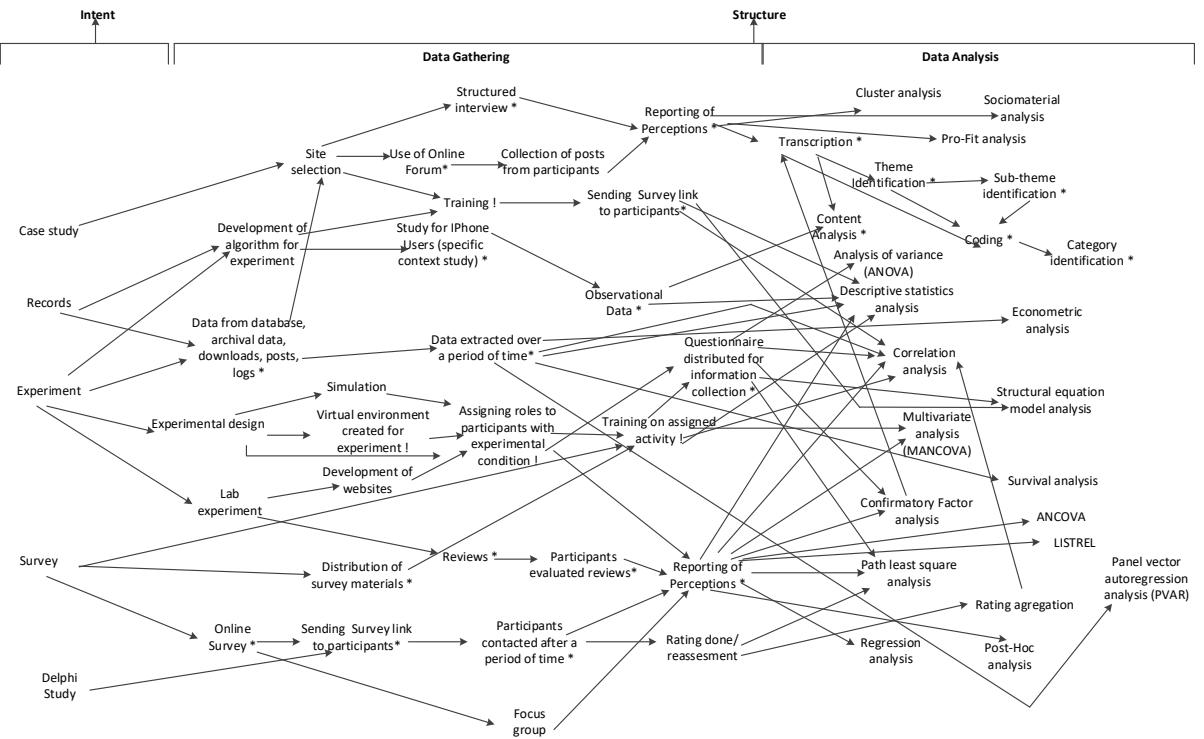


Figure 2.3. Activities and pattern systems for theory testing (* indicates the activity can be crowdsourced)

To emphasise the structural features of the collected data, we organised the pattern systems in two categories: data gathering and data analysis. These categories reflect the steps in theory testing discussed in Section 2.1.

Assessment of crowdsourcing requirements

We used the requirements shown in Table 2.2 to make a systematic decision if an activity could be CS or not. An activity is selected then a requirement check is made, if any of the mandatory requirement is not met and there is no way to modify or get an alternative, then that activity cannot be CS, but if all requirement is met, then the desirable requirements are considered. Such activity can be CS if some desirable requirement is not met, but it is advisable to consider those requirements for good output, this can be seen in figure 2.4.

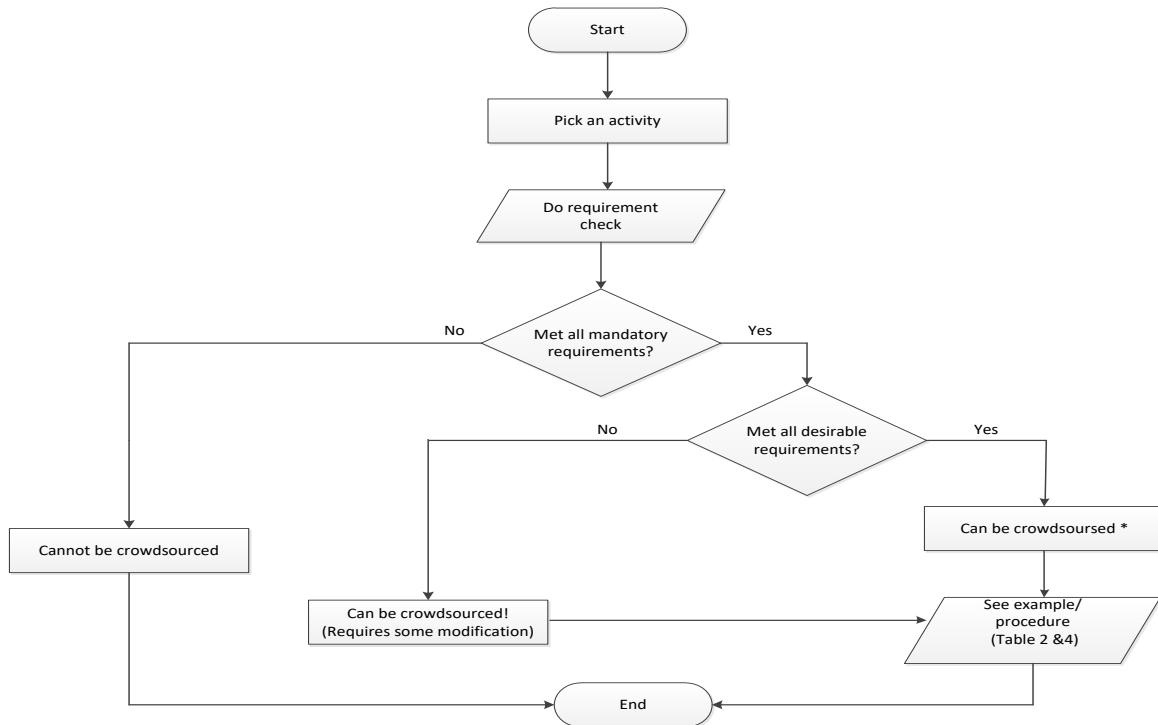


Figure 2.4.

Flowchart for Requirement Check

Some examples have been cited in table 2.2 and table 2.4 as a guideline, some examples could not be sourced for some activities, because CS has not been considered in such aspects, which is one of the main contributions of this study, to explore such areas as long as the CS requirements can be met. With the help of the flowchart and the requirement check, decisions were made for each activity to determine if it can be CS or not.

Take for instance the “case study” intent. We have to make a decision on each activity linked to that intent. Consider then “site selection”. We determine if this activity can be CS or not by going through the list of requirements. In this case, the activity is not decomposable and cannot be remotely executed using the Internet. Therefore, we can conclude this activity cannot be CS. We then pick another activity, which is “interview”. This activity can be decomposed in independent tasks, has clear inputs and outputs, can be remotely executed and communicated via the Internet, can be done within a defined time frame, and the domain knowledge of the crowd may be assessed in the interview. We can therefore conclude that this activity can be CS, which is denoted with an asterisk in Figure 2.3.

The results from this assessment are summarised in Figure 2.3. The symbol “*” close to an activity indicates that we found the activity as being able to be CS, as every requirement in Table 2.2 can be fulfilled. The symbol “!” close to an activity indicates that the activity can be CS if one or more desirable requirements are violated but the mandatory requirements can still be fulfilled. Although the desirable requirements may be violated and CS still possible, it is advisable to consider these requirements before deciding to CS as some of them are also critical if the best possible results are to be expected.

Throughout this process, we found some cases that could not meet all requirements but with some additional constraints which, if applied by the researcher, would allow for CS to be applied. Consider for instance the training requirement. The requirement if a study requires complex training, or training that cannot be recorded and uploaded but necessitates the researcher to do some sort of demonstration for the participants physically, then the activity cannot be CS. Though such a decision depends more on the researcher than on specific the characteristics of CS and theory testing. We denote these activities with (!) in the model shown in Figure 2.3.

Table 2.4. Crowdsourcing Activities and Their Examples

Activity	Description	Example/Procedure
Structured Interview	Presentation of interview with the same questions in the same order	(Brabham, 2010; De Vaus, 2013)
Perceptions recorded	Recording of perceptions from participants	(Snoek et al., 2010)
Online forum	Use of any form of online forum for data collection	(Brabham, 2009; Lampe et al., 2014)
Transcription	Process of producing a written copy of something	(Eveleigh, 2014; Holley, 2010)
Data from posts, database, downloads etc	Collection of data from posts, databases, downloads, logs etc.	(Lampe et al., 2014, Eveleigh, 2014, Chernova et al., 2011)
Virtual environment created for experiment	Use of virtual environment for experiment	(Chernova et al., 2011)
Assigning roles to participants	Roles are assigned to participants based on experimental condition	
reviews	Participants conducting reviews assigned to them	(Su et al., 2013)
Participants evaluates reviews		
Online survey	Use of online survey for data collection	

Survey link sent to participants	Sending of survey link to participants, through which data is collected	
Participants contacted after a period of time		(Brabham, 2008b; Zheng et al., 2011)
Training	Teaching, demonstrating or undertaking a course of exercise for skill needed for a particular task.	(Filatova, 2012), (Biemann, 2013)

2.4 FURTHER ANALYSIS AND DISCUSSION

Several researchers in IS and other disciplines have been adopting CS for delivering surveys (Kevin Crowston, 2012; Ghose et al., 2010; Goodman & Malkoc, 2012; Kim, 2012; Steelman et al., 2012; Ward & Broniarczyk, 2011). Our analysis suggests that CS may be used beyond that activity. As summarised in Table 2.5, we identified 5 intents involving 52 different activities. Table 2.5 shows in more detail how the activities are split between data gathering and data analysis, and within each category between CS and non-CS types. We observe that only 10 activities related to data gathering were classified as not able to be CS, and only 7 activities related to data analysis were classified as being able to be CS.

In Table 2.5 we provide a more detailed analysis of the theory testing systems of patterns using the notion of pathway. A pathway is a set of linked activities that instantiate an intent from beginning to end. Once again, we can split the exercise between data gathering and data analysis. Furthermore, we can divide pathways between fully CS, partially CS, and non-CS. For instance, a fully CS data gathering pathway allows researchers to CS every data gathering activity, while a partial CS pathway only allows them to CS some of the required data gathering activities.

Through this analytic process, we show that none of the 5 intents can be fully CS (both data gathering and data analysis). However, ‘experiments’, ‘surveys’ and ‘Delphi’ can fully CS the data gathering stage, while ‘case studies’, ‘records’ and ‘experiments’ can fully CS the data analysis stage. The detailed breakdown for the number of pathways that can be fully, partially or non-CS can be seen in Table 2.6.

Table 2.5. *Crowdsourcing Pathways*

Intents	# Path	# DG	# DA
Case study	17=0(F) + 17(P)	0(F) + 3(P)	5(F) + 6(N)
Records	10=0(F) + 10(P)	0(F) + 3(P)	1(F) + 8(N)
Experiment	98=0(F) + 98(P)	1(F) + 19(P)	6(F) + 47(N)
Survey	36=0(F) + 36(P)	3(F) + 2(P)	0(F) + 45(N)
Delphi	11=0(F) + 11(P)	1(F) + 1(P)	0(F) + 11(N)

$$Path(F) = DG(F) \wedge DA(F) \quad Path(P) = DG(P) \vee DA(P)$$

Where *Path* = pathway, *F* = Fully Crowdsourced, *P* = partially crowdsourced, *N* = Non-crowdsourceable, *DA* = Data Analysis, *DG* = Data Gathering.

We therefore suggest that the developed model, with its various pattern systems, gives researchers different pathways for testing their theories, which allow for circumventing activities that cannot be CS. Take for instance a ‘case study’. This pattern system considers all activities starting with site selection and then having training, online forum or structured interviews. We observe that parts of this pattern system may not be viable for CS, for example where training might be too demanding. However, we can still CS by circumventing ‘training’, which can be done by either adopting an online forum activity or sending survey links to participants.

Table 2.6. *Crowdsourcing Activities*

Intents	Data Gathering		Data Analysis	
	# CS Activities	# Non-CS activities	# CS Activities	# Non-CS activities
Case study	4	2	6	6
Records	5	1	1	8
Experiment	8	4	-	14
Survey	8	2	-	13
Delphi	3	1	-	11

Another interesting finding is the fact that some aspects of data analysis can be CS, even though most of the activities relate to qualitative methods, since quantitative methods for analysis do not seem amenable for horizontal decomposition. This finding

can be helpful for those engaging in data analysis using any of these activities. Using CS for these activities may increase the reliability of the research process and could also help conducting validity checks by getting different people to code the same data to reduce potential researcher bias and increase inter-coder reliability.

Relating the results to existing related research, this study contributes to previous research by identifying conceptual elements of theory testing. Often theory testing is embedded in theory building, which makes it difficult to handle it as an autonomous process. To avoid this problem, we clearly identify the core elements of theory testing, independently from theory building. We also explored how theory testing has been conducted by researchers. An understanding of the types of theory testing tasks lays a foundation for a discussion on how to crowdsource them. We also conceptually characterised crowdsourcing such that potential matches to theory testing tasks is identified. We identified the fundamental requirements for activities to be a candidate for crowdsourcing support. We then used these requirements to systematically assess what types of theory testing tasks have already been crowdsourced or can be crowdsourced. Finally, we made recommendations for research. We particularly focus on researchers developing theory who are considering different strategies on how to test their theory. For this purpose, we discussed various theory testing patterns and identified what patterns can be fully crowdsourced, partially crowdsourced, or not crowdsourced at all.

2.5 CONCLUDING REMARKS

Theory testing, which involves the validation or falsification of a theory, can be difficult to accomplish, because a good theory needs good data for validation and good data is usually hard to obtain. CS can be used for theory testing in cases where humans are involved in the testing process. However, one has to consider how to align the theory testing goals with the constraints imposed by CS. This alignment problem has been the focus of our study.

To better understand the problem, we analysed how theories have been tested in the past. MISQ was selected as a target for a systematic literature review, because of its high reputation and editorial focus on theory. 248 articles published in the last ten years in MISQ were analysed. In our analysis, we focussed on the operational

dimension of theory testing and not on the particulars of each theory. We then identified the intents and pattern systems of theory-testing activities adopted by each study. The activities were then organised in a model that underlines three groupings: intent, data gathering and data analysis. We also analysed how the activities were actioned in their respective studies, we identified a set of links and pathways for theory testing. The combination of activities and links identifies a set of pattern systems. We then analysed which activities and pattern systems could be CS or not.

This study provides two important insights. One is highlighting how CS can be used across a multitude of activities related to theory testing, using data from prior research. The other one is suggesting that most theory testing intents can be fully CS. Even though some intents may involve pathways that cannot be CS (fully or partially), our analysis shows that in most cases there are alternative pathways that can be CS. Researchers can therefore make the decision to choose pathways that are suitable for their study. This study does not cover the aspect of decision-making, i.e. motives for crowdsourcing and when it's best to crowdsource. For future study, developing an algorithm and decision tree for pathway selection is an aspect to investigate.

2.6 REFERENCES

- Andriole, S. J. (2010). Business impact of Web 2.0 technologies. *Communications of the ACM*, 53(12), 67-79.
- Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of management review*, 14(4), 496-515.
- Bates, J., & Lanza, B. (2013). Conducting psychology student research via the Mechanical Turk crowdsourcing service. *North American Journal of Psychology*, 15(2), 385.
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior research methods*, 43(3), 800.
- Bernstein, M. S., Brandt, J., Miller, R. C., & Karger, D. R. (2011). *Crowds in two seconds: Enabling realtime crowd-powered interfaces*. Paper presented at the Proceedings of the 24th annual ACM symposium on User interface software and technology.
- Bhattacherjee, A. (2012). Social science research: principles, methods, and practices. *Textbooks Collection, Book 3*(http://scholarcommons.usf.edu/oa_textbooks/3).
- Biemann, C. (2013). Creating a system for lexical substitutions from scratch using crowdsourcing. *Language Resources and Evaluation*, 47(1), 97-122.

- Bitektine, A. (2008). Prospective case study design: qualitative method for deductive theory testing. *Organizational Research Methods*, 11(1), 160-180.
- Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11), 977-984.
- Brabham, D. C. (2008a). Crowdsourcing as a model for problem solving an introduction and cases. *Convergence: the international journal of research into new media technologies*, 14(1), 75-90.
- Brabham, D. C. (2008b). Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application. *First monday*, 13(6).
- Brabham, D. C. (2009). Crowdsourcing the public participation process for planning projects. *Planning Theory*, 8(3), 242-262.
- Brabham, D. C. (2010). Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. *Information, Communication & Society*, 13(8), 1122-1145.
- Chernova, S., DePalma, N., Morant, E., & Breazeal, C. (2011). *Crowdsourcing human-robot interaction: Application from virtual to physical worlds*. Paper presented at the RO-MAN, 2011 IEEE.
- Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal*, 50(6), 1281-1303.
- Crowston, K. (2012). Amazon mechanical turk: A research tool for organizations and information systems scholars *Shaping the Future of ICT Research. Methods and Approaches* (pp. 210-221): Springer.
- De Boer, V., Hildebrand, M., Aroyo, L., De Leenheer, P., Dijkshoorn, C., Tesfa, B., & Schreiber, G. (2012). *Nichesourcing: Harnessing the power of crowds of experts*. Paper presented at the International Conference on Knowledge Engineering and Knowledge Management.
- De Vaus, D. (2013). *Surveys in social research*: Routledge.
- Estellés-Arolas, E., & González-Ladrón-De-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information science*, 38(2), 189-200.
- Eveleigh, A. (2014). Crowding out the archivist? Locating crowdsourcing within the broader landscape of participatory archives. *Crowdsourcing our Cultural Heritage*, 211-212.
- Filatova, E. (2012). *Irony and Sarcasm: Corpus Generation and Analysis Using Crowdsourcing*. Paper presented at the LREC.
- Følstad, A. (2008). Living labs for innovation and development of information and communication technology: A literature review. *eJOV: The Electronic Journal for Virtual Organization & Networks*, 10.
- Ford, R. C., Richard, B., & Ciuchta, M. P. (2015). Crowdsourcing: A new way of employing non-employees? *Business Horizons*, 58(4), 377-388.
- Geerts, S. (2009). Discovering crowdsourcing: theory, classification and directions for use. *unpublished Master of Science in Innovation Management thesis, Eindhoven University of Technology*.
- Geiger, D., Seedorf, S., Schulze, T., Nickerson, R. C., & Schader, M. (2011). *Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes*. Paper presented at the AMCIS.

- Gerring, J. (2006). *Case study research: Principles and practices*: Cambridge University Press.
- Ghose, A., Ipeirotis, P. G., & Li, B. (2010). *Designing Ranking Systems for Hotels on Travel Search Engines to Enhance User Experience*. Paper presented at the ICIS.
- Goode, W. J., & Hatt, P. K. (1952). Methods in social research.
- Goodman, J. K., & Malkoc, S. A. (2012). Choosing here and now versus there and later: The moderating role of psychological distance on assortment size preferences. *Journal of Consumer Research*, 39(4), 751-768.
- Gregor, S. (2002). A theory of theories in information systems. *Information Systems Foundations: building the theoretical base*, 1-20.
- Gregor, S. (2006). The nature of theory in information systems. *MIS quarterly*, 611-642.
- Gura, T. (2013). Citizen science: Amateur experts. *Nature*, 496(7444), 259-261.
- Heer, J., & Bostock, M. (2010). *Crowdsourcing graphical perception: using mechanical turk to assess visualization design*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Holley, R. (2010). Crowdsourcing: How and why should libraries do it? *D-Lib Magazine*, 16(3/4 Ma).
- Howe, J. (2006). The rise of crowdsourcing. *Wired magazine*, 14(6), 1-4.
- Hyde, K. F. (2000). Recognising deductive processes in qualitative research. *Qualitative market research: An international journal*, 3(2), 82-90.
- Jarmolowicz, D., Bickel, W., Carter, A., Franck, C., & Mueller, E. (2012). Using crowdsourcing to examine relations between delay and probability discounting. *Behavioural processes*, 91(3), 308-312.
- Jarmolowicz, D. P., Bickel, W. K., Carter, A. E., Franck, C. T., & Mueller, E. T. (2012). Using crowdsourcing to examine relations between delay and probability discounting. *Behavioural processes*, 91(3), 308-312.
- Kafura, D. (1998). *Object-oriented software design and construction with C++*: Prentice-Hall, Inc.
- Khan, J. A. (2011). *Research methodology*: APH Publishing.
- Kim, J. (2012). The effect of design characteristics of mobile applications on user retention: an environmental psychology perspective.
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1), 7-15.
- Kittur, A., Chi, E. H., & Suh, B. (2008). *Crowdsourcing user studies with Mechanical Turk*. Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems.
- Kleemann, F., Voß, G. G., & Rieder, K. (2008). Un (der) paid innovators: The commercial utilization of consumer work through crowdsourcing. *Science, technology & innovation studies*, 4(1), PP. 5-26.
- Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American journal of occupational therapy*, 45(3), 214-222.
- Lampe, C., Zube, P., Lee, J., Park, C. H., & Johnston, E. (2014). Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly*, 31(2), 317-326.
- Lee, A. S., & Hubona, G. S. (2009). A scientific basis for rigor in information systems research. *MIS quarterly*, 237-262.

- Lin, J., Amini, S., Hong, J. I., Sadeh, N., Lindqvist, J., & Zhang, J. (2012). *Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing*. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.
- Lokke, A., & Sorensen, P. (2014). Theory testing using case studies. *Electronic Journal of Business Research Methods*, 12(1), 66-74.
- Lynham, S. A. (2002). The general method of theory-building research in applied disciplines. *Advances in developing human resources*, 4(3), 221-241.
- MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, 35(2), 293-334.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods*, 44(1), 1-23.
- Mingers, J. (2003). The paucity of multimethod research: a review of the information systems literature. *Information Systems Journal*, 13(3), 233-249.
- Muntermann, J., Nickerson, R., & Varshney, U. (2015). Towards the Development of a Taxonomic Theory. Paper presented at the 21st Americas Conference on Information Systems, Puerto Rico.
- Niederman, F., & March, S. (2015). Reflections on Replications. *AIS Transactions on Replication*.
- Nunamaker Jr, J. F., Chen, M., & Purdin, T. D. (1990). Systems development in information systems research. *Journal of management information systems*, 7(3), 89-106.
- Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts Work. Pap. Inf. Syst*, 10, 26.
- Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. *European Journal of Information Systems*, 1-16.
- Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183-199.
- Park, S., Shoemark, P., & Morency, L.-P. (2014). *Toward crowdsourcing micro-level behavior annotations: the challenges of interface, training, and generalization*. Paper presented at the Proceedings of the 19th international conference on Intelligent User Interfaces.
- Parvanta, C., Roth, Y., & Keller, H. (2013). Crowdsourcing 101 a few basics to make you the leader of the pack. *Health promotion practice*, 14(2), 163-167.
- Penker, M. (2000). Business Modeling with UML: Business Patterns at Work: New York, NY: John Wiley & Sons.
- Piper, S. (2006). Qualitative theory testing as mixed-method research. *Journal of Research in Nursing*, 11(3), 183-193.
- Ross, J., Zaldivar, A., Irani, L., & Tomlinson, B. (2009). Who are the turkers? worker demographics in amazon mechanical turk. *Department of Informatics, University of California, Irvine, USA, Tech. Rep.*
- Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics & Management*(1), 93-107.
- Smith, D., Manesh, M. M. G., & Alshaikh, A. (2013). How can entrepreneurs motivate crowdsourcing participants? *Technology Innovation Management Review*, 3(2).

- Snoek, C. G., Freiburg, B., Oomen, J., & Ordelman, R. (2010). *Crowdsourcing rock n'roll multimedia retrieval*. Paper presented at the Proceedings of the 18th ACM international conference on Multimedia.
- Steelman, Z., Soror, A., Limayem, M., & Worrell, D. (2012). Obsessive compulsive tendencies as predictors of dangerous mobile phone usage. Paper presented at AMCIS.
- Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing Samples in Cognitive Science. *Trends in cognitive sciences*, 21(10), 736-748.
- Stinchcombe, A. L. (1987). *Constructing social theories*: University of Chicago Press.
- Su, A. I., Good, B. M., & van Wijnen, A. J. (2013). Gene Wiki Reviews: marrying crowdsourcing with traditional peer review. *Gene*, 531(2), 125.
- Thuan, N. H., Antunes, P., & Johnstone, D. (2016). Factors influencing the decision to crowdsource: A systematic literature review. *Information Systems Frontiers*, 18(1), 47-68.
- Turner, A. M., Kirchhoff, K., & Capurro, D. (2012). Using crowdsourcing technology for testing multilingual public health promotion materials. *Journal of medical Internet research*, 14(3).
- Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, 48(4), 865-885.
- Vukovic, M., Laredo, J., & Rajagopal, S. (2010). *Challenges and experiences in deploying enterprise crowdsourcing service*. Paper presented at the International Conference on Web Engineering.
- Vukovic, M., Lopez, M., & Laredo, J. (2010). *Peoplecloud for the globally integrated enterprise*. Paper presented at the Service-Oriented Computing. ICSOC/ServiceWave 2009 Workshops.
- Ward, M. K., & Broniarczyk, S. M. (2011). It's not me, it's you: How gift giving creates giver identity threat as a function of social closeness. *Journal of Consumer Research*, 38(1), 164-181.
- Wexler, M. N. (2011). Reconfiguring the sociology of the crowd: exploring crowdsourcing. *International Journal of Sociology and Social Policy*, 31(1/2), 6-20.
- Whitla, P. (2009). Crowdsourcing and its application in marketing activities. *Contemporary Management Research*, 5(1).
- Witschey, J., Murphy-Hill, E., & Xiao, S. (2013). *Conducting interview studies: Challenges, lessons learned, and open questions*. Paper presented at the Conducting Empirical Studies in Industry (CESI), 2013 1st International Workshop on.
- Wu, W., Tsai, W.-T., & Li, W. (2013). Creative software crowdsourcing: from components and algorithm development to project concept formations. *International Journal of Creative Computing*, 1(1), 57-91.
- Yang, J., Adamic, L. A., & Ackerman, M. S. (2008). *Crowdsourcing and knowledge sharing: strategic user behavior on taskcn*. Paper presented at the Proceedings of the 9th ACM conference on Electronic commerce.
- Yuen, M.C., King, I., & Leung, K.S. (2011). *A survey of crowdsourcing systems*. In 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing (pp. 766-773). IEEE.

Zhao, Y., & Zhu, Q. (2012). Exploring the motivation of participants in crowdsourcing contest. Paper presented at the 33rd International Conference on Information Systems, Orlando.

Zheng, H., Li, D., & Hou, W. (2011). Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce*, 15(4), 57-88.

CHAPTER 3

TOWARDS THE DEVELOPMENT OF A DSS SUPPORTING THE INTEGRATION OF CROWDSOURCING IN THEORY TESTING: CONCEPTUAL FRAMEWORK AND MODEL

ABSTRACT

The information systems (IS) discipline has not accorded the same attention to theory testing as it has to theory building. Further, crowdsourcing presents rich opportunities for the theory testing process that have not been fully explored. This paper builds on previous work, employing a design science research (DSR) paradigm in order to develop a decision support system artefact that will help early career researchers identify viable theory testing approaches, and how crowdsourcing can help facilitate the testing process. As part of the DSR build/evaluate cycle, this paper presents a conceptual framework and model of theory testing in IS, and the problem frame in which they are situated is evaluated using Schön's theory of reflective practice and problem/solution framing. Data collected from PhD students revealed an incomplete level of knowledge of theory testing, and a lack of awareness of the possibilities provided by adopting a crowdsourcing strategy.

Keywords Theory building, theory testing, crowdsourcing, decision support.

3.1 INTRODUCTION

Theory testing is concerned with establishing or refuting the validity of a theory's core propositions, which are tested by determining to what degree they provide a close fit to empirical data (Colquitt & Zapata-Phelan, 2007). The key foundations to advance scientific knowledge are testing and re-testing theory (Niederman & March, 2015b).

Often, the theory testing process involves human participation. People can play different roles in theory testing, such as collecting data on behalf of the researcher, checking data quality, mediating access to data, analysing data, etc. Although many robust methods have already been developed to test theories, including interviews, surveys, ethnography and many others, the adoption of crowdsourcing (CS) as a

means to facilitate these methods has recently started to receive considerable attention (Lowry et al., 2016).

CS is an emerging strategy that fragments a single task into multiple tasks delivered to a large group of people in an open call (Howe, 2006). Several reasons have lead researchers towards considering the adoption of CS in theory testing. For instance, Lowry et al. (2016) advocated CS as a way to increase the quality of data collection. In particular, the authors reviewed some criticisms raised against CS and concluded that most of these were also associated with other data collection methods. Steelman et al. (2014) also highlighted the advantages brought by CS concerning the demographics, psychometrics, and structural properties of data samples. Their research also provided initial empirical evidence that CS can tap into large samples of participants and reach a wide variety of demographics.

Especially in the psychology field, researchers have been open-minded on the use of CS (Lowry et al., 2016). Therefore, this field has been leading the way, particularly with the use of the Amazon Mechanical Turk (AMT) platform for data collection. In the information systems (IS) field, CS has also started to be seen as a valid medium for this purpose. For instance, Steelman et al. (2014) identified 20 quality IS publications that used CS.

Nevertheless, it is one thing to know that others have already successfully used CS in their research, and it is another to know *how* CS can actually be involved in the research process, especially because of the diversity of the IS field. We believe there is a need for helping researchers to understand and make decisions on how to integrate CS into theory testing. This is especially true with PhD students, because often they are engaged for the first time in theory testing, and theory testing is typically on the critical path to conclude their PhD.

This type of support is usually associated with a category of tools designated as Decision Support Systems (DSSs) (Arnott & Pervan, 2005; Hosack et al., 2012a). Therefore, the challenge we undertake is the *development of a DSS that helps IS researchers making decisions about how to integrate CS in the process of testing IS theory*. We adopted the design science paradigm and Schön's (1983) theory of practice to address this challenge. They both emphasize a developmental viewpoint

where the goal is to iterate problem and solution frames. Considering the complexity of this challenge, in this paper, we only report on two problem frames necessary to build the DSS, which include the development of a *conceptual framework* and a *model*. This would be followed by the development of a DSS artefact, in the form of a DSS tool for theory testing, possibly in the form of a simple spreadsheet or a more complex decision tool (Thuan et al., 2015).

The remainder of this paper is organized as follows. Section two describes the adopted research method. Section three describes the conceptual framework and model. In Section four, we assess our problem frames using the card sorting method. Finally, Section five highlights the research contributions and draws implications for the DSS development.

3.2 APPROACH

The design science paradigm was adopted for this research. Design science is a problem-solving paradigm that seeks to create innovative IS artefacts that solve practical problems (Gregor & Hevner, 2013). Design science involves building and evaluating artefacts while making significant research contributions (Peffers et al., 2007). By artefacts we mean anything that can be transformed into a material existence as an artificially made object or process (Gregor & Hevner, 2013). Artefacts may also include social innovations (Aken, 2004), or new properties of technical, social, or informational resources (Järvinen, 2007). In essence, any innovative artefact providing a solution to a research problem can be seen as design science (Peffers et al., 2007). We adopted this paradigm because: 1) It reflects a developmental perspective over the identified challenge and proposed solution; 2) It is considered the most common approach to DSS development (Arnott & Pervan, 2012); and 3) It has been accepted by the IS community as a mainstream research paradigm Hevner & Chatterjee, 2010).

Along with design science, we adopted the iterative problem/solution framing suggested by Schön (1983). According to Schön's theory of practice, design is a reflective conversation with the design situation (Schön, 1992), where problem frames suggest solutions and experience with solutions suggest new problem frames. This

iterative process leads to the development of primary and secondary artefacts, the former addressing the fundamental research challenge and the latter guiding and supporting the journey (Hevner & Chatterjee, 2010; Hevner et al., 2004). Secondary artefacts are relevant because they provide actionable knowledge (Argyris, 1996), and in their own way they also frame the problem. In our case, the primary artefact is the DSS, while the secondary artefacts consist of a conceptual framework and model.

Design science research (DSR) involves two primary research activities: build and evaluate (March & Smith, 1995). The build activity involves developing an artefact based on a problem frame. The build activity is followed by an evaluation activity, which confronts the artefact with the problem frame and suggests new problem frames, until a satisficing solution is obtained (Sein et al., 2011). Considering the particular nature of our research challenge and its specific context, we can then delineate a more detailed research method, which is described next.

3.2.1 Development Method

The adopted method was based on a detailed analysis of two methods previously described in the related literature and specifically used to develop DSSs. One such method has been suggested by Arnott and Pervan (2012). The other method was developed by Thuan et al. (2016). The method developed by Thuan et al. (2016), named SCoA, regards design as a verb (activities) and considers the following steps: 1) scope the knowledge sources; 2) develop a conceptual model; 3) develop an ontology; and 4) develop a DSS that supports users in exploring the ontology. To ensure rigour, SCoA scaffolds the development within existing knowledge sources. It also ensures that the researcher adequately frames both the problem and solution using a conceptual model and ontology, from which the DSS then logically emerges.

The method discussed by Arnott and Pervan (2012) is more centred on design as a noun (artefacts). It suggests several categories of artefacts: 1) constructs; 2) models; 3) methods; and 4) instantiations. These categories directly reflect a categorization of design science artefacts proposed by March and Smith (1995). Arnott and Pervan (2012) further suggest that there are logical and purposeful relationships between constructs, models and instantiations, which define an iterative construction method. That is, constructs lead to models, which then lead to instantiations. We may therefore

regard instantiations as primary artefacts, and constructs and models as secondary artefacts.

Our method defines four conceptual elements mixing verbs and nouns: 1) conceptual model; 2) scoping knowledge source; 3) model; and 4) DSS instantiation. These elements are graphically represented in Figure 3.1. The sequential linkage represents the development order, i.e. the conceptual model precedes the scoping knowledge source, which precedes the model. This method has similarities with SCoA, with a shift between the scoping of knowledge source and conceptual model. This change is necessary because our problem is framed by existing theory and epistemology on theory building and theory testing. Regarding differences to the method discussed by Arnott and Pervan (2012), we essentially bring a verb (scoping knowledge source) in between the collection of nouns. The similarities and differences between these methods are summarized in Figure 3.1. Next, we provide additional details about these steps.

Step 1 – Build a conceptual framework. This step aims to identify the main concepts and constructs of the application domain, thus providing a holistic problem frame (Sonnenberg & vom Brocke, 2011). As there are multiple ways in which a problem can be framed, an attempt to solve a problem should start with a particular viewpoint (Holmström et al., 2009). In our particular case, the conceptual framework highlights theory testing as a distinct component of theory building.

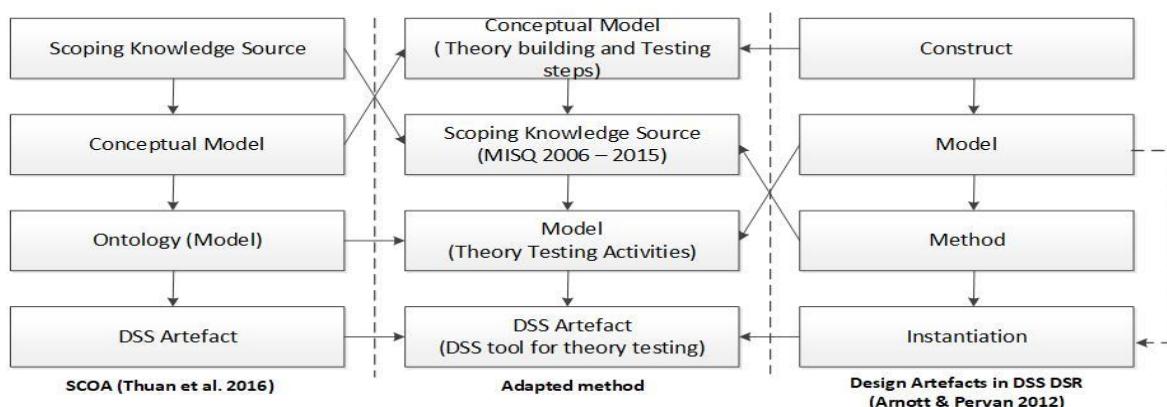


Figure 3.1. The adapted method based on SCoA and DSS DSR

Step 2 - Scope the knowledge sources. Scoping the knowledge sources seeks to extract and articulate the existing knowledge related to the DSS that one seeks to develop (Thuan et al., 2016). It establishes a relationship between the DSS and the state-of-the-art. This is considered important for design science because it contributes to rigour, which should be on par with relevance (Gregor & Hevner, 2013; Peffers et al., 2007).

The descriptive literature review method was adopted to identify the different ways in which theory has been tested in the IS field. This method was selected because it helps to determine the extent to which a body of empirical knowledge supports or reveals any interpretable trends or patterns with respect to pre-existing findings (Paré et al., 2015). This was done by collecting, codifying and analysing data that reflects the frequency of relevant topics found in the literature (Paré et al., 2015).

A systematic approach to data collection was adopted to increase rigor and transparency (Kitchenham et al., 2009; Paré et al., 2016). We also followed the data collection guidelines suggested by Kitchenham et al. (2009): 1) select articles; 2) filter articles; 3) data extraction; and 4) data synthesis. These steps are explained in detail in a prior paper Enwereuzo et al., 2018 and will not be repeated here.

Step 3 – Build the model. The developed model represents the theory testing patterns and activities identified in the previous step. This is essentially the result of an analytic process. Our interest was not so much on the conceptual parts of theory testing, but more on the activities and patterns involved in the theory testing process, as presented by researchers to the community through published articles. Questions like what types of participants were used, what types of resources or instruments were used, what types of research methods and techniques were used, data collection methods, procedures followed during testing, and what were the outcomes of theory testing, were our concern when building the model. The model also highlights where CS can be used within the patterned activities.

Step 4 – Build the DSS. This step concerns the development of a solution addressing the framed problem. This can be realized in a variety of ways and can also be iterative (Thuan et al., 2015). For instance, the solution may go through conceptual design,

prototyping and instantiation. As previously noted, in this paper, we do not detail this step.

3.2.2 Evaluation Method

Evaluation is a primary consideration in DSR, as it determines the utility of the developed artefacts (Hevner et al., 2004). Since design science is still a relatively young paradigm, there has been some ongoing discussion regarding the breadth, depth and scheduling of evaluation in DSR (Sonnenberg & vom Brocke, 2011). In particular, should the evaluation be centred on the last stage of the research, or done multiple times during the project? Should it concern the primary artefact or also consider secondary artefacts, and in the latter case, what is the purpose of evaluating secondary artefacts? Furthermore, what is the appropriate balance between rigour and utility?

In fact, DSR evaluation does not have to be limited to a single, summative evaluation of the primary artefact. Evaluation actions can be conducted along with the development of secondary artefacts. Furthermore, since secondary artefacts essentially serve to iteratively frame the problem, the evaluation actions may assume a more formative purpose, generating justificative knowledge (Sonnenberg & vom Brocke, 2011). This justificative knowledge serves to either consolidate or adjust the problem frame. For these reasons, and to avoid misconceptions in understanding the purpose and target of the evaluation, we will use the term problem frame assessment (PFA). In this paper, we report on PFA regarding the conceptual framework and model.

3.3 DEVELOPMENT STEPS

3.3.1 Conceptual Framework

The conceptual framework is based on literature concerning both theory building and theory testing. We adopted the propositions by De Vaus (2013) and Bitektine (2007), which focus on the abstract set of activities necessary to build a theory, and then isolated the activities that specifically concern theory testing. Theory testing can then be characterised as two consecutive activities: operationalization and validation.

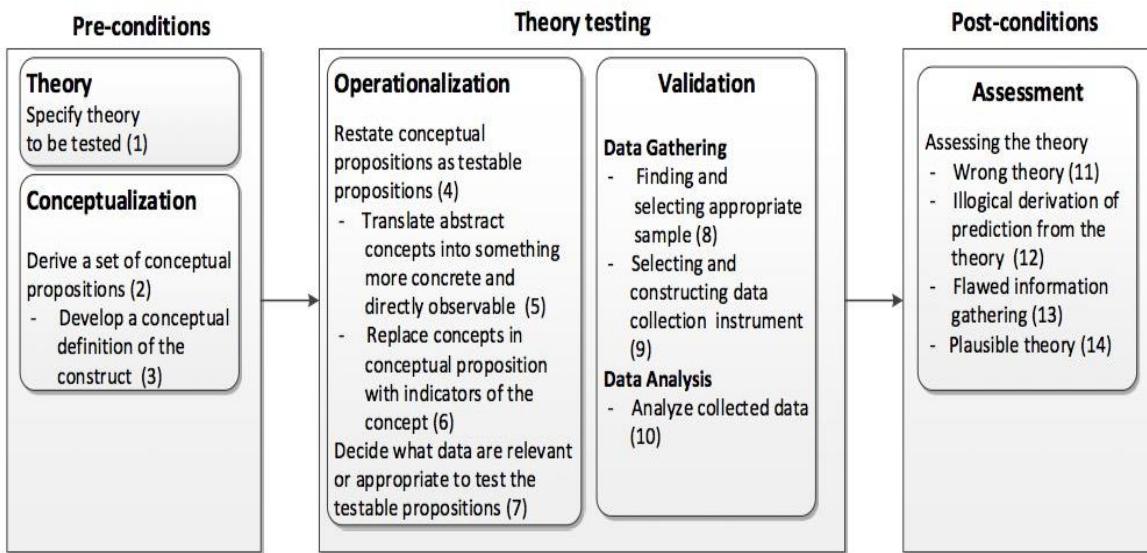


Figure 3.2. Conceptual framework, adapted from (Bitektine, 2007; De Vaus, 2013; MacKenzie et al., 2011; Niederman & March, 2015)

Considering operationalization, one has to restate the conceptual propositions as testable propositions. This involves translating abstract concepts into concrete, observable and measurable variables (De Vaus, 2013). Then, one has to decide what data are relevant or appropriate to test the propositions. Regarding validation, one has to collect relevant data and analyse it. Data are analysed to see: a) how much support there is for the testable propositions; b) how much support there is for the conceptual propositions; and c) how much support there is for the whole theory. The framework is shown in Figure 3.2.

Though addressing theory testing independently from theory building seems a reasonable way to frame the problem, we should nevertheless recognise that there is a level of dependence on theory building. Therefore, we define pre- and post-conditions to highlight the dependencies on theory testing. Furthermore, since many theories in social sciences are qualitatively tested, especially using the case study method, we integrate the recommendations from Bitektine (2007) into the conceptual framework, which specifically concern pre- and post-conditions related to case studies.

As pre-conditions, we consider: 1) briefly state what the theory is and what it intends to achieve; 2) provide a set of conceptual propositions (statements that show relationships between factors); and 3) when the research involves a case study, identify the case(s), unit of analysis and subjects.

As post-conditions, we consider: 1) state an assessment of the whole theory; and 2) when the research involves a case study, analyse the internal and/or external validity, realising that single-case studies tend to generate type-1 errors (accept false hypotheses) while cross-case studies tend to generate type-2 errors (reject true hypotheses) (Gerring, 2006).

3.3.2 Model

The model is based on a scoping literature review conducted to understand the diversity of theory testing activities reported in the literature. The review uses a set of 248 papers published in a ten-year period in MISQ, which is considered the leading journal in the IS field and has a particular editorial focus on theory building.

The model organises the identified activities in a pattern system (Figure 3.3). A pattern is a sequence of regular and intelligible ways in which something is done. It can also be seen as a generalized solution to recurrent problems (Penker, 2000).

Some vital elements associated to patterns are *intent* and *structure* (Penker, 2000). Intent summarizes the general purpose of a study, taking a theory testing perspective, which does not concern the specific problem under investigation (Penker, 2000). Structure considers how researchers put together some activities to reach their research intents. An intent may have one or more patterns, each one defining a particular sequence that allows fulfilling the intent. These patterns do not originate from a theoretical perspective but from practice, based on what researchers have reported in the literature. The whole collection of intents, activities and links then defines a complex, comprehensive pattern system, which suggests different ways of doing.

From the data gathered in the scoping literature review, we identified a large number of intents, activities and patterns. We found five intents: case study, records, experiment, survey, and Delphi study. We also found 53 different activities, which have been divided into two groups: data gathering and data analysis. The activities were then linked together, with the intents, to show the patterned characteristics of theory testing (Figure 3.3).

The model emphasises the diversity of ways in that a theory can be tested. For instance, the survey intent can start in three different ways: training on assigned activity, distribution of survey material, and online survey. Then, it can take many different pathways. For instance, the survey can be followed by a survey link sent to participants or a focus group.

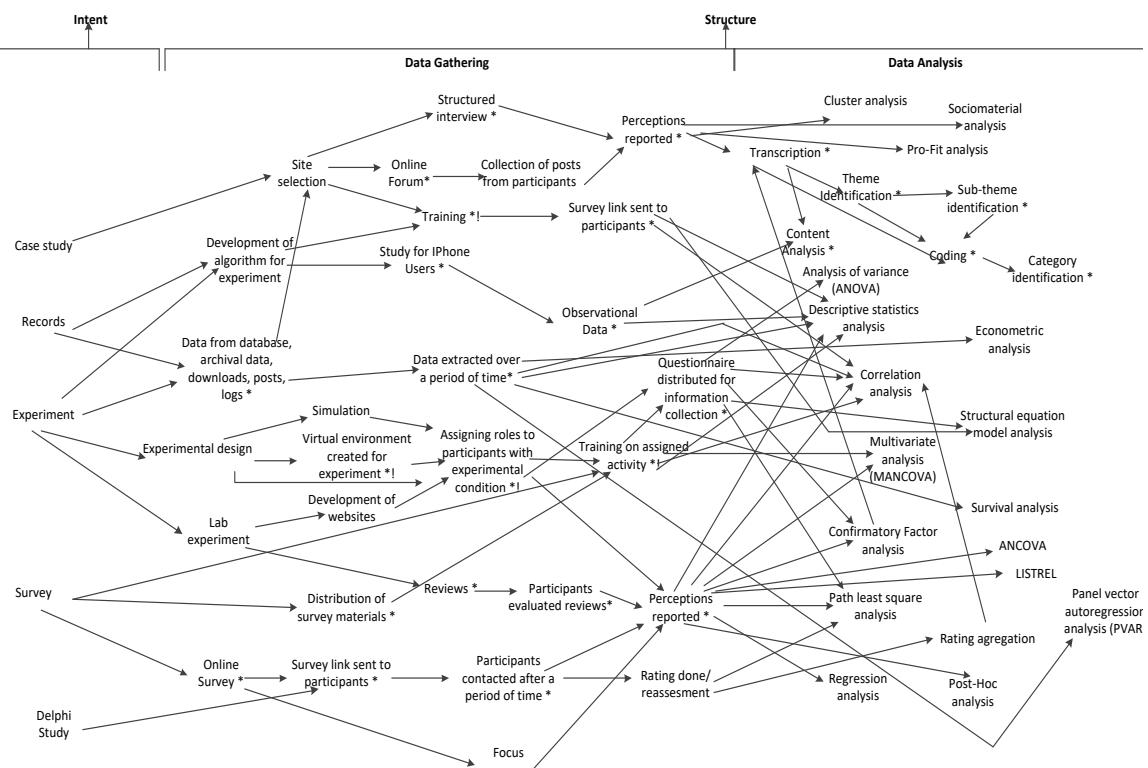


Figure. 3.3. Model showing theory testing patterns and activities (* means the activity can be crowdsourced)

A final step in model building concerned analysing which activities can be CS. The analysis was based on a set of rules: 1) the activity should be decomposable in a number of simple CS tasks; 2) the CS tasks should be executed in a one-off event; 3) the CS tasks should be simple to understand; 4) the CS tasks should have clear inputs and outputs; 4) limited interaction with the crowd should be required; 5) the CS task should be remotely executed; and 6) the CS task should be completed in a bounded period of time. The activities that can be crowdsourced are identified in Figure 3 using an asterisk (*).

3.4 PROBLEM FRAME ASSESSMENT

Card sorting is an inexpensive and reliable method that has been widely used in various fields such as psychology, knowledge engineering, and software engineering (Barrett & Edwards, 1995). The method helps understanding the people we are designing for (Donna Spencer, 2009), considering in particular how they structure information and action (Spencer & Warfel, 2004). For instance, it can be used to assess the users' needs and priorities, how they deal with information, and how they react to tool features. Finally, it can offer more insights into the users' view about a problem and a solution. All in all, card sorting seems adequate to assess if our problem frame relates to what exists in the participants' minds and if our solution would be useful to them.

We structured PFA using four card sorting exercises followed by interviews. The exercises were conceived to acquire justificatory knowledge regarding:

E1: How familiar the participants are with theory testing activities. By allowing the participants to externalise their own conceptual frameworks, we can assess if they are close to best practices reported in the literature. A negative result suggests the DSS could help better planning theory testing.

E2: How familiar the participants are with theory testing intents and patterns. Once again, by allowing the participants to externalise their own views, we can assess if they are closer to what others have reported in the literature. A negative result suggests the DSS could help selecting theory testing patterns.

E3: What activities the participants think could be CS. This data may again suggest if the DSS could help integrating CS in theory testing.

E4: Which difficulties the participants face during theory testing and how a DSS could be more beneficial to their research. This exercise contributes a list of requirements for DSS development.

Setting and participants. The card sorting exercises were conducted in a meeting room, which had a large table for space and convenience. The participants were PhD



Figure 3.4. Participants sorting cards

students conducting their research in the IS field. PhD students were selected because they represent the main target audience for the DSS. The participants were selected by convenience.

Card sorting materials. The card sorting materials involved sets of 90 cards with words on them and a unique identifier number for recording purposes. Blank cards were also provided with pen to write if needed.

Card sorting procedure. The exercises were done in one-on-one sessions moderated by one of the authors (Figure 3.4). One-on-one sessions were adopted to promote personal opinions, to get detailed feedback, and also to cater for the diversity of research problems and methods that are typical in the IS field. The method was operationalised according to the following steps.

A brief explanation about card sorting and verbal instructions were given to participants at the beginning. The purpose of each exercise was explained before handing over a deck of cards to the participants. The participants were given some time to read through the cards for familiarization of contents before the exercises began. The first exercise (E1) contained 14 cards with theory testing steps, which the participants should group in five categories. Both the categories and steps were taken from the conceptual framework (Figure 3.2).

The second exercise (E2) showed five theory testing intents (Figure 3.3) and required the participants to, for each intent, delineate how theory testing activities should be

organised. In the third exercise (E3), we gave the participants a deck of cards with all activities and asked them to select which ones they thought could be CS. In the last exercise (E4), we gave the participants two decks of cards, one with potential problems faced during theory testing and another with possible features supported by the DSS. The participants had to organise the cards by order of importance. Blank cards were given to include any problems and features that were not in the decks.

After all exercises were completed, a brief interview was conducted to gather the participant's views regarding the exercise and how beneficial a DSS would be to help in making decisions on how to integrate CS in theory testing, and also to know what other features the participant might want the DSS to have that were not mentioned. The interviews were recorded and transcribed. All participants completed the entire procedure and each session took an average of 40 minutes, even though some exceeded 60 minutes.

3.5 RESULTS

3.5.1 Familiarity with theory testing (E1)

To recap, the participants were offered a deck of 14 activities and were requested to place them in the five categories defined by the conceptual framework (theory, conceptualisation, operationalisation, validation, and assessment) and an additional category named other. The “optimal” distribution of activities by categories, as defined by the conceptual framework, is shown at the top of Table 3.1. Below, we show the participants’ distributions. This approach allows measuring the deviations between the participants’ sorts and the conceptual framework. Category deviations were calculated by measuring the distance between the categories where an activity belongs and where it was placed by the participant. For instance, placing activity #1 in the theory category has a distance of 0, if placed in the conceptualisation has a distance of 1, if placed in the operationalisation category has a distance of 2, and so forth. These results were then averaged.

As shown at the bottom of Table 1, the best approximation was obtained for the assessment category, and the worst was obtained for the validation category. These results suggest the participants are not sufficiently familiarised with the logical

progression of theory testing activities, especially regarding the early stages. These results support the goal to develop the DSS as a way to improve knowledge about the structure of theory testing activities.

Table 3.1. Familiarity with Theory Testing

'Optimal' distribution (suggested by conceptual framework)						
Categories	Theory	Conceptual.	Operational.	Validation	Assessment	
Activities	1	2,3	4,5,6,7	8,9,10	11,12,13,14	
Participants' distributions						
Categories	Theory	Conceptual.	Operational.	Validation	Assessment	Other
P1	1	3,5	2,6,4,9	8,10,13,7	12,11,14	
P2	1	2,3,5		8,9,7,10,6,4,13	14,11,12	
P3	1,2	3,5,6,4,		8,9,7,10	11,13,12,14	
P4	1,14	2,5,3		7,9,8,10	11,13,12	4,6
P5	1	2,3	5,6,4	9,8,7,10,13	11,14,12	
P6	9,13	10,4,3	2,7,8	5,1,	12,6,11,14	
P7	1	3,2,4	6,5	8,9,7,13,10	14,11,12	
P8	1	2,12,4,3,6	5,7,8,9	13,10,	14,11	
P9	2,5	3,4,	1,12	8,7,9,13,10	6,14,11	
P10	1	2, 3	4,7	8,9,6,5,10	13,11,14,12	
Deviation	1.5	1.6	1	1.9	0.4	

3.5.2 Familiarity with model constructs (E2)

We adopted a contents analysis approach to evaluate the outcomes of this exercise. The ways in which the participants arranged the theory testing activities for each intent were then analysed and contrasted with the model shown in Figure 3.3. The results indicate that the participants had clear understanding of the different nature of the intents. However, we observed that the participants' arrangements were less diverse than the patterns suggested by the model. For instance, one participant characterised the case study intent as: site/case selection, interview, participants contacted after a period of time, and collection of posts from participants. This corresponds to one of eight possible patterns identified in the model. These results suggest the DSS may bring a more diverse view over theory testing.

3.5.3 Selection of activities that could be crowdsourced (E3)

We adopted frequency distribution to identify what activities the participants considered able to CS. The top selected activities are listed in Table 3.2.

Table 3.2. Top activities identified by the participants as able to CS

Collection of posts from participants	8
Questionnaires distributed for information collection	7
Online forum, transcription	5
Coding, collection of data over time, develop website, study for iPhone	4
Virtual environment created for experiment, site selection, survey link sent to participants	3

3.5.4 Problems and features (E4)

We adopted frequency distribution to analyse the problems and features reported by the participants. Figure 3.5 shows the obtained distribution of problems. Most of the identified problems are issues that CS can address, notably, recruiting participants and collecting data, which took the top of the list. With CS, one can recruit many subjects, and they can contribute to data collection. The subjects may even be willing to give out private information, because CS allows for anonymity.

We also adopted frequency distribution to analyse which DSS features were most well-regarded by the participants. As shown in Figure 3.6, the highest priority was given to the capability to show different patterns to theory testing and the different pathways that can be adopted by a research project.

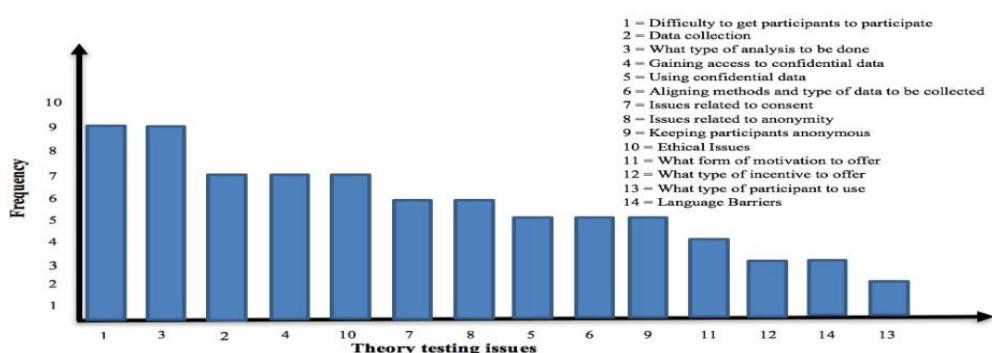


Figure 3.5. Frequency distribution of problems with theory testing identified by participants

3.5.5 Feedback from the interviews

The participants responded positively to the card sorting exercises. The first exercise was an eye opener to some as they were not so familiar with the theory testing steps. Most of the participants said they have been overly exposed to theory building concepts, both in the literature they have read and the classes that have taken during their Ph.D. studies. However, theory testing has not been that prominent, even though they realised that it is important. In that context, the participants perceived that a tool illustrating how different theory testing activities could be related to an intent would be useful.

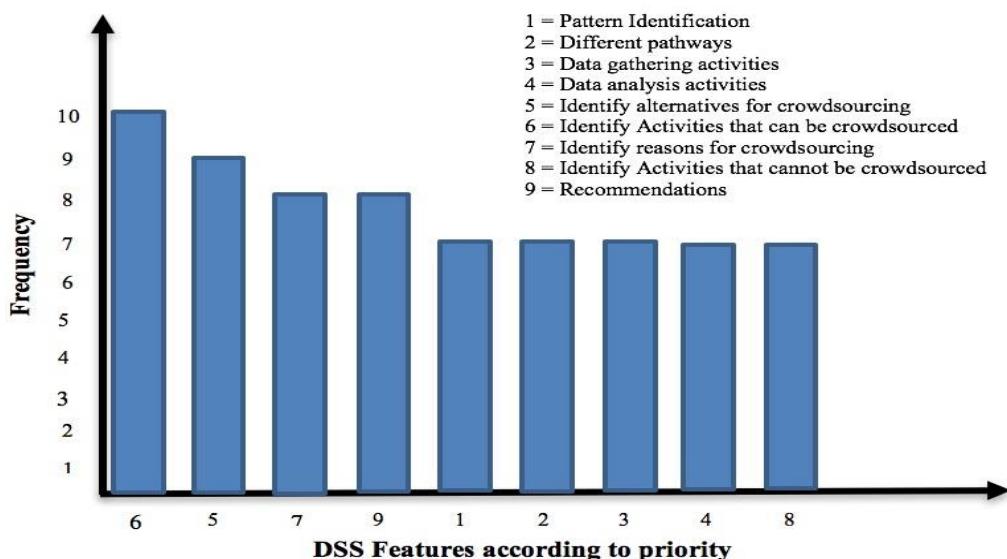


Figure 3.6. Frequency distribution of DSS features prioritized by the participants

A lot of positive feedback was received from the participants about the card sorting exercises. They particularly liked the way of prioritising features they would like to see in the DSS. The card sorting exercises were also found to be effective to compare their thoughts with the conceptual framework.

3.6 DISCUSSION AND CONCLUSIONS

To the best of our knowledge, no DSS is currently available to help Ph.D. students making decisions about their theory testing endeavours, especially considering how

to integrate the CS strategy in the process. This paper describes first steps in that direction and therefore represents a true innovation.

This paper also describes and justifies with significant detail the steps leading to the DSS. Our method uses the design science paradigm and Schön's (1983) problem/solution framing. We believe the two perspectives are very synergetic. On the one hand, design science emphasises the combination of rigour and relevance in the development of IS artefacts. On the other hand, Schön's viewpoint integrates design decisions into IS artefacts. Bringing them together, we completely scaffold the development of the conceptual framework and model. The framework is very important, as it provides a solid frame for the DSS: the framework is based on relevant literature about theory building, which integrates aspects of theory testing. The model not only builds upon the conceptual framework, but it also integrates elements from a comprehensive scoping literature review. Seen together, the two artefacts are logically related and solidly established in the related literature.

Our integration of the DSR build/evaluate cycle and Schön's problem/solution framing, has also been extended to evaluation, through problem frame assessment. Instead of evaluating the IS artefact against its users, we evaluated our problem frame against the users. This approach seems very adequate to evaluate secondary artefacts, as it provides justificative knowledge necessary to later on develop the DSS. This approach does not preclude adopting a more summative approach to evaluate the DSS. Instead, it contributes to support the DSS development on a realistic foundation, which takes both users and designers into consideration.

The problem frame assessment revealed that, even though PhD students were knowledgeable about theory building, they were not familiar with theory testing. The PFA also showed that PhD students were not aware of the variety of patterns that can be used to test theory. They also seemed to be unaware of the diversity of theory testing activities that can be CS. All in all, these results suggest the conceptual framework and model constitute a solid foundation for developing the DSS in the future.

Regarding the present, we note this research already brings some interesting contributions. One is highlighting that PhD students may be currently underexposed

to theory testing, when compared to theory building. Another important contribution is the identification of a variety of patterns that can be used for theory testing, along with the identification of a set of activities that can be CS. Both the conceptual framework and model developed by this research may contribute to increase the attention to theory testing as a phenomenon of interest. The integration of theory testing with CS may increase attention to this strategy in research design.

3.7 REFERENCES

- Aken, J. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of management studies*, 41(2), 219-246.
- Argyris, C. (1996). Actionable knowledge: Design causality in the service of consequential theory. *The Journal of Applied Behavioral Science*, 32(4), 390-406.
- Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of information technology*, 20(2), 67-87.
- Arnott, D., & Pervan, G. (2012). Design science in decision support systems research: An assessment using the Hevner, March, Park, and Ram Guidelines. *Journal of the Association for Information Systems*, 13(11), 923.
- Barrett, A., & Edwards, J. (1995). Knowledge elicitation and knowledge representation in a large domain with multiple experts. *Expert Systems with Applications*, 8(1), 169-176.
- Bitektine, A. (2007). Prospective case study design: Qualitative method for deductive theory testing. *Organizational Research Methods*, 11(1), 160-180.
- Colquitt, J., & Zapata-Phelan, C. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal*, 50(6), 1281-1303.
- De Vaus, D. (2013). *Surveys in social research*: Routledge.
- Gerring, J. (2006). *Case study research: Principles and practices*: Cambridge University Press.
- Gregor, S., & Hevner, A. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 37(2), 337-355.
- Hevner, A., & Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol. 22): Springer Science & Business Media.
- Holmström, J., Ketokivi, M., & Hameri, A. (2009). Bridging practice and theory: A design science approach. *Decision Sciences*, 40(1), 65-87.
- Hosack, B., Hall, D., Paradice, D., & Courtney, J. (2012). A look toward the future: decision support systems research is alive and well. *Journal of the Association for Information Systems*, 13(5), 315.
- Howe, J. (2006). The rise of crowdsourcing. *Wired magazine*, 14(6), 1-4.
- Järvinen, P. (2007). Action research is similar to design science. *Quality & Quantity*, 41(1), 37-54.

- Kitchenham, B., Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1), 7-15.
- Lowry, P. B., D'Arcy, J., Hammer, B., & Moody, G. D. (2016). "Cargo Cult" science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels. *The Journal of Strategic Information Systems*, 25(3), 232-240.
- MacKenzie, S., Podsakoff, P., & Podsakoff, N. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, 35(2), 293-334.
- March, S., & Smith, G. (1995). Design and natural science research on information technology. *Decision support systems*, 15(4), 251-266.
- Niederman, F., & March, S. (2015). Reflections on Replications. *AIS Transactions on Replication*, 1.
- Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. *European Journal of Information Systems*, 25(6), 1-16.
- Paré, G., Trudel, M., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183-199.
- Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.
- Penker, M. (2000). Business Modeling with UML: Business Patterns at Work: New York, NY: John Wiley & Sons.
- Schön, D. (1992). Designing as reflective conversation with the materials of a design situation. *Knowledge-based systems*, 5(1), 3-14.
- Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS quarterly*, 35(1), 37-56.
- Sonnenberg, C., & vom Brocke, J. (2011). *Evaluation patterns for design science research artefacts*. Paper presented at the European Design Science Symposium.
- Spencer, D. (2009). *Card sorting: Designing usable categories*: Rosenfeld Media.
- Spencer, D., & Warfel, T. (2004). Card sorting: a definitive guide. *Boxes and arrows*. Retrieved from
- Steelman, Z., Hammer, B., & Limayem, M. (2014). Data Collection in the Digital Age: Innovative Alternatives to Student Samples. *Mis Quarterly*, 38(2), 355-378.
- Thuan, N., Antunes, P., Johnstone, D., & Duy, N. (2015). *Establishing a Decision Tool for Business Process Crowdsourcing*. Paper presented at the International Conference on Future Data and Security Engineering.
- Thuan, N. H., Antunes, P., & Johnstone, D. (2016). *A Design Science Method for Emerging Decision Support Environments*. Paper presented at the Australasian Conference on Information Systems, Adelaide, Australia.
- Von Alan, R., Hevner, A., March, S., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1), 75-105.

CHAPTER 4

PATTERNS OF TESTING THEORY WITH HUMAN SUBJECTS: A DESIGN SCIENCE PERSPECTIVE

ABSTRACT

This paper addresses the challenges of theory testing. Adopting the design science paradigm, we address this challenge by delimiting the line that separates theory building from theory testing by conceptually characterizing its essential aspects: the relationships between humans, organizations and technology, which led to the development of a conceptual framework for theory testing. Practically, the methods, processes, instruments and tools needed to acquire and analyze data for theory testing are considered, leading to the development of a pattern system. Card sorting was used as an evaluation strategy for the conceptual framework and the pattern model.

Keywords: Theory, Theory testing, Design Science, Pattern Theory.

4.1 INTRODUCTION

The key foundations to advance scientific knowledge include the testing and re-testing of theory (Niederman & March, 2015). Theory testing is concerned with establishing or refuting the validity of a theory's core propositions by determining to what degree they provide a close fit to empirical data (Colquitt & Zapata-Phelan, 2007). Theory testing thus requires collecting empirical data either supporting or not-supporting the set of propositions articulated by a theory (Lokke & Sorensen, 2014). Empirical data can be collected from various types of sources and using various types of instruments, exemplified by variations among disciplines.

In this paper we address the challenges of testing theory in the IS discipline. From a conceptual viewpoint, we have to consider the diversity of variables that characterise the essential aspects of the phenomena of interest covered by the IS field: the relationships between humans, organisations and technology. Identifying and

characterising these variables is critical to both rigorously testing existing IS theory and exploring new IS theory.

From a more practical viewpoint, we have to consider the methods, processes, instruments, and tools that have to be used to acquire data necessary for theory testing (Bhattacherjee, 2012). The confluence between humans, organisations and technology generates varied and rich types of data, which the researcher has to analyse in different ways to determine whether the propositions articulated by a theory are supported or not by the data.

Our main goal is to help researchers understand the complexities of theory testing and act upon them with sound, robust theory testing processes. To accomplish this, we set two objectives. The first objective is to define a conceptual framework for theory testing in the IS field. This framework lays a foundation for the second objective: to define a model based on theory testing patterns, which codify the structural elements of theory testing and help researchers link these structural elements to their specific research contexts. Another way of looking at these two objectives is that the framework allows researchers to conceptualise theory testing, and patterns allow researchers to apply theory testing.

To accomplish these objectives, we apply design science research (DSR). DSR is a research paradigm that emphasises the development of innovative and useful IS artefacts (Hevner & Chatterjee, 2010). In this case, we regard the conceptual framework and pattern model as IS artefacts. This paper contributes to IS research in three ways. First, we propose two artefacts supporting theory testing in the IS field. More broadly, this research contributes to theory generation by supporting theory validation.

The remainder of this paper is structured as follows. In the next section we discuss the research approach in more detail. In Section 3 we develop our first artefact: a conceptual framework of theory testing. Section 4 uses the framework to review a body of knowledge on how theory has been tested in practice and synthesises that knowledge into a pattern model of theory testing. In Section 5 we evaluate the developed artefacts to demonstrate their significance and utility. Finally, in Section 6 we discuss the research contributions and provide some concluding remarks.

4.2 RESEARCH APPROACH

We adopt the DSR paradigm for this study (Hevner et al., 2004). DSR supports problem-solving research centred on the creation of innovative artefacts. By artefact we mean anything that can be developed as an artificially made object (Gregor & Hevner, 2013). In that category we include abstractions such as the framework and model proposed in this study (Gregor & Hevner, 2013).

The DSR paradigm articulates three research cycles: relevance, rigour, and development & evaluation. The relevance cycle ensures the problem concerns a relevant application environment; and that the artefact solution delivers utility in that environment (Hevner & Chatterjee, 2010). In our case, the application environment concerns researchers wanting to test theory; and relevance concerns improving research processes involving theory testing.

The rigour cycle ensures the research is grounded on a solid knowledge base. This is done in two directions: first, applying relevant knowledge to understand the problem and construct a solution artefact; and second, contributing to the knowledge base with new knowledge in the form of innovative artefacts and principles of artefact implementation (Gregor & Jones, 2007). We ground the artefact construction on two literature reviews. The first review concerns the problem domain: it seeks to develop a clear understanding of how theory is generated. The second review addresses the solution domain by considering how theory has been tested in studies published in top IS journals (the “basket of eight”).

The artefact development and evaluation cycle is an iterative process where an artefact is developed based on a problem frame and then evaluated against the application environment, leading to the adjustment of the problem frame and artefact until a satisficing solution is obtained (Sein et al., 2011). By combining this generic structure with the specific application environment of our research, we can then delineate a more detailed research method. This approach to artefact development is described in Figure 4.1.

In this study we develop two interrelated artefacts, a conceptual framework and a pattern model. The former defines a problem frame while the latter corresponds to the

artefact solution. The conceptual framework positions theory testing in the wider domain of theory generation and highlights which activities pertain to each domain. Its innovative aspect lies in a clear separation between theory building and theory testing.

The pattern model of theory testing is a synthesised output of a descriptive literature review on how theory has been tested. Our interest in building this model is not so much on articulating the conceptual parts of theory testing, but more on articulating the purposeful activities involved in theory testing. The model adopts pattern theory (Alexander, 1999), which originated in architecture and has been used in software engineering to describe generalised solutions to recurring design problems (Zigurs & Khazanchi, 2008). In the DSR context, existing theory used in artefact development has been designated kernel theory (Gregor & Jones, 2007). The adoption of kernel theory in artefact development is considered important to strengthen the relationship between artefact and knowledge base (Gregor & Jones, 2007).

The two artefacts are related through what we could designate as a cycle of mutual justification. On the one hand, the pattern model is justified by the conceptual framework, which delimits the domain of applicability and provides the fundamental concepts and relationships. On the other hand, the conceptual framework is justified through the realisation of the pattern model.

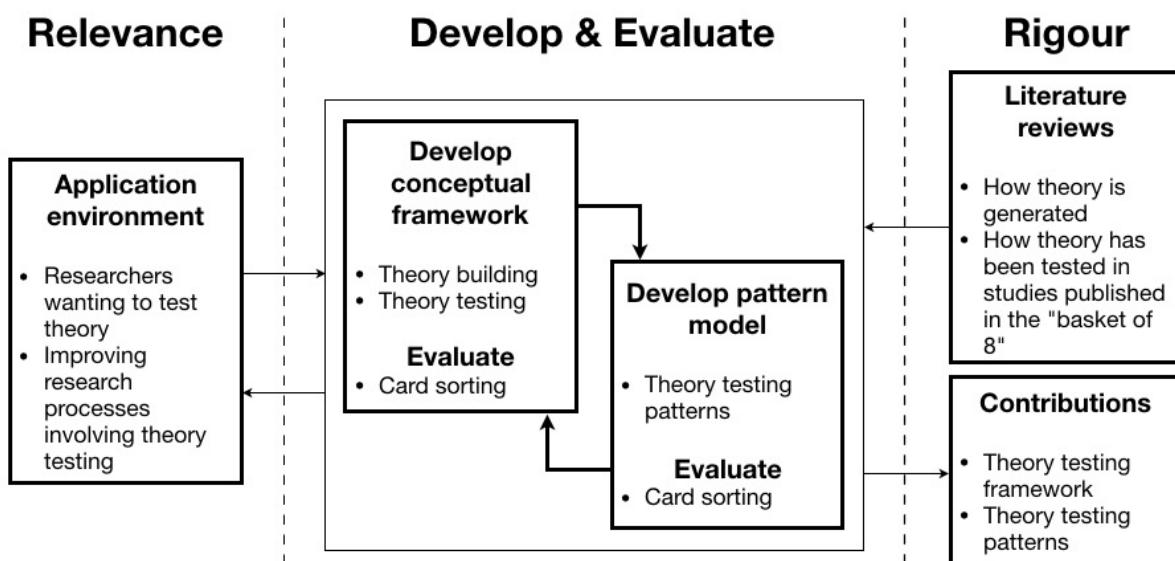


Figure 4.1. Research approach

In DSR, artefact evaluation seeks to deliver two complementary types of contributions (Sonnenberg & vom Brocke, 2011). One is utility, with a focus on solutions to problems relevant in an application domain (Hevner et al., 2004). The other is justificatory knowledge, providing explanations derived from artefact construction which improve general knowledge about artefact development (Gregor & Jones, 2007). In particular, justificatory knowledge contributes to design theory by adding technological rules, principles of implementation, and prescriptions to the knowledge base (Bichler et al., 2016; Gregor , 2006; Gregor & Jones, 2007). In this study, we adopt this dualistic view on the evaluation of artefacts. We apply the card sorting technique to evaluate both the conceptual framework and pattern model.

Card sorting involves participants sorting, grouping and prioritising sets of cards with pieces of information (e.g. concepts). The sorting process externalises the participant's thinking and allows them to analyse how they would structure information (categories, priorities, relationships, groups, etc.). In this study, we use card sorting to analyse how researchers relate their own expertise to the conceptual framework, which indicates the closeness between problem frames: as perceived by researchers, and as articulated in the conceptual framework. We also use card sorting to analyse how researchers relate their goals to the pattern model. This provides an indication of how researchers value the model.

Next, we discuss the development/evaluation of the first artefact in more detail.

4.3 CONCEPTUAL FRAMEWORK

Theory helps scientists logically describe and explain a phenomenon, process or sequence of events. (Gregor, 2002, 2006) suggests a classification into five theory types summarized in table 4.1. A Type I theory is descriptive, answering the “what is” question by seeking to characterise elements of the phenomena (e.g. through classification). A Type II theory is explanatory, answering the “how and why” question in relation to the phenomena of interest. A Type III theory is predictive, answering the question “what will be” without necessarily understanding how or why the phenomena occurs (for example by employing statistical methods). A Type IV theory is both explanatory and predictive, answering the questions “how”, “why” and “what will be”.

Finally, a Type V theory is prescriptive, answering the question “how is it done”, by focusing on the methodologies, methods, principles, and guidelines that shape the construction of artificial artefacts (Gregor , 2002).

Table 4.1. Gregor's classification of theory (Gregor, 2006)

Theory Type		Description
Type I	Analysing & Describing	Describes and classifies “what is”, usually using a set of dimensions associated to a phenomenon of interest.
Type II	Understanding	Explains “how” and “why” something occurs.
Type III	Predicting	Aims at predicting “what will be”. Provides testable propositions.
Type IV	Explaining & Predicting	Concerns “what is”, “how”, “why”, and “what will be”. Provides testable propositions with causal explanations.
Type V	Design and Action	Concerns “how something is done”. This includes methodologies, methods and tools.

In this paper we focus on Type III and Type IV theories because they both involve predictive elements focused on testable propositions. Goode & Hatt (1952) define a testable proposition as an imaginative idea, a guess, and a statement of a solvable problem or any thinking that can be put to test to determine its validity. A sound testable proposition should 1) refer to observable and measurable events, 2) provide a solution to a defined problem, 3) be formulated in a way that can be tested, 4) be related to the existing body of knowledge, and 5) have logical unity (Khan, 2011). A testable proposition can take the form of an hypothesis: a tentative generalization which has to be tested to determine its validity (Khan, 2011).

The test of a theory is concerned with establishing or refuting the validity of the theory's testable propositions. A theory is tested by determining to what degree it provides a close fit to empirical data (Colquitt & Zapata-Phelan, 2007). These outcomes are the basis for codifying and accumulating scientific knowledge, as suggested by Popper's falsifiability requirement (Popper, 2014). The continuous theory building and testing and re-testing is considered a foundation towards establishing scientific knowledge (Niederman & March, 2015).

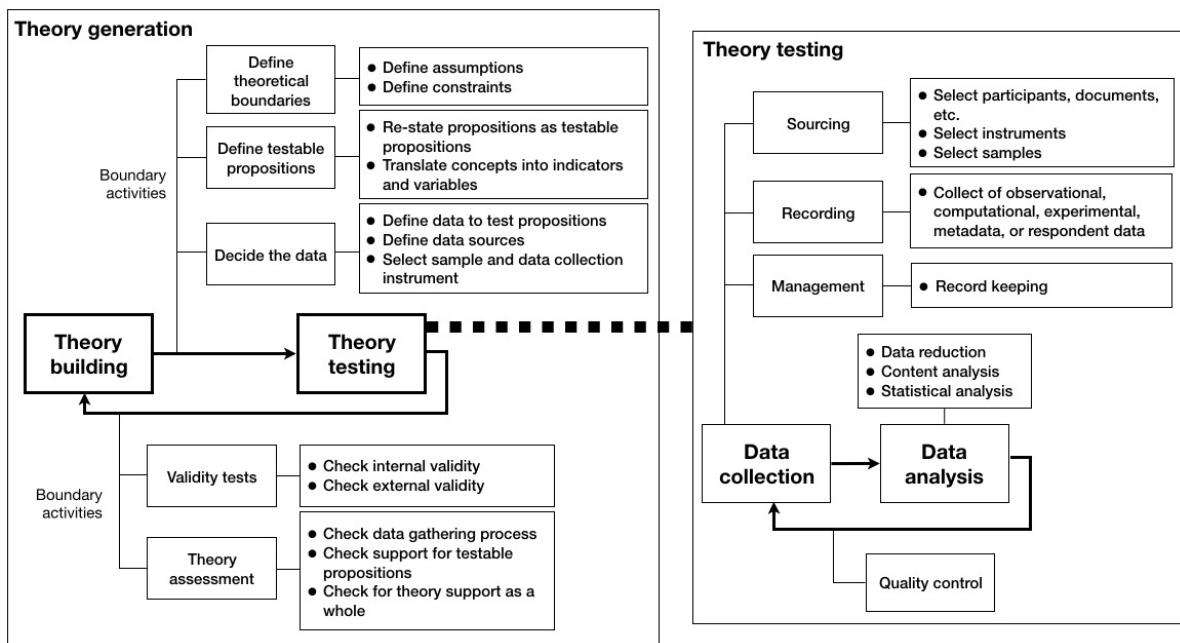


Figure 4.2. Conceptual framework of theory testing

Researchers spend time building theory; and then spend more time testing theory (Viswesvaran & Ones, 1995). De Vaus (2013) and Bitektine (2007) developed conceptual frameworks that describe the major activities involved in theory testing. They include important concepts such as the definition and operationalisation of theoretical constructs, data selection, definition of hypotheses, and formulation of criteria for outcome evaluation. However, a problem with these frameworks is that they combine theory building with testing. Since we are exclusively focussed on theory testing, we need to disentangle the two concepts.

To accomplish that goal, we first conceptualise the theory generation cycle (Figure 4.2), where theory is consecutively built, checked against empirical data and adjusted to better reflect the real world (Lynham, 2002). Theory adjustments may reflect changes to the way the theory is expressed and/or the context where it applies. This, in turn, leads to further testing. Putting theory in use through testing creates informal knowledge that allows the understanding, explaining and predicting of what happens in the world (Lynham, 2002).

This simple, mutually re-enforcing, relationship between building and testing allows us to determine the activities involved in each when generating theory. However, we note that these activities will vary depending on epistemological assumptions and

methodological choices, and it will not always be clear how the outputs of building activities will readily translate into inputs for testing activities. To address this concern, we consider the notion of a boundary activity. According to Maaninen-Olsson et al. (2008), a boundary activity exists with the purpose of integrating knowledge from two different contexts. The concept was developed in the knowledge management field to understand how a community-of-practice codifies and transfers knowledge between different contexts. Here, we use the concept to discuss how certain boundary activities are necessary to integrate theory building and theory testing.

Regarding the theory generation cycle in the direction that goes from theory building to testing, we can identify three important boundary activities: define theoretical boundaries, define testable propositions, and decide upon the test data. The first boundary activity concerns setting the boundaries of the theory, identifying the value-laden assumptions taken by the researcher and the constraints to its applicability (Bacharach, 1989).

The second boundary activity concerns the transformation of theoretical propositions into testable propositions. A theoretical proposition is a conjecture about the world, which is claimed by the researcher (Wacker, 1998). A testable proposition is a theoretical proposition that is stated in a way that can be tested to determine its validity (Goode & Hatt, 1952). Testable propositions make explicit which indicators and (qualitative or quantitative) variables can be used for empirically testing the proposed conjecture about the world. Therefore, restating conceptual propositions as testable propositions involves translating abstract concepts into observable variables, a step which is often called operationalization (De Vaus, 2013). The more conceptual propositions to be tested, the stronger the test of a theory (Stinchcombe, 1987).

The third boundary activity is to decide upon the data, which includes: defining the type of empirical data necessary to test the propositions, identifying the test data sources, and establishing appropriate samples and collection instruments.

The direction that goes from theory testing back to theory building involves two boundary activities. One focuses on activities establishing the validity of the testing regime, taking two forms: internal and external. Internal validity checks the strength of the testing process, seeking to confirm the testing activities will appropriately test the

propositions. External validity checks the generalisability of the testing outcomes across different contexts.

The other boundary activity to consider is theory assessment, in particular checking the support given by empirical data to each testable proposition, and checking the support given by empirical data to the theory as a whole, taking the set of propositions together (De Vaus, 2013).

We now move our attention to theory testing, increasing the level of detail. As a result of the post-theory building boundary activities, we have a set of testable propositions, statements of any theoretical assumptions and constraints that must be met, and an outline of the data to be sampled, including where it can be found, and how it will be sampled. Collectively, these form the basis of a testing plan that will specify the testing activities to follow.

Testing can be conceptualised as a cycle comprising two activities: data collection and data analysis. Data collection is an umbrella term for a variety of secondary activities. In particular, we consider: data sourcing from participants, documents and other sources; data recording, which may include observational (e.g. counting events), computational (e.g. online processing), experimental (e.g. outputs generated from assigned tasks), metadata (collecting data about existing data), and respondent data (surveys) (Borgman et al., 2007); and data collection management, which includes record keeping activities.

Data analysis includes three important secondary activities: data reduction (e.g. coding), content analysis and statistical analysis.

Finally, to complete the cycle, we consider one boundary activity between data analysis and data collection, where the outcomes of data analysis inform adjustments to the data collection plan based on a data quality control function, which may serve to improve overall internal and/or external validity.

The evaluation of this artefact was done using card sorting (see Section 4.5).

4.4 PATTERN MODEL

We now address our second research goal: to define a model with theory testing patterns, which codifies the structural elements of theory testing and helps researchers to link these structural elements to their specific research contexts. To accomplish this goal, we systematically review the literature on how theory has been tested in the IS discipline. The adoption of a systematic¹ approach stems from the DSR principle that, when designing an artefact, the researcher should rigorously ground the whole process in the existing knowledge base (Hevner et al., 2004). Furthermore, the adoption of well-defined and well-structured protocols contributes to research transparency (Kitchenham et al., 2009; Paré et al., 2016).

4.4.1 Review Procedure

We employ a descriptive literature review (Paré et al., 2015) to examine the different ways in which theory has been tested in the IS field. This method helps to determine the extent to which a body of empirical studies supports or reveals any interpretable trends or patterns with respect to pre-existing findings (Paré et al., 2015). This is done by collecting, coding and analysing data that reflects the frequency of topics found in the literature.

The review procedure includes the following stages (Kitchenham et al., 2009; Okoli & Schabram, 2010): 1) select articles; 2) filter articles; 3) data extraction; and 4) data synthesis. Details about these steps are presented below.

Select articles

We are interested in articles that test theory. A two-year period of journals from the “Basket of Eight” is reviewed. These journals were selected because they are recognised as publishing some of the best research in the IS field. By covering this set, we accommodate a range of research traditions in IS, while targeting the more obvious outlets for theory testing.

¹ Note that being ‘systematic’ while conducting a literature review is not the same as conducting a ‘systematic literature review’ (Pare et al. 2016).

We use a 2-year period for the review because this frame seems sufficient to develop a comprehensive view over the topic. Furthermore, during the review we found that we got close to a saturation point, where not much new data is collected (Fusch & Ness, 2015). The initial selection of articles based only on publication outlets resulted in 508 items.

The final selection of articles is obtained through keyword search, a procedure that other published reviews have also used (Estelles-Arolas & Gonzalez-Ladron-de-Guevara, 2012). We consider articles containing any one of the following keywords: ‘theory’, ‘test’, ‘theoretical model’, ‘theoretical framework’, ‘conceptual model’, and ‘conceptual framework’. This search is done on the full body of text of each article. The final selection contained 506 papers, which means that only two papers were excluded. This should not be a surprise considering that the selected publication

Table 4.2. Summary of articles after selection and filtering stages

	Journal	No. of articles	Excluded based on theory type	Selected
MISQ	2016	29	7	22
	2017	51	15	36
JSIS	2016	19	6	13
	2017	15	5	10
JMIS	2016	42	10	32
	2017	42	6	36
JIS	2016	25	8	17
	2017	21	2	19
JAIS	2016	26	4	22
	2017	32	3	29
ISR	2016	49	11	38
	2017	46	13	33
ISJ	2016	24	5	19
	2017	23	1	22
EJIS	2016	32	4	28
	2017	32	4	28
TOTAL		508	104	404

outlets have a strong theoretical nature, and the selected keywords considered typical theoretical constructs.

Filter articles

The purpose of this stage is to eliminate articles unrelated to the study. As mentioned earlier, we are interested in articles where theory is tested one way or the other, which means that theories should belong to type III and IV categories. The filtering procedure is as follows. An article is taken from the pool and the methodology section is checked to determine if it contains a theory with testable propositions (Khan, 2011; Muntermann et al., 2015). After this stage, we had a collection of 404 articles to review. In Table 4.2 we summarise the results from the selection and filtering stages.

Data extraction

This stage involves a detailed reading of specific sections of each article, and then coding relevant data for posterior analysis. According to Paré et al. (2016), there are three main operational aspects to data extraction. The first one is deciding what parts of an article should be targeted. In our review, we focus on the methodology, discussion and concluding sections, as they are critical to assessing theoretical contributions, where theory testing typically takes a central role in supporting the core propositions.

Coding of research activities is the second operational aspect to consider. Here we focus more on the theory testing process than on conceptual content. We adopt a combination of open and closed codes. We use a data extraction form with both predefined data fields and open fields that allow new codes to be added if necessary. This is used to balance efficiency with openness, ensuring that the review is not confined to a small set of questions and assumptions (see figure 4.3).

The final operational aspect is managing the extracted data in preparation for analysis. We used NVivo to manage the extracted data. Coding was done by the first author. To increase the reliability of the analysis (Krefting, 1991), an independent researcher coded a sample selection of articles in parallel and the codes were compared and adjusted for clearer understanding.

Data synthesis

After data extraction, we have a collection of raw data that has to be synthesised into a usable artefact. To synthesise the review data, we adopt pattern theory, which is discussed below.

4.4.2 Pattern Theory

According to pattern theory, a pattern is a regular and intelligible form or sequence in which something is done (Alexander, 1999). It can also be seen as a generalized solution to a recurrent problem, which is presented in a template form (Kampffmeyer & Zschaler, 2007). Patterns have a strong affinity with DSR since researchers, when developing artefacts, do not seek to develop single solutions to unique problems, but instead seek to develop generalised solutions to classes of problems (Bider et al., 2013).

Pattern theory has been proposed in the field of architecture with the same purpose (Alexander, 1999). Pattern theory concerns the design domain, i.e. “how to” find a satisficing solution to a problem that involves material construction by humans, and therefore may have an infinite number of solutions (Simon, 1996). Professional fields such as architecture, engineering and computer science have to make constant design

Coding								
General Information		Data Collection			Data Analysis			Other
Source	Title	Method of Data collection	Recruitment Method	Data Quality Check Method	Data Reduction Method	Content Analysis Method	Statistical Analysis Method	Crowdsourcing
MISQ 2016 starts	A TREE-BASED APPROACH FOR ADDRESSING SELF-SELECTION IN IMPACT STUDIES WITH BIG DATA	Experiment * Simulation * Observation * survey	Eligibility Criteria			Performance analysis propensity score subclassification (PSS) propensity score matching (PSM)		
	AN EXPLORATORY STUDY OF THE FORMATION AND IMPACT OF ELECTRONIC SERVICE FAILURES	Survey	Invitations e-mailed to members	Used the filtering question inserted at the start of the questionnaire	Removal of ambiguous responses, coding	Content analysis (sorting, chi-square test and Crisp-Set Qualitative Comparative Analysis c3QA)	Descriptive Statistics, Post-Hoc analysis	
	ARE SOCIAL MEDIA EMANCIPATORY OR HEGEMONIC? SOCIETAL EFFECTS OF MASS MEDIA DIGITALIZATION IN THE CASE OF THE SOPA DISCOURSE	Case study * tweets	Records from social media sites	manipulation control on task characteristics through contextual framing of survey questions	Use of keywords, description, categories, coding	Transcription, Interpretive analysis		
	CAPTURING THE COMPLEXITY OF MALLEABLE IT USE: ADAPTIVE STRUCTURATION THEORY FOR INDIVIDUALS	Survey * questionnaire	Smart phone users		Removal of responses with Missing data		PLS, Post-Hoc analysis	

Figure 4.3. Fragment of the data extraction form used in the review

decisions and therefore may benefit from pattern theory. They often identify recurrent categories of problems for which template solutions have to be initially designed and then repeatedly used (Zigurs & Khazanchi, 2008). In the professional context, a pattern expresses the abstract relationship between a category of problems and a template solution.

Key elements associated with patterns are: *intent*, *structure* and *implementation* (Penker, 2000). *Intent* defines the general purpose of a pattern and is also used to name the pattern. *Structure* regards the organisation of activities within a pattern, which define a coherent system of knowing what to do (Alexander, 1999). And *implementation* considers how the pattern realises a solution to the problem.

Pattern structure involves the arrangement and relations between the pattern activities, links and pathways of the pattern model. The structure includes two main activities – the primary activities and the secondary activities. A primary activity represents the intent of the pattern, giving meaning and purpose to other activities, and usually appears at the start of the pattern. The primary activity generates data and gives rise to other activities. Secondary activities are non-data generating activities but serve as a support. They are activities that need to be carried out to achieve the desired intent.

Implementation considers how the solution of a problem is realized based on the pattern. It also involves rules and procedures regarding how the structure works or is implemented. A rule has: an intent and a tangible purpose; associated secondary activities; and the possibility of linking secondary activities to the primary activity in a coherent format. Procedures will: first seek to identify and select primary activities/intents; then identify relevant secondary activities associated with the identified intent; then identify different pathways (if any) that can be associated with the selected intent; and finally link the activities accordingly based on the proposed intent.

Applying the Rules and Procedure to an Example Case

This example uses an article written by Burton-Jones and Volkoff (2017), which was drawn from the pool of reviewed papers. The study adopted an inductive approach to theory building and the case study approach to theory testing (as noted in the paper's introduction). As discussed in more detail in the paper's Section 3, theory testing comprised the demonstration of the proposed theoretical constructs in a specific case dealing with a community-based healthcare system.

The authors obtained their data through focus groups, interviews, observations of system use by shadowing workers, and document analysis (Section 3.3). Matching these specific theory testing activities to the pattern system shown in Figure 4, we can infer which of the patterns best suit what was done. We first note that a particular environment was selected (a community care unit of a regional health authority in Canada). The paper does not provide details about other activities done between selecting the environment and selecting the participants. The selection of participants was very diverse, since data collection was done over a period of four years. The selection of data sources considered a wide range of criteria including experience with the system (new users versus long-term users), roles (e.g. nurses, physicians and social workers), specialities (community care and information technology staff), and organisational levels. Participants in the initial focus groups were selected according to job category to ensure that participants could speak without their supervisors being present. The selection of participants for interviews followed-up participation in focus groups. After the participant selection, the authors report they got group and individual perceptions. Furthermore, the paper has explicit references to moderating the focus groups.

Regarding document analysis, the paper provides a brief account of the types of artefacts that were selected for data extraction (e.g. business cases, terms of reference and project descriptions). Finally, the paper provides evidence that specific outcomes of interest and environments were selected for observing workers (shadowing). We can therefore identify the following data collection patterns that have been used:

Intent (Primary activities)

There were multiple intents identified from the paper

- Get group perception
- Get individual perceptions
- Data extractions from artefacts
- Observation

Procedure

Applying the above-mentioned procedure, we have the following outcomes:

- Select environment - [No other activity reported] - Select participants - Moderate discussion - Get group perceptions
- Select environment - [No other activity reported] - Select participants - Get individual perceptions
- Select records - Data extraction from artefacts
- Select outcome of interest - Select environment - Observation

Burton-Jones and Volkoff (2017) analysed data by transcribing the outputs from focus groups, interviews and shadowing, coding all data, and finally performing content analysis. We can therefore identify the two data analysis patterns that have been used:

- Transcribe - Code - Content analysis
- Code - Content analysis

Based on these rules and procedures, a model representing a pattern system with 32 theory testing activities, and the relationships between them, was developed. The model is presented in Figure 4.4 and discussed below.

4.4.3 Pattern System

The analysed body of research exhibits a clear separation of concerns between data collection and data analysis, which supports the theory testing cycle defined in Section 4.3 and illustrated in Figure 4.2. We start the literature synthesis by discussing the

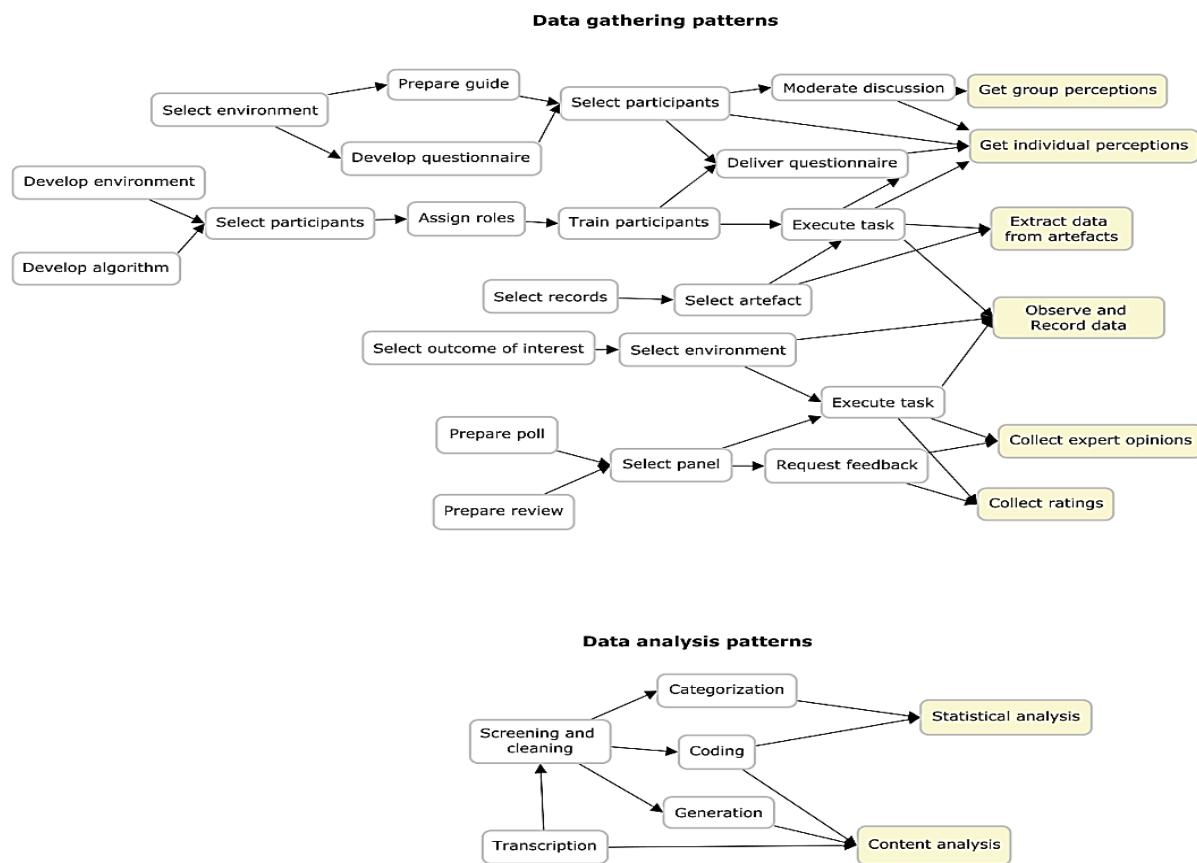


Figure 4.4. Synthesis of the literature review showing theory testing patterns. Patterns are divided between data collection and data analysis. Primary activities are presented with dark background, while secondary activities have a clear background.

data collection patterns. We found two different types of data collection activities reported in the literature. The first type concerns primary activities, which generate data, while the other type concerns secondary activities, which support data gathering activities but do not generate data by themselves. We found six primary activities: “get group perceptions”, “get individual perceptions”, “extract data from artefacts”, “observe and record data”, “collect ratings”, and “collect expert opinions”. We found that each one of these primary activities has a very specific purpose and therefore should correspond to a distinct intent. Consequently, we also have six different patterns describing how to conduct data collection in research.

To build the patterns, we started from the primary activities and then analysed their relationships with secondary activities by moving backwards. In establishing these relationships, we kept the requirement that all secondary activities linked to a primary activity should contribute to the same intent.

To illustrate the procedure, having identified the “get group perceptions” primary activity and intent, we observed that the preceding secondary activity was “moderate discussion”, which does not generate data by itself but is nevertheless necessary to have a group of participants generating data. We then found that the activity preceding “moderate discussion” was “select participants”, which concerns identifying and selecting people to participate in the research. And finally, we found that, preceding “select participants”, there were two possible activities, either “prepare guide” or “develop questionnaire”. The former activity concerns preparing the topic, set of initial questions, and group rules, while the latter activity concerns the design of a questionnaire. Since the “get group perceptions” pattern can be structured according to different secondary activities, and the primary activity in itself is agnostic regarding the research paradigm and method, we observe the pattern can be used to implement different research methods, e.g. focus groups and semi-structured group interviews – provided they all pursue the same intent.

The six data gathering patterns that were identified in this way can start with different activities and often share activities between them. Therefore, the best way to analyse these patterns is to start from the end and move backwards.

Regarding data analysis, we found two different types of activities. Once again, one type can be regarded as primary, because it addresses an intent, while the other concerns secondary activities, which support the primary ones.

We found in our data a myriad of approaches to data analysis that could be classified as primary activities, e.g. cluster analysis, sociomaterial analysis, interpretive analysis, correlation analysis, econometric analysis, analysis of variance, and so forth. However, we consider it would be very difficult to define distinct intents for each one of these. Therefore, we decided to condense these different activities into two major intents: “statistical analysis” and “content analysis”. These two intents reflect the well-known and heavily discussed distinctions between qualitative and quantitative analyses (Kumar, 2005). Regarding secondary activities, we also found a wide range of possibilities in the literature, which we condensed into five activities: transcription, screening and cleaning, categorization, coding, and generation. Screening and cleaning comprehend more specific activities such as filtering noise (e.g. discarding

duplicates), applying selection criteria to exclude or include data (e.g. excluding non-homogeneous transactions (Lu et al., 2016), removing information that is not in English (Zhang et al., 2016)), and excluding data based on screening questions (Roberts et al., 2016). We note that coding can be applied to both qualitative and quantitative data ((Selander & Jarvenpaa, 2016) and (Li et al., 2017) are two examples), and can use various techniques, such as open and axial coding (Miles & Huberman, 1994). Finally, generation considers the generation of data from data (e.g. generating a network model from a dataset of online comments (Zhang et al., 2016b)).

We finally note that the reviewed literature provides ample evidence that the relationships between data collection and data analysis are manifold. For instance, there is no rationale for assuming that the results from observation necessarily lead to content analysis. In fact, we found examples in the literature where observational data was analysed quantitatively, qualitatively, or both (we refrain from designating the latter approach as mixed method because it brings forward the discussion on research methods, which we regard as a different concern). In fact, we attempted the exercise of linking data collection and analysis using the review data and found examples connecting most patterns in different ways. Because of this, we do not link the two categories. We only emphasise they are logically bound through the theory building/testing framework in Figure 4.2.

4.5 EVALUATION

Card sorting is a qualitative evaluation method that has been widely used in various fields such as psychology, knowledge engineering, and software engineering (Barrett & Edwards, 1995), as well as DSR (Prat et al., 2015). The method helps us understand the relationship between an artefact and the people the artefact is designed for (Spencer & Warfel, 2004). In particular, the method can help explain how users structure information and action mediated by the artefact (Spencer & Warfel, 2004). For instance, card sorting can be used to assess the users' needs and priorities, how they deal with information, and how they react to features provided by the artefact. Finally, through card sorting we can also gather insights on how users view the problem and the solution addressed by an artefact. This latter consideration is extremely important in the DSR context, since DSR addresses wicked problems for

which there are no single best problem definitions and solutions (Rittel & Webber, 1973). All in all, we use card sorting to evaluate how our ensemble artefact relates to what exists in the researchers' minds. We structured the card sorting approach as a combination of card sorting exercises and interviews. They were conceived to acquire expository knowledge regarding:

- E1:** Testing the understanding of theory generation and its cycles based on the developed framework;
- E2:** Testing the ease of understanding of the pattern model and how users interact/relate with it based on their research project;
- E3:** Determining which elements defined by the pattern model the participants find difficult to understand;
- E4:** Establishing how useful the participants find the pattern model in order to conceptualise their research project.

Setting and participants. The card sorting exercises were conducted in a meeting room, which had a large table where to lay out the cards. The participants were PhD students conducting research projects in the IS field and IS Academics. PhD students were selected because they represent the main target audience for this research, while the IS Academics were also used to get their opinions and feedback based on their years of experience. 15 participants (10 PhD students and 5 academics) were selected by convenience.

Materials. The card sorting materials involved sets of 90 cards with words on them and a unique identifier number for recording purposes. Blank cards were also provided to write on if needed.

Procedure. The exercises were done in one-on-one sessions moderated by the first author (Figure 4.5). At the beginning of each session, the participant was informed about the purpose of the exercises, the artefacts under evaluation, and the card sorting technique. The participants were then encouraged to express their thoughts and opinions during the exercises, to provide detailed feedback about what they were thinking and doing, and to cater for the diversity of research problems and methods



Figure 4.5: Card sorting exercise

that are typical in the IS field. The exercises were then operationalised according to the following steps.

The purpose of each exercise was explained before handing over a deck of cards to the participant. The participant was then given some time to read through the cards for familiarization of contents. The first exercise contained a deck of cards with theory testing activities, which the participant should group in five categories, from very difficult to understand to very easy to understand. The second exercise, the pattern model was shown, and the participants were required to delineate her/his research according to the model.

Along with information about the card sorts, we also gathered the participant's reactions and comments made throughout the exercises, focussing on the positive and negative reactions. The exercises were audio recorded and captured to ease later analysis. All participants completed the entire procedure and each session took an average of 30 minutes, even though some exceeded 50 minutes.

4.5.1 Understanding of Theory Generation and its Cycles (E1)

Participants were told to sort the cards into two categories: theory testing and theory building activities. The cards included theory building, testing and boundary activities. The cards had unique numbers, making it easy to know how many cards were

mismatched. As a result, over 70% were able to sort the cards into the right categories (see figure 4.6).

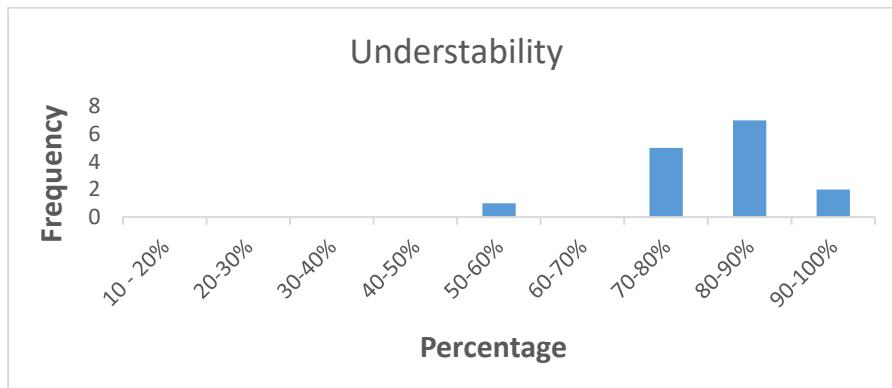


Figure 4.6. Card sorting exercise result showing Framework Understability

Two interesting issues emerged from this exercise. First, card arrangements/sorting was based primarily on the participant's area of research specialization. For example, researchers that focused on grounded theory found it difficult placing the theory testing cards, because they felt it can 'go both ways', since they usually collect data to build or develop their theory. Those with an engineering background couldn't really place where the cards belong, as they have less experience with theory development. Those who were more familiar with quantitative methods could place the cards appropriately, and only had issues with the boundary activities, which could go both ways. Consequently, the exercise confirms the usefulness of this framework, both for new and more experienced researchers, as it helps them navigate from the conceptual to the operational.

Secondly, many participants were uncertain about the boundary activities, also feeling they could 'go both ways'. This uncertainty provides further support for the notion they are best classified as boundary activities, nestled between the building and testing domains.

4.5.2 Understanding of Pattern Model (E2)

This step evaluated how easy it was for participants to interact and relate with the pattern model. The model was presented to the participants, and instructions were given on how to use the model. Participants were told to consider any part of their research project, focusing on the activities they have either carried out, or will be carrying out, and then map those activities to the model. Blank cards were given so they could write down any activities not represented in the model. We also encouraged participants to consider patterns they haven't used or are not familiar with, and then try using the model to grasp how easy it was for them to understand the links and their associated activities.

As a result, 30% of the sample said that it was very easy to understand the model and relate that to their research project, while 60% said it was somewhat easy, another 10% responded that it was not easy at all to understand (see figure 4.7). We interviewed the participants that said it wasn't easy to understand to know the reason behind their response, and most of them said it was due to how the model was positioned, having to work 'backwards' from right to left. When asked to suggest how it could be repositioned for better understanding, they tried different styles, but concluded it was better off how it was. Generally, most of the participants found it easy to understand.

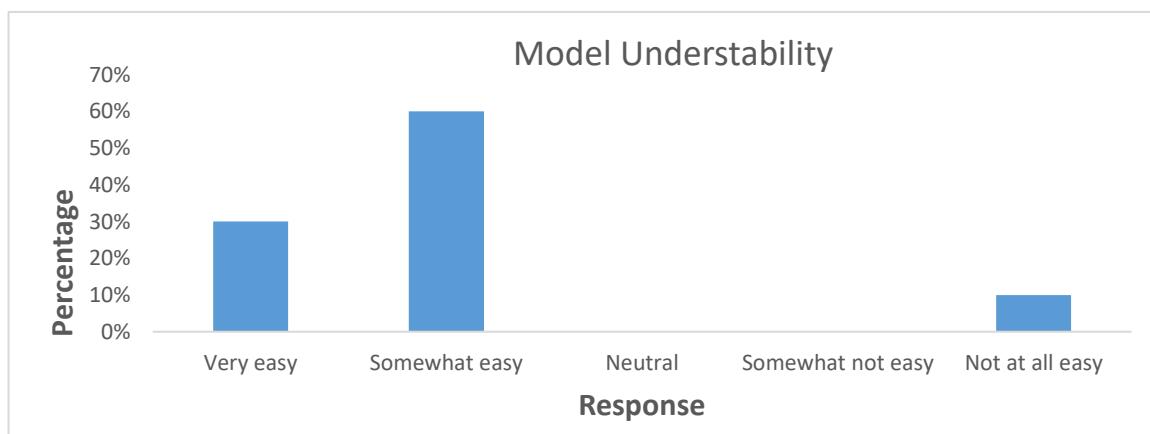


Figure 4.7. Card sorting exercise result showing Model Understability

4.5.3 Understanding of the Elements defined by the Pattern Model (E3)

This step evaluated how easy it was for participants to understand the elements described within the pattern model. Participants were given blank cards and were told to go through all the elements and labels used in the model, and then they were to write down elements that they found very difficult to understand. Based on content analysis, we found more than half of the participants fully understood the elements, but some confusion clearly remained. One participant picked out the “execute task” card, and said he didn’t fully understand it, as the word “execute” sounded like a computer programming phrase. Another participant picked the card labelled “assign role”, indicating she had never used it, and didn’t fully grasp what it meant. Where participants were unclear about specific card labels, their responses largely indicated a lack of familiarity with the research methods.

4.5.4 Usefulness of Pattern Model (E4)

The last exercise performed by the participants was aimed at determining if the participants found the pattern model useful to conceptualise their research project. From the result, 40% of the participants said that the model was very useful, while 60% said it was somewhat useful (see figure 4.8).

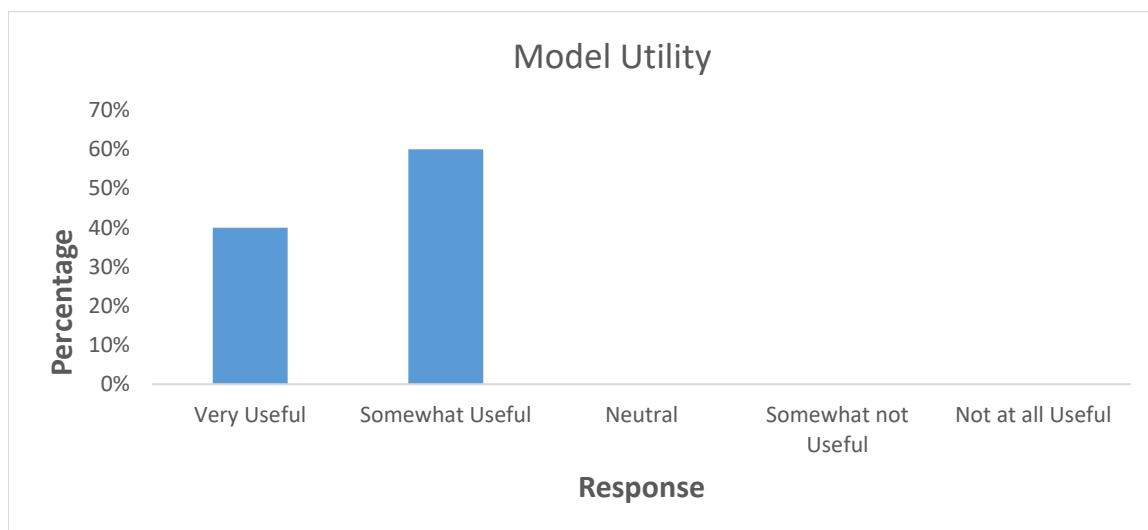


Figure 4.8. Card sorting exercise result showing Model Usability

Some of the participants were thrilled about the model, as it was pulled from top journals in the IS field, giving them an insight of what is expected from them if they

aspire to get their papers accepted in such journals. One participant, a relatively inexperienced researcher, said the model was very useful for practitioners wanting to carry out some form of testing, as it gave them an idea about different intents, patterns and pathways. More experienced researchers thought the model will benefit new researchers and felt comfortable recommending it to their students. The exercise encouraged the participants, experienced or not, to think carefully about testing, which was a positive outcome. Finally, there was also considerable positive feedback about the card sorting exercises, as it provided an effective opportunity to compare their thoughts with the conceptual framework.

4.6 DISCUSSION AND CONCLUSIONS

This section is structured around the two artefacts developed by this research and their contributions to research and practice. The first artefact we developed was the theory building/testing framework. This was initially developed as an intermediate artefact, necessary to define the perimeter of the research and core conceptual elements of the phenomenon of interest. However, during the research, we realised that very often theory testing *per se* has not been considered as an independent phenomenon of interest. It is present in most research endeavours but is often entangled with theory building and taking a subsidiary role.

Our first undertaking consisted in delimiting the line that separates theory building from theory testing, which resulted in the two related cycles, the first one encompassing theory building and theory testing, and the second cycle partitioning theory testing into data collection and data analysis. We suggest this separation contributes to developing an operational view of theory testing as an independent phenomenon, which in turn provides some interesting design-science contributions to research and practice. A fundamental goal of design science in the IS tradition is to contribute with artefact knowledge (Gregor & Hevner, 2013): principles and technological rules that can be repeatedly used in multiple applications. In other words, the framework regards research as a recurrent activity, which may benefit from the adoption of operational principles and architectures about the research activity.

Of course, the nature of research itself has long been the subject of investigation, e.g. in the fields of philosophy of science, epistemology, ontology, and methodology. The novel viewpoint proposed in this paper is using design science to propose an operational framework that not only clearly distinguishes theory building from theory testing, but also identifies a set of boundary activities that have to be considered by the researcher when aligning theory building and testing. This artefact provided a firm basis for the development of the second artefact.

The second artefact developed by this research is a model describing a set of theory testing activities organised as a pattern system. The model was developed from a review of 404 papers published in the last two years in the most prestigious journals in the IS field. The model distils the different theory testing approaches adopted in these papers into a set of patterns, which have been divided into two groups considering data gathering and data analysis, and two other groups considering primary and secondary activities. The model contributes to understand theory testing as a pattern system, which promotes order and flexibility.

Once again, the value of this second artefact is strongly related with the fundamental goals of design science: to contribute design knowledge that will help practitioners (in this case, researchers) to utilise principles and technological rules in their endeavours (in this case, theory testing). Pattern systems reveal best practices in design. They offer a multitude of patterns that can be adopted according to different constraints and intents. Furthermore, pattern systems are conceptually open to further additions and developments. Such openness is particularly important, given the limitation that the collection of papers that were considered is just a sample of the research that has been conducted in the past, and is neither representative of the whole body of research undertaken to date nor expected to be undertaken in the future.

Nevertheless, the openness of the pattern model allows for further additions and refinements to the patterns described. Therefore, while neither wholly complete nor prescriptive, these patterns elucidate what best practices have been adopted, and provide a certain coherence about theory testing that emphasises an operational viewpoint. By adopting such a pattern system, researchers may benefit from stronger links to best practices, as well as stronger transparency and coherence when reporting

their research. By expanding or contradicting such a pattern system, researchers may also benefit from a frame of reference against which they can claim the novelty and strengths of their research approaches.

4.7 REFERENCES

- Alexander, C. (1999). The origins of pattern theory: The future of the theory, and the generation of a living world. *IEEE software*, 16(5), 71-82.
- Bacharach, S. (1989). Organizational theories: Some criteria for evaluation. *Academy of management review*, 14(4), 496-515.
- Barrett, A., & Edwards, J. (1995). Knowledge Elicitation and Knowledge Representation in a Large Domain with Multiple Experts. *Expert Systems with Applications*, 8(1), 169-176.
- Bhattacherjee, A. (2012). Social science research: principles, methods, and practices. Textbooks Collection. 3. http://scholarcommons.usf.edu/oa_textbooks/3
- Bichler, M., Frank, U., Avison, D., Malaurent, J., Fettke, P., Hovorka, D., . . . Thalheim, B. (2016). Theories in business and information systems engineering. *Business & Information Systems Engineering*, 58(4), 291-319.
- Bider, I., Johannesson, P., & Perjons, E. (2013). Design science research as movement between individual and generic situation-problem-solution spaces *Designing Organizational Systems* (pp. 35-61). Berlin, Heidelberg: Springer.
- Bitektine, A. (2007). Prospective case study design: Qualitative method for deductive theory testing. *Organizational Research Methods*, 11(1), 160-180.
- Borgman, C., Wallis, J., & Enyedy, N. (2007). Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries. *International Journal on Digital Libraries*, 7(1-2), 17-30.
- Burton-Jones, A., & Volkoff, O. (2017). How can we develop contextualized theories of effective use? A demonstration in the context of community-care electronic health records. *Information Systems Research*, 28(3), 468-489.
- Colquitt, J., & Zapata-Phelan, C. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal*, 50(6), 1281-1303.
- De Vaus, D. (2013). *Surveys in social research*: Routledge.
- Estelles-Arolas, E., & Gonzalez-Ladron-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189-200. doi:10.1177/0165551512437638
- Fusch, P., & Ness, L. (2015). Are we there yet? Data saturation in qualitative research. *The qualitative report*, 20(9), 1408.
- Goode, W., & Hatt, P. (1952). Methods in social research. New York: Megraw-Hill Book Company.
- Gregor, S. (2002). A theory of theories in information systems *Information Systems Foundations: building the theoretical base* (pp. 1-20). Canberra: Australian National University.
- Gregor, S. (2006). The nature of theory in information systems. *MIS quarterly*, 30(3), 611-642.

- Gregor, S., & Hevner, A. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 37(2), 337-355.
- Gregor, S., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association of Information Systems*, 8(5), 312-335.
- Hevner, A., & Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol. 22): Springer Science & Business Media.
- Hevner, A., March, S., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *Management Information Systems Quarterly*, 28(1), 75-105.
- Kampffmeyer, H., & Zschaler, S. (2007). Finding the pattern you need: The design pattern intent ontology *International Conference on Model Driven Engineering Languages and Systems* (pp. 211-225). Berlin, Heidelberg: Springer.
- Khan, J. (2011). *Research methodology*: APH Publishing.
- Kitchenham, B., Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1), 7-15.
- Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American journal of occupational therapy*, 45(3), 214-222.
- Kumar, R. (2005). *Research methodology*. Australia: Pearson.
- Li, M., Tan, C., Wei, K., & Wang, K. (2017). Sequentiality of product review information provision: an information foraging perspective. *MIS Quarterly*, 41(3), 867-892.
- Lokke, A., & Sorensen, P. (2014). Theory testing using case studies. *Electronic Journal of Business Research Methods*, 12(1), 66-74.
- Lu, Y., Gupta, A., Ketter, W., & Van Heck, E. (2016). Exploring Bidder Heterogeneity in Multichannel Sequential B2B Auctions. *Management Information Systems Quarterly*, 40(3), 645-662.
- Lynham, S. (2002). The general method of theory-building research in applied disciplines. *Advances in developing human resources*, 4(3), 221-241.
- Maaninen-Olsson, E., Wismén, M., & Carlsson, S. (2008). Permanent and temporary work practices: knowledge integration and the meaning of boundary activities. *Knowledge Management Research & Practice*, 6(4), 260-273.
- Miles, M., & Huberman, A. (1994). *Qualitative Data Analysis*. Thousand Oaks, California: SAGE.
- Muntermann, J., Nickerson, R., & Varshney, U. (2015). Towards the Development of a Taxonomic Theory. Paper presented at the 21st Americas Conference on Information Systems, Puerto Rico.
- Niederman, F., & March, S. (2015). Reflections on Replications. *AIS Transactions on Replication*, 1.
- Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts Work. Pap. Inf. Syst.*, 10, 26.
- Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. *European Journal of Information Systems*, 25(6), 1-16.
- Paré, G., Trudel, M., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183-199.
- Penker, M. (2000). *Business Modeling with UML: Business Patterns at Work*: New York, NY: John Wiley & Sons.
- Popper, K. (2014). *Conjectures and refutations: The growth of scientific knowledge*: Routledge.

- Prat, N., Comyn-Wattiau, I., & Akoka, J. (2015). A taxonomy of evaluation methods for information systems artifacts. *Journal of Management Information Systems*, 32(3), 229-267.
- Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, 4(2), 155-169.
- Roberts, N., Campbell, D., & Vijayasarathy, L. (2016). Using information systems to sense opportunities for innovation: integrating postadoptive use behaviors with the dynamic managerial capability perspective. *Journal of Management Information Systems*, 33(1), 45-69.
- Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS quarterly*, 35(1), 37-56.
- Selander, L., & Jarvenpaa, S. (2016). Digital action repertoires and transforming a social movement organization. *MIS Quarterly*, 40(2), 331-352.
- Simon, H. (1996). *The Sciences of the Artificial*. Cambridge, USA: The MIT Press.
- Sonnenberg, C., & vom Brocke, J. (2011). *Evaluation patterns for design science research artefacts*. Paper presented at the European Design Science Symposium.
- Spencer, D., & Warfel, T. (2004). Card sorting: a definitive guide. *Boxes and Arrows*, 2. Retrieved from <https://boxesandarrows.com/Card-sorting-a-definitive-guide/>
- Stinchcombe, A. (1987). *Constructing social theories*: University of Chicago Press.
- Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, 48(4), 865-885.
- Wacker, J. (1998). A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of operations management*, 16(4), 361-385.
- Zhang, K., Bhattacharyya, S., & Ram, S. (2016). Large-Scale Network Analysis for Online Social Brand Advertising. *MIS Quarterly*, 40(4).
- Zigurs, I., & Khazanchi, D. (2008). From profiles to patterns: A new view of task-technology fit. *Information systems management*, 25(1), 8-13.

CHAPTER 5

TOWARDS THE DEVELOPMENT OF A DSS SUPPORTING THE INTEGRATION OF CROWDSOURCING IN THEORY TESTING: ANALYTICAL FRAMEWORK DESIGN

ABSTRACT

Crowdsourcing can be an adopted strategy for researchers where tasks are distributed to internet users to harness different forms of data, which adds to the reliability and validity of the research process. As theory testing is an essential part of the research process, involving more of operationalization than conceptualization, of which different forms of activities that most times needs input from different and diverse participants is required, the need to consider what theory testing activities can be crowdsourced is paramount. Adopting a design science paradigm to manage this challenge, we design an analytic framework which comprises of important attributes that need to be considered if crowdsourcing is to be used for any of the theory testing activities. The framework which was justified by the use of sample cases, gives us an insight to what attributes make such activity to be easy to crowdsource, difficult to crowdsource or impossible to crowdsource. The value of this artefact lies in its capacity to help researchers making complex design decisions, in this case regarding how they can utilise crowdsourcing to their advantage.

Keywords: Crowdsourcing, Theory testing, Design Science, Framework

5.1 INTRODUCTION

In academia, a discipline will benefit from having theories that provide a basis for categorizing, understanding, predicting, and utilizing the knowledge accumulated in the field. At any given point in time, the set of theories available to a given field will be tentative and subject to continual improvement. Young fields with less tested theory should not be harshly viewed for having less theory simply as a function of insufficient time to mature. Theory is an important element among the many contributions a discipline can make to our common body of knowledge and, relative to business, our ability to apply that knowledge to practice. (Niederman & March, 2015). Theory helps

scientists to describe and explain a process or sequence of events. It also helps scholars to explain the complexities of the empirical world by providing a linguistic tool for organizing it (Colquitt & Zapata-Phelan, 2007).

The test of a theory is concerned with establishing the strength of the relationship between underlying concepts as they act in the "real world". As a theory is developed (theory building) and tested, the resulting outcome either becomes the basis for codification of knowledge and application to new situations both for further testing and for use in practice or offers an opportunity for further studies if diverse results are observed. This is because testing and retesting theory can lead to the accumulation of knowledge within a field of study(Niederman & March, 2015).

Theory testing which is an integral part of theory building is very important to the development of Information Systems (IS) discipline, just as Bacharach (1989) said, "if it is not testable, no matter how profound or aesthetically pleasing it may be, it is not theory". We are therefore interested in supporting theory testing in the IS field and how crowdsourcing (CS) – which is a strategy in which a single task can be fragmented into multiple tasks which can then be delivered to a large group of people in the form of an open call (Howe, 2006), can be used to support the theory testing process.

In previous paper (Enwereuzo et al., 2017), we analysed how theories have been tested in IS and we proposed a conceptual framework and a model. We further tested the framework and model using card sorting and the outcome showed that the framework is sound and that students (PhD) couldn't comprehend the theory testing process and the activities involved. The outcome also indicates that even though PhD students were knowledgeable about theory building, they were not familiar with theory testing. The experiment conducted showed that PhD students were not aware of the variety of pathways that can be used to test theory. They also seemed to be unaware of the diversity of theory testing activities that can be crowdsourced.

Therefore, it is one thing to know the steps to take; it is another to know what types of activities are involved in those steps, especially during the validation stage, which comprises of data collection and analysis. Having identified those activities, there are different pathways to choose when carrying out the testing process. A pathway is a set of linked activities that instantiate an intent from begin to end, while an Intent

summarizes the general purpose of a study taking a theory testing perspective, which does not concern the specific problem under investigation (Penker, 2000).

Having studied the conceptual framework and model which constitute a solid foundation for this study, the next logical step in our research is the development of a support system which aims at assisting researchers to know the theory testing steps, theory testing pathways and activities, crowdsourcing requirements and how to make decisions as to which theory testing activity can be crowdsourced. This type of support is usually associated with the Decision Support Systems (DSSs) (Arnott & Pervan, 2005; Hosack et al., 2012). The research objective of this paper is to *analyse and design a CS analytic framework for supporting the integration of CS in theory*. This is a step towards developing the primary artefact- a DSS, which is a future study.

We adopted the design science paradigm to address this challenge. It emphasizes a developmental viewpoint where the goal is to iterate problem and solution frames. To provide a solid knowledge base for the development of this artefact, the study will rely on the developed *conceptual framework* and a *model* which are two problem frames necessary to build the DSS, previously developed in previous project (Enwereuzo et al., 2017). The last task involved in building the DSS would be the development of a DSS artefact, which would be in the form of a DSS framework for theory testing. The framework consolidates existing research knowledge in a structured manner.

The remainder of this paper is organized as follows. Section two describes the research background of the study. Section three describes the method adopted for the study. In Section four, we look at the tool architecture, development and evaluation. Finally, Section five highlights the research contributions.

5.2 RESEARCH BACKGROUND

CS is an emerging strategy that fragments a single task into multiple tasks delivered to a large group of people in an open call (Howe, 2006). Several researchers have considered the adoption of CS in theory testing. For instance, Lowry et al. (2016) advocated CS as a way to increase the quality of data collection. The advantages brought by CS concerning demographics, psychometrics, and structural properties of data samples were highlighted by Steelman et al. (2014). Their research also provided

initial empirical evidence that CS can tap into large samples of participants and reach a wide variety of demographics.

We note that using crowdsourcing for data collection has been around for some time. Crowdsourcing is becoming a very common, if not the predominant, way to deliver questionnaires in behavioral research (Bates & Lanza, 2013; Behrend et al., 2011; Jarmolowicz et al., 2012; Peer et al., 2017). In user studies, it has also started to be used as an instrument to collect user data (Kittur et al., 2008; Stewart et al., 2017). And it is also regarded as an important component of citizen science, where it supports the distributed collection of research data (Bonney et al., 2009; Gura, 2013). However, the relationship between crowdsourcing and theory testing with human participation has not yet been systematically explored and established. Establishing this relationship is important from both theoretical and practical perspectives. From a theoretical viewpoint, finding innovative ways to research humans is paramount to rigorously test existing theories and advance new theories, while from a more practical viewpoint, existing crowdsourcing platforms open up opportunities to develop relatively easy-to-implement methods to conduct theory testing.

The possibilities brought by crowdsourcing seem even more relevant in a research landscape where problems are becoming increasingly complex and the time available for completing doctoral studies remains constant, if not shrinking (Blagojević et al., 2017). This is especially true in the Information Systems (IS) field where theory building, and therefore theory testing, have a fundamental role in shaping the identity of the field (Gregor, 2006). Furthermore, it has been reported that participation rates in scientific studies have been declining (Galea & Tracy, 2007; Van Gelder et al., 2010). Finally, as alternative dissertation formats, such as the three-paper dissertation, are becoming increasingly popular, the need for multiple theory testing efforts is likely to increase (Fong, 2017; Jump, 2015). The crowdsourcing strategy offers a new way to incentivize humans to participate in research. Some major advantages of using crowdsourcing above other theory testing methods are:

- a) Logistics: is the management of the flow of things between the point of origin and the point of completion in order to meet set requirements. This includes

cutting costs, managing recruitment, selecting participants. These forms of logistics can be easily done when using crowdsourcing.

- b) Controlling the environment: such as enforcing boundaries, defining/forcing certain actions.
- c) Controlling the experiment: these includes dynamic interventions, experimental manipulations etc.
- d) Diversity/heterogeneity of experimental subjects

The attributes of CS are essential to theory testing, and when incorporated makes the theory testing process much easier and with less bias.

5.2.1 Crowdsourcing Requirements in Respect to Theory Testing

Requirements are important and should be clearly defined because it is essential for both the technical designers and the users of the new system or developed tool to be absolutely clear on what they want and expect the system to do (Mumford, 1985).

Requirements here tends to lean more to the non-functional requirements, which deals more on the constraints, qualities, characteristics and properties that a system or tool should possess (Malan & Bredemeyer, 2001). For crowdsourcing to be used for theory testing, these set of requirements must be considered both by the user and for designing of the tool. This will serve as a checklist when designing and for recommendations of the different theory testing activity that can be crowdsourced. The requirements can be categorized into two: Mandatory requirements – requirements that must be met before CS can be used, they are based on platform conditions and limitations, and are what CS is made up of. Desirable requirements – are requirements that are flexible depending on what aspect the researcher is exploring. They don't have to be met for CS to be considered.

Mandatory Requirements

- Horizontal decomposability: This is a requirement that states that tasks must be decomposable in several simple, independent tasks that can be executed in parallel by the crowd. Task decomposition leads to faster task turn-around time,

which is the main advantage of CS (Heer & Bostock, 2010; Vukovic, et al., 2010)

- One-off: This requirement entails that the crowd task must be executed only once. Repeated tasks are usually not allowed (platform limitation) (Mason & Suri, 2012; Yang et al., 2008).
- Simplicity: The crowd task should be meaningful, easy to delineate and have clear inputs and outputs. Simplicity is considered one of the most important conditions for the success of CS (Geerts, 2009).
- Limited interaction: The crowdsourcer and the crowd have a limited channel for interaction, this implies that communication loops are usually not possible (platform limitation) (Wexler, 2011; Wu et al., 2013).
- Computer mediation: The crowd task should be remotely executed, using the Internet for coordination and communication. CS is mainly carried out on different platforms, and most CS platforms are web-based (Andriole, 2010; Kleemann et al., 2008).

Desirable Requirements

- Time-bound-ness: It should be possible to complete the crowd task in a bounded period of time. One of the advantages of using CS is having and obtaining a quick turn around on task, therefore, timing is crucial to any CS task (Bernstein et al., 2011; Kittur et al., 2008).
- Skills matching: It should be possible to match the domain knowledge required by the crowd task with the skills of the selected members of the crowd. Mismatch of skills and tasks leads to low quality results (Geiger et al., 2011; Schenk & Guittard, 2011).
- Training: Training, if required, should not be complex or prolonged. In current CS platforms, training is still considered challenging (Park et al., 2014).

5.3 RESEARCH METHOD

The design science paradigm was adopted for this research. We adopted this paradigm because design science is a problem-solving paradigm that seeks to create and evaluate IS artefacts to solve identified organizational problems (Gregor &

Hevner, 2013). Design science involves the designing of artefacts, making research contributions, evaluating the designed artefacts, and communicating the results to appropriate audiences (Peffers et al., 2007). By artefacts we mean anything that can be transformed into a material existence as an artificially made object (e.g., model, instantiation) or process (e.g., method, software) (Goldkuhl, 2002; Gregor & Hevner, 2013). Artefacts may also include social innovations (Aken, 2004), or new properties of technical, social, or informational resources (Järvinen, 2007). In essence, any designed object with an embedded solution to a research problem could be seen as design science (Peffers et al., 2007). We therefore adopted this paradigm because it is:

- Adequate to research goals
- Common approach to DSS development (80% according to Pervan)
- Accepted in IS field (journals, DESRIT conf)

Design science is focussed on the iterative design of IS artefacts. In our case, our iterative development started with the development of the conceptual model which is a secondary artefact, then we proceeded to the model development, which is also a secondary artefact, and the DSS tool is the primary artefact. The primary artefact which is the DSS tool evolved from the development of the secondary artefacts and will be used to provide users' support and guidance (Hevner & Chatterjee, 2010; Von Alan et al., 2004). These secondary artefacts which has already been developed (Enwereuzo et al, 2017) serves as a knowledge base and foundation to the development of the primary artefact.

The design science paradigm which views the to-be-built tool as an artefact distinguishes two primary research activities: build and evaluate (March & Smith, 1995) (see Figure 5.1). The build activity involves generating an initial artifact based on a preliminary problem frame and a set of requirements, which then lead to the artefact design. The build activity is then followed by an evaluation activity, which confronts the design against the objectives and suggests a new problem frame, which may then lead towards another design cycle, until a satisficing solution is obtained (Sein et al., 2011).

When choosing an evaluation criteria or method, it is important as a design researcher to pay attention to balance the interests of the practitioners and researchers (Sonnenberg & vom Brocke, 2011), while the practitioners are interested in the applicability and usefulness of the artefact, the researchers are interested in the validity of the artefact. Bearing this in mind led to the selection of the evaluation methods. Literature review was used for the conceptual framework and scoping knowledge source because its focus is more on validity than utility, while the model and DSS artefact, while we are still interested in their validity, its usefulness is much more important, which led to the use of card sorting as an evaluation method.

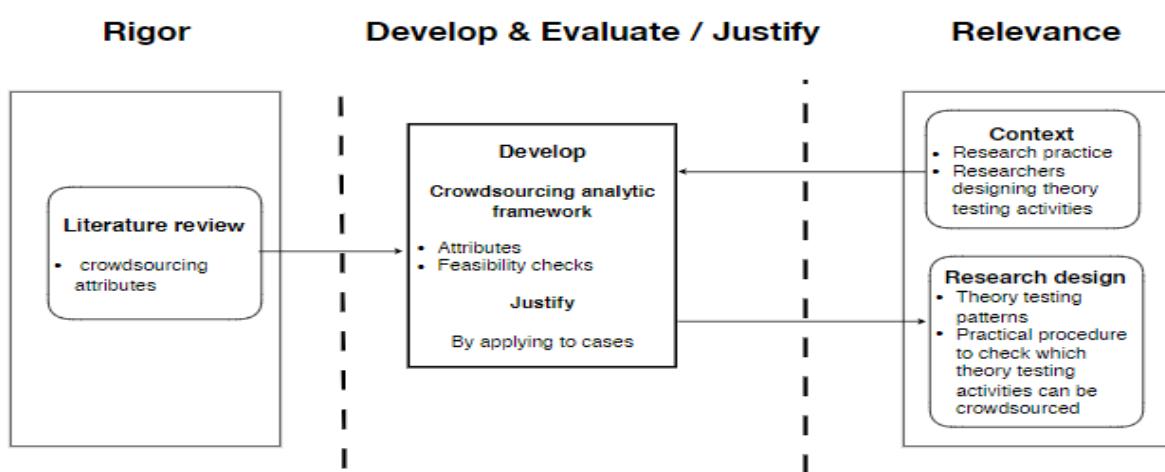


Figure 5.1. Build and evaluate activities within a DSR method (Adapted from (Peffers et al., 2007b)

The evaluation for this artefact will be done using example cases from existing literature (Sonnenberg & vom Brocke, 2011). This evaluation serves the purpose of showing that the artefact design incorporates the solution to the stated problem (Sonnenberg & vom Brocke, 2011).

5.4 DECISION SUPPORT CROWDSOURCING ANALYTICAL FRAMEWORK

The Decision Support Crowdsourcing Analytic Framework Development (DSCAFD) is based on the two secondary artefacts, which are the conceptual model and the model (Enwereuzo et al., 2017). The model feeds the DSCAFD and serves as a knowledge base to the DSCAFD. The DSCAFD basically allows users to play with the model using a set of restrictions and requirements.

5.4.1 Crowdsourcing Attributes

The analysis of which activities can be crowdsourced or not needs to be systematic and anchored on a transparent procedure. To increase systematicity and transparency, we first define the set of crowdsourcing attributes considered in this study. The attributes are organized in four categories related to the task, medium, crowd, and quality. We consider these to be the core attributes characterising crowdsourcing (See figure 5.2).

Task. The task corresponds to “what the crowd has to do” (Estelles-Arolas & Gonzalez-Ladron-de-Guevara, 2012). Some tasks are completely independent of the context in which they are executed, e.g. classifying photos according to a set of given categories, which means they are *context-free*. Other tasks depend on a specific context to be executed, such as a physical location, operational environment, or virtual place. Examples include gathering data on public libraries, while driving a car, or from commonly used social media. In these cases, tasks are *context-dependent*. Finally, some tasks may have to be performed on restricted contexts, which may not be easily accessible or reproducible. Examples include gathering data from hospitals, coding a data set using a proprietary tool, and selecting records from a private database. In these cases, we designate the tasks as *context-restricted*. We consider that context-free tasks are easy to crowdsource, while context-dependent tasks are difficult to crowdsource, and context-restricted tasks cannot be crowdsourced. The rationale for considering context-dependent tasks difficult to crowdsource is that certain constraints have to be enforced, which make tasks more difficult to specify, control and execute. In principle, we consider context-restricted tasks as not crowdsource-able. This is because, if would be possible to circumvent the restrictions, then they would be dependencies. Consider as an example that you would like to crowdsource data collection in an hospital. Sending the crowd to the hospital without consent would be unethical, and therefore the task would be context-restricted. However, if you can get a consent from the hospital, then the task becomes context-dependent.

Tasks can also be classified according to complexity, structure and interdependence. Regarding complexity, a task can be classified as *micro* if it can be done quickly, uses repetition, or requires little cognitive effort; or it can be classified as *macro* if it requires

considerable cognitive effort (e.g. planning what to do and making decisions (Kittur et al., 2011)). Structure classifies the task's problem and solution as either *well-defined* (e.g. classifying photos using a set of tags) or *ill-defined* (e.g. generating new research ideas) (Nakatsu et al., 2014). Interdependence classifies how the task outputs are generated. The task outputs can be *pooled* (e.g. individual ideas can be pooled together), *coordinated* (e.g. gather data first, and then analyse data), or *shared* (e.g. collaborative writing) (Crowston, 1997; Malone & Crowston, 1994). We consider that well-defined, pooled and micro-tasks can be easily crowdsourced, while ill-defined, coordinated or shared tasks can be difficult to crowdsource.

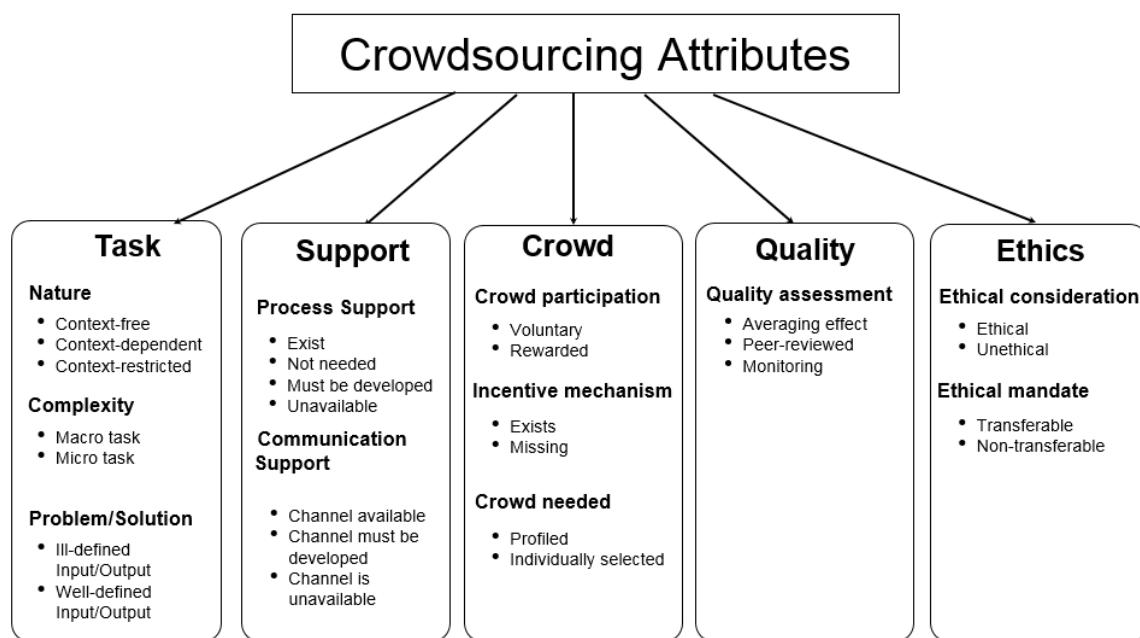


Fig. 5.2 *Crowdsourcing Attributes*

Support. Firstly, we consider communication support, which allows sending the inputs and getting the outputs from the crowd. Regarding this attribute, we consider the following values: *channel available* (e.g. using a crowdsourcing platform or social media), *channel must be developed* (e.g. a mobile application must be created to communicate with the crowd), or *channel is unavailable* (e.g. the function must be communicated within proprietary software, which cannot be accessed by the crowd).

Secondly, we consider process support, which allows the crowd to perform the function. Regarding this attribute, we consider the following values: *process support*

not needed when the function can be carried out by the crowd with no additional support (e.g. counting birds in a forest), *process support exists* when there is an existing technology that can be used by the crowd to execute the function (e.g. a crowdsourcing platform such as MTurk, InnoCentive, Threadless, and iStockphoto), *develop process support*, when such support must be developed by the researcher (e.g. a website must be created to classify an ecosystem), and *process support unavailable*, when process support is needed but cannot be developed.

Finally, we have to consider what type of coordination support is required to collect the outputs generated by the crowd. Regarding this attribute, we consider the following values (Crowston, 2012; Malone & Crowston, 1994): *pooled outputs* (e.g. individual ideas can be pooled together), *coordinated outputs* (e.g. gather data first, and then analyse data), or *shared outputs* (e.g. collaborative writing).

We consider that crowdsourcing cannot be done if no channel is available or process support is unavailable. If the communication channel is available, the outputs can be pooled, and process support exists or is not needed, then it is easy to crowdsource. Also, if the communication channel is available, process support exists, and outputs are coordinated or shared, then it is easy to crowdsource. All other combinations of attributes should be seen as difficult to crowdsource. If the communication channel must be developed or process support must be developed, it is difficult to crowdsource.

Crowd. The crowd is the collection of people who execute the task. The participation of the crowd is either *voluntary* (e.g. open software development) or *rewarded* (e.g. prizes and money). We should also consider if either an *incentive mechanism exists* (e.g. most crowdsourcing platforms offer reward mechanisms) or an *incentive mechanism is missing*.

We consider it is easy to crowdsource if an incentive mechanism exists and the participation is rewarded, and if an incentive mechanism is missing but participation is voluntary. An example of the latter case is the successful use of crowd mapping in disaster relief actions, which depend on voluntary contributions (Gao et al., 2011).

The crowd may have to be either *individually selected* or *profiled* according to a set of criteria (e.g. generic skills, geographical locations, or answers to a questionnaire (Amazon, 2011)). We consider that it is easy to crowdsource when the crowd can be profiled, while individually selecting the participants makes it difficult to crowdsource.

Quality. Considering the loose characteristics of crowdsourcing, it seems natural that the task outputs should be subject to quality assessment. Quality assurance may rely on the *averaging effect*, which uses multiple data sources to increase accuracy and trust (Brunt & Meidell, 2018). It can also be *peer-reviewed*, where quality assurance is actually done by the crowd. And finally, it may also rely on *quality monitoring*, where either a sample or the entire data set is checked for accuracy by the researcher. We consider that, if the averaging effect or peer-reviews are viable, then the task is easy to crowdsource. Monitoring has a neutral impact in relation to crowdsourcing, as it is a traditional approach to quality assessment.

Ethics. We should also consider if it is either *ethical* or *unethical* to execute a task. We suggest that an unethical task cannot be crowdsourced (e.g., gathering data from inside a hospital without consent). If the task is ethical, we need to further consider if the *ethical mandate is transferable* (i.e., if the mandate obtained by the researcher can be transferred to the crowd) or the *ethical mandate is non-transferable* (i.e., only the researcher has the mandate, which cannot be transferred to anyone else). We suggest that neither unethical tasks nor non-transferable mandates should be crowdsourced.

5.4.2 Analytic Framework Development

Based on the set of crowdsourcing features discussed above, we now propose an analytic procedure to check if a task can be crowdsourced or not. The procedure is applied to each individual task that one wishes to crowdsource and checks systematically how the attributes related to the task, structure, crowd, quality, and ethics determine the feasibility to crowdsource. As illustrated in Figure 5.3, the feasibility test may have three results: 1) we cannot crowdsource; 2) we can crowdsource, but it may be difficult to accomplish; and 3) it is easy to crowdsource.

The first check to do regards the impossibility to crowdsource (Figure 5.3, left). We discussed five conditions under which it is not possible to crowdsource. They are

related to the task context, channel availability, process support, and ethics. The next feasibility check determines if crowdsourcing is difficult to accomplish (Figure 5.3, centre). We found seven conditions making a task difficult to crowdsource. They are related to the task context, type of task, types of input and output, channel and support development, and crowd selection.

The final feasibility check determines if crowdsourcing is easy to accomplish (Figure 5.3, right). Here, various combinations of conditions have to be fulfilled to determine that a task is easy to crowdsource. Overall, they consider 19 different conditions. We observe that some combinations of conditions may lead to conflicting results. For instance, if the task is classified as macro, then it is difficult to crowdsource. And if the crowd has to be profiled, then it is easy to crowdsource. In these cases, we adopt a pessimistic approach and suggest that any conditions leading towards the impossibility to crowdsource should take precedence over any other conditions; and any conditions leading towards the difficulty to crowdsource should take precedence over any other conditions indicating it is easy to crowdsource. Finally, in the eventual cases where a combination of conditions would lead towards none of the above, we suggest the default should be that it is difficult to crowdsource.

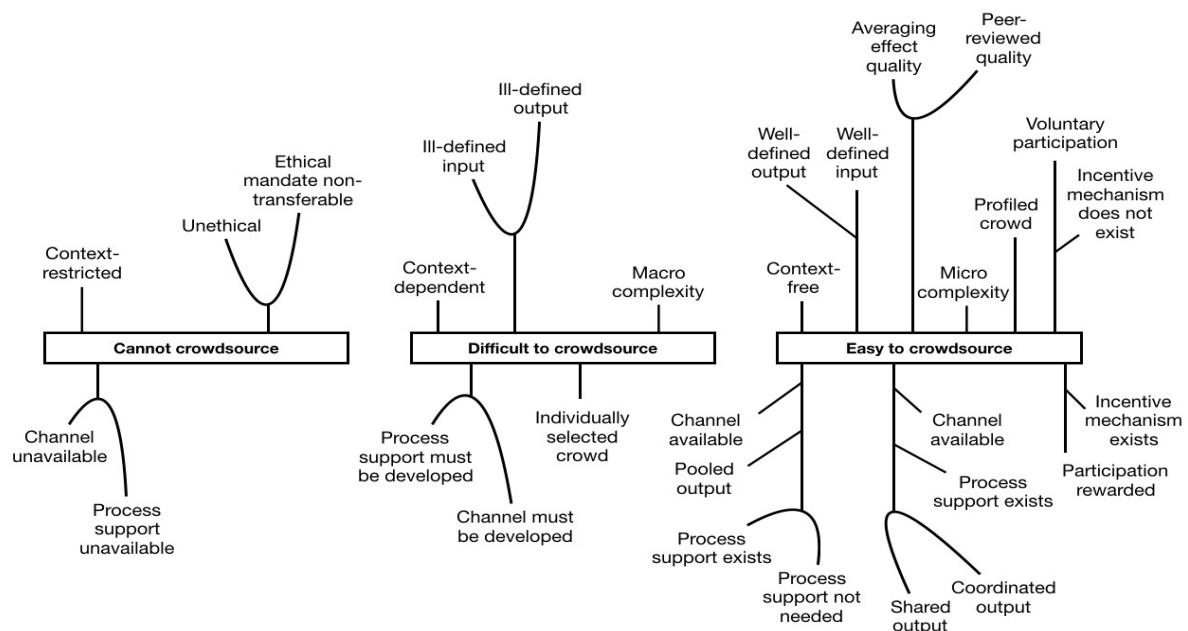


Figure 5.3

Feasibility check that determines if a function can be crowdsourced (A V-branch corresponds to a logical AND while a U-branch corresponds to an OR).

We note that the procedure shown in Figure 5.3 could have been more formally presented as a decision tree. However, because of the high number of attributes and conditions to consider, such a tree would be too large to be useful. The adopted graphical scheme seeks to summarise the various decisions that have to be considered in a more compact way, which is done at the cost of not showing every possible combination of attributes and conditions in the whole decision path. To better understand the adopted graphical scheme, we need to define the meaning of the various graphical elements used. Boxes represent the assessment, which determines if a particular activity can be crowdsourced or not, and if yes, if it is easy or difficult. V-branches represent a logical “AND”, while U-branches represent a logical “OR”. For instance, if we have well-defined outputs AND well-defined inputs, then it is easy to crowdsource. If we have ill-defined inputs OR ill-defined outputs, then it is difficult to crowdsource. Next, we use this procedure in some example cases.

5.5 JUSTIFICATION BY APPLYING THE PROCEDURE: EXAMPLE CASES

We now apply the proposed procedure to concrete examples of theory testing reported in the literature. The examples were selected to illustrate different approaches to theory testing (see figure 5.4).

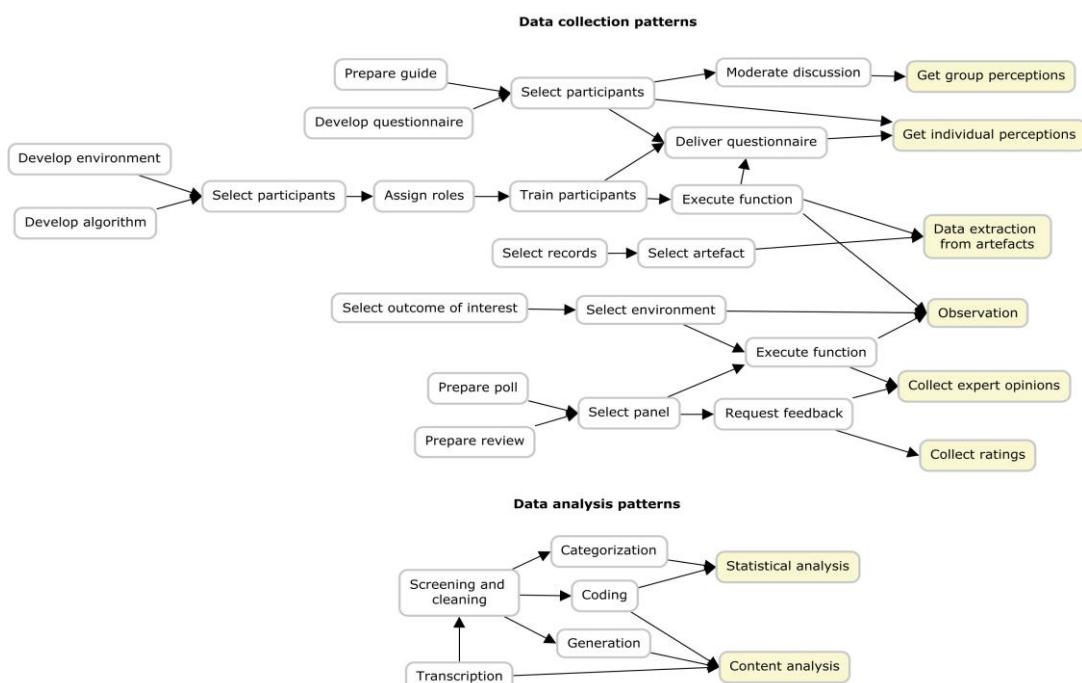


Figure 5.4 Theory testing patterns (Adapted from (Enwereuzo et al., 2017))

Example 1

This example uses an article written by Burton-Jones and Volkoff (2017), which was drawn from the pool of reviewed papers. The study adopted an inductive approach to theory building and the case study approach to theory testing (as noted in the paper's introduction). As discussed in more detail in the paper's Section 3, theory testing comprised the demonstration of the proposed theoretical constructs in a specific case dealing with a community-based healthcare system.

The authors obtained their data through focus groups, interviews, observations of system use by shadowing workers, and document analysis (Section 3.3). Matching these specific theory testing activities to the pattern system shown in Figure 5.4, we can infer which of the patterns best suit what was done. We first note that an environment was selected (a community care unit of a regional health authority in Canada). The paper does not provide details about other activities done between selecting the environment and selecting the participants. The selection of participants was very diverse, since data collection was done over a period of four years. The selection of data sources considered a wide range of criteria including experience with the system (new users versus long-term users), roles (e.g. nurses, physicians and social workers), specialities (community care and information technology staff), and organisational levels. Participants in the initial focus groups were selected according to job category to ensure that participants could speak without their supervisors being present. The selection of participants for interviews followed-up participation in focus groups. After the participant selection, the authors report they got group and individual perceptions. Furthermore, the paper has explicit references to moderating the focus groups.

Regarding document analysis, the paper provides a brief account of the types of artefacts that were selected for data extraction (e.g. business cases, terms of reference and project descriptions). Finally, the paper provides evidence that specific outcomes of interest and environments were selected for observing workers (shadowing). We can therefore identify the following data collection patterns that have been used:

- Select environment - [No other activity reported] - Select participants - Moderate discussion - Get group perceptions
- Select environment - [No other activity reported] - Select participants - Get individual perceptions
- Select records - Data extraction from artefacts
- Select outcome of interest - Select environment - Observation

Burton-Jones and Volkoff (2017) analysed data by transcribing the outputs from focus groups, interviews and shadowing, coding all data, and finally performing content analysis. We can therefore identify the two data analysis patterns that have been used:

- Transcribe - Code - Content analysis
- Code - Content analysis

Crowdsourcing was not used in this article. However, we can discuss how it could have been applied using the schema presented in Figure 5.3.

Select environment: This is a context-restricted activity because the researcher alone has to determine the environment which best suits the study. In this example, the environment was a regional health authority. Considering the nature of the activity and the research goals, we recognise that it cannot be crowdsourced, which is aligned with the results from our procedure.

Select participants: Burton-Jones and Volkoff (2017) note the selection of participants had to be carefully done. They targeted specific sites belonging to the same community care unit. Within those sites, they targeted specific types of participants, including front-line staff with access to the system (clerks and clinicians), middle managers, directors and executives. Furthermore, as noted earlier, the focus groups participants were carefully selected to avoid having supervisors and subordinates in the same group. Based on these constraints, we suggest the activity has macro complexity and therefore is difficult to crowdsource. The activity also seems context dependent, since the selection of participants takes into consideration the characteristics of the target organisation. Therefore, we regard this activity as context-dependent and therefore difficult to crowdsource. Finally, considering all requirements and constraints, it also seems that it would be difficult to profile the crowd. So, the

crowd would have to be individually selected, which again makes the activity difficult to crowdsource.

Moderate discussion: Burton-Jones and Volkoff (2017) do not provide many details about the focus groups, which suggests that a typical approach was adopted, with the researchers conducting the discussion of topics of interest (an observation supported by the statements reported in the paper, which cover a variety of issues with the system). We suggest this is a macro, context-dependent activity. It has ill-defined outputs, because of the complexity associated to moderating a group of people while adapting the script to what the participants report and what the researchers find interesting. This suggests the activity is difficult to crowdsource. However, another concern is about ethics. Reflecting on the sophistication of the topic, i.e. use of patient health records by health professionals, the diversity of stakeholders, and the complexity of data the researchers aim to collect, i.e. data about effective use, do we think it would be ethical to crowdsource such an activity? We suggest the answer should be no and consequently the activity should not be crowdsourced.

Get group perceptions: From what Burton-Jones and Volkoff (2017) report, this activity had macro complexity, was context-dependent, and had ill-defined inputs and outputs, all contributing to being difficult to crowdsource.

Get individual perceptions: Details were not given by Burton-Jones and Volkoff (2017) about the interviews, and how they were conducted, so we assume that a typical approach was adopted using semi-structured interviews. We suggest this activity would be difficult to crowdsource because it has macro complexity and is context-dependent. Burton-Jones and Volkoff (2017) highlight the interviewers had to participate in several preparation sessions to understand the work context, which supports the classification as context-dependent.

Select records: The same argument we used regarding the “select environment” activity applies here. This activity cannot be crowdsourced.

Data extraction form artefacts: Burton-Jones and Volkoff (2017) do not report any specific procedure to extract data from artefacts, which suggests they moved immediately to data analysis.

Select outcome of interest: This is a context-restricted activity because only the researchers knew the outcome of interest, which was to develop and test a theory of effective use. Considering the nature of this activity, we recognise it cannot be crowdsourced.

Observation: As reported by Burton-Jones and Volkoff (2017), observation was done by shadowing front line workers. We suggest that this activity is difficult to crowdsource because it is context-dependent: it must be done in a specific environment, shadowing health workers like nurses, physicians, etc. However, a mandate to gather data using the crowd would have to be secured by the researchers, which seems difficult to obtain in this case, as an unknown crowd would have to wander in a patient care environment. Consequently, we suggest this activity cannot be crowdsourced.

Transcribe: We suggest that this activity is easy to crowdsource because it has micro complexity, is context-free and has a clear output. As reported by Burton-Jones and Volkoff (2017), the gathered data was about the use of a new software system, which suggests the data does not raise privacy issues. Therefore, the transcription could have been crowdsourced.

Code: Burton-Jones and Volkoff (2017) report that coding was done by identifying emerging themes. This activity has micro complexity, which suggests that it could have been crowdsourced. However, the output is ill-defined, which suggests that instead it is difficult to crowdsource.

Content analysis: We suggest this activity would be difficult to crowdsource because it has an ill-defined output, which is based on the analysed data. Furthermore, it has macro complexity, since it requires a skill to interpret the data using a specific framework.

All in all, the research reported by Burton-Jones and Volkoff (2017) had 17 theory testing activities, of which six primary and eleven secondary. Of the primary activities, one could not have been crowdsourced, and five would be difficult to crowdsource. Regarding secondary activities, four could not have been crowdsourced, four would have been difficult to crowdsource, and one could have been easy to crowdsource.

Example 2

In this example we consider a study by Sanyal (2016), which was also drawn from the pool of reviewed papers. The author studied the effects of bidding strategies on the economic performance of auctions and used an experimental approach to theory testing. As discussed in detail in Section 4 of the paper, theory testing consisted in the demonstration of the proposed theoretical constructs in the laboratory using different auction designs.

Experimental data was obtained by the author in the laboratory (Section 4.2). An environment was developed for the experiment (a bidding interface). The paper mentions the participants were undergraduate university students. The data gathering procedure adopted the following pattern:

- Develop environment - Select participants - Run function - Data extraction from artefacts

In this particular case, no specific roles had to be assigned to the participants, since they were all bidders. Training was also not considered in the experiment.

Regarding data analysis, Sanyal (2016) report that some activities we identify in the pattern system were automated by the bidding interface. In particular, the interface automatically removed dead bids from the data set, which is an automated form of screening and cleaning. The interface also automatically generated data in two groups, from symmetric and asymmetric bidders, which is a form of categorisation. Considering that these two activities were automated, the data analysis pattern ended up having one single activity to consider:

- [Automated screening and cleaning] - [Automated categorisation] - Statistical analysis

Crowdsourcing was not used in this article. However, we can discuss how it could have been applied using our procedure.

Develop environment: According to Sanyal (2016), a Web site with a bidding user-interface had to be developed for the experiment. Therefore, this is a context-

dependent and macro activity because it requires specific skills to develop the software. Considering the nature of the activity, we suggest it would be difficult to crowdsource.

Select participants: This activity has micro complexity. According to Sanyal (2016), there were no criteria for selecting the participants. For that reason, this is also a context-free activity. Therefore, we suggest this activity would be easy to crowdsource.

Run function: This activity falls into micro complexity because it is repetitive and requires a single skill, which is to place bids. Being only focussed on bidding, it is also context-free. Process support would have to be developed to run the experiment. However, since we have already considered that decision element in the “develop environment”, here we have to consider that process support exists. Regarding the communication channel, the researcher used email to recruit the participants on campus, which means a communication channel is available. Sanyal (2016) noted the participants were rewarded through cash payments, which suggests an incentive mechanism would be needed. However, the same argument regarding process support applies here: since the environment has to be developed, the incentive mechanism can be integrated in the software. Therefore, we regard this activity as easy to crowdsource.

Data extraction from artefacts: Specific details were not given by Sanyal (2016) about how the data was extracted, but we can infer from the paper that the Web site stored the user data generated during the experiments. Therefore, this activity has micro complexity because it only involves retrieving stored data. It also has well-defined inputs and outputs. We suggest that this activity would be easy to crowdsource.

Statistical analysis: Sanyal (2016) compared the effects of price in jump bidding using statistical techniques. We suggest this activity has macro complexity and therefore is difficult to crowdsource.

Example 3

An article written by Van Looy et al. (2017) was selected as our last example. The study develops a set of criteria for comparing business process maturity models. The Delphi technique was used for theory testing (validation and evaluation as noted in the methodology section). As discussed in detail in the paper's Section 3.2, theory testing focused on demonstrating the applicability of the developed set of criteria.

The authors obtained their data through expert opinion collection. The process considered a preparatory phase, panel selection (participants were required to be experts), followed by the Delphi study. We can therefore identify the data collection patterns that have been used:

- Prepare poll - Select panel - Run function - Collect expert opinions

Then the authors analysed data applying grounded theory, which included coding and content analysis. We can therefore identify the analysis patterns that have been used:

- Coding - Content analysis

Crowdsourcing was not used in this article. However, we once again discuss how it could have been applied.

Prepare poll: Van Looy et al. (2017) report the study had a preparatory phase in which the criteria developed in the theory building phase were used to prepare the poll. This is a context-restricted activity because the researcher alone has the details and prior knowledge of the preceding part of the study (theory building), and therefore knows what is needed for this part of the study. Considering the nature of the activity, we recognise that it could not have been crowdsourced.

Select panel: The Delphi technique involves careful selection of panel members, who must be experts in area of study. According to Van Looy et al. (2017), the participants were selected between practitioners and academics with a genuine interest in the topic. This suggests the activity was context-free, but the crowd could have been profiled. We can therefore suggest that this activity would be easy to crowdsource.

Run function: This activity required reaching consensus about what criteria to use when comparing business process maturity models. It therefore had well-defined inputs and outputs. Van Looy et al. (2017) do not provide details on what communication and process support were used but we can infer from the description that electronic mail was used as communication medium, and no process support was required. Therefore, we suggest this activity would be easy to crowdsource.

Collect expert opinions: This activity has micro complexity, as it involves the collections of the output from a Delphi round. Since the outputs provide responses to questionnaires, the activity was also context-free. Therefore, we suggest this activity would also be easy to crowdsource.

Coding: Van Looy et al. (2017) report that, for the purposes of theory testing, coding was done by the Delphi participants in consecutive rounds, until consensus was reached about including or excluding a criterion. This was done through consecutive Delphi rounds. Considering this, we suggest this was a context-free activity with well-defined inputs and outputs. Therefore, we suggest this activity would also be easy to crowdsource.

Content analysis: Van Looy et al. (2017) report that after coding, the Delphi participants were requested to determine which criteria were more important in pairwise comparisons. We consider that this activity would be easy to crowdsource because it has well-defined inputs and outputs, and it has micro complexity, which could be performed in a repetitive way.

All in all, of the six theory testing activities that were considered in this study, one could not have been crowdsourced, while the other five would be easy to crowdsource.

5.6 DISCUSSION AND RECOMMENDATIONS

This section is structured around the artefact developed by this research and its contributions to research and practice. The artefact we have developed is the mechanism to check if a theory testing activity can be crowdsourced or not. Once again, the value of this artefact lies in its capacity to help researchers making complex design decisions, in this case regarding how they can utilise crowdsourcing to their

advantage. Gathering and analysing data to prove a theory can be a daunting endeavour. The complexity of today's world, as well as the sophistication of the research undertaken today, suggests researchers need to collect very large amounts of data pertaining to complex environments, considering a multitude of factors, contextual elements and stakeholders, which require new, innovative approaches.

The proposed mechanism helps researchers systematically checking if crowdsourcing can be applied to a wide range of theory testing activities organised as a pattern system. So, our proposition is not just to check if data collection per se can be crowdsourced or not. Our proposition is to check if every activity related to theory testing can be crowdsourced or not.

We applied the proposed mechanism to three examples illustrating very different research endeavours. In the three examples, we could make a rational recommendation regarding the decision to crowdsource or not using simple questions, which ask about specific characteristics of the research project in consideration to attributes required by crowdsourcing.

Despite the potential values described above, we should also recognise some limitations of this research. One limitation to consider is that the decision to crowdsource may extend beyond the intrinsic characteristics of theory testing activities, for instance, taking into consideration the specific characteristics of the research undertaken. That introduces a contingency factor in the decision to crowdsource that complicates the decision-making process.

Furthermore, our procedure takes into consideration a set of criteria that cannot be considered complete. Many other criteria could eventually be added, some of them addressing operational issues such as the characteristics of specific platforms used to communicate with the crowd and to perform the function. Other criteria could also be added related to a more comprehensive taxonomy of crowdsourcing functions. Another issue to consider, which is illustrated in our examples, is that the development of process support may automate some theory testing activities. However, automation was not considered in our research.

Finally, we should also consider that researchers may also find creative ways to overcome some of the constraints addressed by the procedure, e.g. adopting games, role playing, etc. Therefore, we suggest that both the pattern system and procedure described in this paper are just initial approaches to the endeavour of bringing design science into theory building and theory testing, and also bringing crowdsourcing into theory testing.

5.7 CONCLUSION

We began our study based on the output of the Problem Frame Assessment (PFA) done in previous paper (Enwereuzo et al., 2017), which indicated a need for the development of this framework to guide researchers make decisions on activities (especially theory testing activities which is our main focus) that can be crowdsourced or not. We justified the framework by applying sample cases from published literature. Results indicates that the framework is useful and usable in decision making as to theory testing activities that can be crowdsourced by considering some crowdsourcing attributes.

From a research point of view, our work extends existing research by showing that crowdsourcing can be used for different activities apart from surveys, which most of the existing literature that used crowdsourcing focused on. Through the sample cases, we were able to operationalize and demonstrate how some of the activities gotten from those papers which were not CS could be CS creating awareness and enlightenment.

Future work that can be done involves transforming this framework into a tool which will operationalise the framework and provide further support to researcher. This development will contribute to a better understanding of the framework. After this is done, the utility of the tool needs to be confirmed. Thus, the future research should provide further details about the framework regarding information, data and technical structures.

5.8 REFERENCES

- Aken, J. E. v. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of management studies*, 41(2), 219-246.
- Amazon. (2011). *Requester Best Practices Guide*. Amazon Web Services.
- Andriole, S. J. (2010). Business impact of Web 2.0 technologies. *Communications of the ACM*, 53(12), 67-79.
- Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of information technology*, 20(2), 67-87.
- Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of management review*, 14(4), 496-515.
- Bernstein, M. S., Brandt, J., Miller, R. C., & Karger, D. R. (2011). *Crowds in two seconds: Enabling realtime crowd-powered interfaces*. Paper presented at the Proceedings of the 24th annual ACM symposium on User interface software and technology.
- Brunt, L., & Meidell, E. (2018). When Are Crowdsourced Data Truthful, Accurate, and Representative. *The Journal of Business Inquiry*, 17(1), 55-71.
- Burton-Jones, A., & Volkoff, O. (2017). How can we develop contextualized theories of effective use? A demonstration in the context of community-care electronic health records. *Information Systems Research*, 28(3), 468-489.
- Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal*, 50(6), 1281-1303.
- Crowston, K. (1997). A coordination theory approach to organizational process design. *Organization Science*, 8(2), 157-175.
- Crowston, K. (2012). Amazon mechanical turk: A research tool for organizations and information systems scholars *Shaping the Future of ICT Research. Methods and Approaches* (pp. 210-221): Springer.
- Enwereuzo, I., Antunes, P., & Johnstone, D. (2017). Towards the Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: Conceptual Framework and Model. Paper presented at ACIS.
- Estelles-Arolas, E., & Gonzalez-Ladron-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189-200. doi:10.1177/0165551512437638
- Gao, H., Barbier, G., & Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems*, 26(3), 10-14.
- Geerts, S. (2009). Discovering crowdsourcing: theory, classification and directions for use. *unpublished Master of Science in Innovation Management thesis, Eindhoven University of Technology*.
- Geiger, D., Seedorf, S., Schulze, T., Nickerson, R. C., & Schader, M. (2011). *Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes*. Paper presented at the AMCIS.
- Goldkuhl, G. (2002). *Anchoring scientific abstractions—ontological and linguistic determination following socio-instrumental pragmatism*. Paper presented at the Proceedings of European Conference on Research Methods in Business, Reading.
- Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 37(2), 337-355.

- Heer, J., & Bostock, M. (2010). *Crowdsourcing graphical perception: using mechanical turk to assess visualization design*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Hevner, A., & Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol. 22): Springer Science & Business Media.
- Hosack, B., Hall, D., Paradice, D., & Courtney, J. F. (2012). A look toward the future: decision support systems research is alive and well. *Journal of the Association for Information Systems*, 13(5), 315.
- Howe, J. (2006). The rise of crowdsourcing. *Wired magazine*, 14(6), 1-4.
- Järvinen, P. (2007). Action research is similar to design science. *Quality & Quantity*, 41(1), 37-54.
- Kittur, A., Chi, E. H., & Suh, B. (2008). *Crowdsourcing user studies with Mechanical Turk*. Paper presented at the Proceedings of the SIGCHI conference on human factors in computing systems.
- Kittur, A., Smus, B., Khamkar, S., & Kraut, R. (2011). Crowdforge: Crowdsourcing complex work *Proceedings of the 24th annual ACM symposium on User interface software and technology* (pp. 43-52): ACM.
- Kleemann, F., Voß, G. G., & Rieder, K. (2008). Un (der) paid innovators: The commercial utilization of consumer work through crowdsourcing. *Science, technology & innovation studies*, 4(1), PP. 5-26.
- Malan, R., & Bredemeyer, D. (2001). Defining non-functional requirements.
- Malone, T., & Crowston, K. (1994). The Interdisciplinary Study of Coordination. *ACM Computing Surveys*, 26(1), 87-119.
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251-266.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods*, 44(1), 1-23.
- Mumford, E. (1985). Defining system requirements to meet business needs: a case study example. *The Computer Journal*, 28(2), 97-104.
- Nakatsu, R., Grossman, E., & Iacovou, C. (2014). A taxonomy of crowdsourcing based on task complexity. *Journal of Information Science*, 40(6), 823-834.
- Niederman, F., & March, S. (2015). Reflections on Replications. *AIS Transactions on Replication Research*, 1(1), 7.
- Park, S., Shoemark, P., & Morency, L.-P. (2014). *Toward crowdsourcing micro-level behavior annotations: the challenges of interface, training, and generalization*. Paper presented at the Proceedings of the 19th international conference on Intelligent User Interfaces.
- Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.
- Penker, M. (2000). Business Modeling with UML: Business Patterns at Work: New York, NY: John Wiley & Sons.
- Power, D. J. (2008). Decision support systems: a historical overview. *Handbook on Decision Support Systems* 1, 121-140.
- Pries-Heje, J., Baskerville, R., & Venable, J. R. (2008). *Strategies for Design Science Research Evaluation*. Paper presented at the ECIS.
- Sanyal, P. (2016). Characteristics and Economic Consequences of Jump Bids in Combinatorial Auctions. *Information Systems Research*, 27(2), 347-364.
- Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics & Management*(1), 93-107.

- Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS quarterly*, 37-56.
- Smith, D., Manesh, M. M. G., & Alshaikh, A. (2013). How can entrepreneurs motivate crowdsourcing participants? *Technology Innovation Management Review*, 3(2).
- Sonnenberg, C., & vom Brocke, J. (2011). *Evaluation patterns for design science research artefacts*. Paper presented at the European Design Science Symposium.
- Van Looy, A., Poels, G., & Snoeck, M. (2017). Evaluating business process maturity models. *Journal of the Association for Information Systems*, 18(6), 461.
- Venable, J., Pries-Heje, J., & Baskerville, R. (2016). FEDS: a framework for evaluation in design science research. *European Journal of Information Systems*, 25(1), 77-89.
- Von Alan, R. H., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1), 75-105.
- Vukovic, M., Lopez, M., & Laredo, J. (2010). *Peoplecloud for the globally integrated enterprise*. Paper presented at the Service-Oriented Computing. ICSOC/ServiceWave 2009 Workshops.
- Wexler, M. N. (2011). Reconfiguring the sociology of the crowd: exploring crowdsourcing. *International Journal of Sociology and Social Policy*, 31(1/2), 6-20.
- Wu, W., Tsai, W.-T., & Li, W. (2013). Creative software crowdsourcing: from components and algorithm development to project concept formations. *International Journal of Creative Computing*, 1(1), 57-91.
- Yang, J., Adamic, L. A., & Ackerman, M. S. (2008). *Crowdsourcing and knowledge sharing: strategic user behavior on taskcn*. Paper presented at the Proceedings of the 9th ACM conference on Electronic commerce.

CHAPTER 6

ANALYSING THE ADOPTION OF CROWDSOURCING IN THE WHOLE THEORY TESTING LIFECYCLE

ABSTRACT

We explore how researchers may adopt the crowdsourcing model in the whole theory testing lifecycle, which could bring a crowd to execute various research project activities ranging from data collection to management, analysis, and quality control. Testing theory is a fundamental component of theory building, which makes this a relevant research problem, in particular in the information systems field where theory building has a fundamental role in shaping the identity of the field. We adopt a design science research perspective over the problem. In particular, we develop three artefacts: a conceptual framework of theory testing, a pattern model of theory testing activities, and a crowdsourcing template to help making decision on which research activities to crowdsource. This study provides contributions to both research and practice. Regarding the former, we contribute artefacts supporting the development of new approaches to test – and ultimately develop – theory. Considering the latter, the developed artefacts offer opportunities to manage the theory testing lifecycle, which in turn may foster new research methods.

Keywords: Theory Building; Theory Testing; Theory Testing Patterns; Crowdsourcing; Design Science Research.

6.1 INTRODUCTION

Chris has been working on her PhD focusing on antecedents to online buying decisions. To test her model, she has been recruiting student subjects to her lab to perform a series of tasks. Chris' data collection is going steady but slow: so far, she has collected only 68 out of 160 required subjects. She fears she may not be able to defend in a long time. Her fears turn to disbelief when she receives an email announcement for Robin's dissertation defence two weeks out. Robin is another PhD student who started her PhD on hostile behaviour in social media in the same week as Chris. How can Robin be done already? Chris approaches Robin to find out. Robin

explains she needed behavioural data from over 240 subjects operating within a realistic social media scenario to test her model. She decided to crowdsource her study and enrolled as a requestor in Amazon's Mechanical Turk (MTurk). She developed a task that has workers go to a discussion forum, generate a specific hostile behaviour, and evaluate the feedback. She created another task for workers to go to the same forum and respond to hostile behaviour. Robin says she got a huge amount of behavioural data in a matter of days. Besides, as the initial results showed the study design needed to be tweaked, she tuned and repeated the tasks a few times until getting good quality data. After some consideration and discussion with her supervisor, Chris abandons her lab efforts and implements her tasks on MTurk. She runs the tasks without a hitch in 3 days and starts drafting her results and discussion chapters straightaway.

This example illustrates the appeal of using crowdsourcing for theory testing. Crowdsourcing is a diffusion model that relies on information technology to outsource a task (or function) to a strategically defined population of human and non-human actors in the form of an open call (Kietzmann, 2017). The task is executed remotely and in parallel mediated by technology. Based on this diffusion model, several crowdsourcing platforms like Wikipedia, MTurk, Upwork, and InnoCentive have been created to mediate the relationship between the requestor and the crowd, dealing with issues such as recruitment, selection of skills, making open calls, and compensating the crowd (Kittur et al., 2008).

The key foundations to advance scientific knowledge are testing and re-testing of theory (Niederman & March, 2015). Theory testing is concerned with establishing or refuting the validity of a theory's core propositions by determining to what degree they provide a close fit to empirical data (Colquitt & Zapata-Phelan, 2007). Theory testing thus requires collecting empirical data either confirming or disconfirming the set of propositions articulated by a theory (Lokke & Sorensen, 2014). Depending on the research area, empirical data can be collected from various types of sources and using various types of instruments. For instance, particle physicians have been gathering data from the Large Hadron Collider to develop theory about the Higgs boson, while web scientists collect data from the world wide web to develop theory on varied issues

like social networking, web marketing and trust, and collaboration researchers collect data from microworlds (Sapateiro et al., 2016).

In this paper we emphasise the challenges of doing research with human subjects. Theory testing involving human participation faces many conceptual and practical difficulties. For instance, researchers must respect ethical constraints regarding the recruitment and involvement of participants in the study (Vitak et al., 2017). Researchers must also consider the diversity of variables which influence human behaviour and thus challenge the decisions on exactly what tasks are human subjects expected to perform and what data is necessary to form robust evidence. Researchers also have to reflect on the adequacy of the selected methods, processes, instruments and tools (Bhattacherjee, 2012). Considerations about engagement, training and data quality may also result in long periods of trial and error (Peer et al., 2017; Witschey et al., 2013).

Human participation in research generates various types of empirical data, which the researcher has to analyse in order to determine whether the propositions articulated by a theory are supported or not by the data. A variety of analytic methods can be used in every field of research. A common perspective splits existing approaches into quantitative, qualitative and mixed methods (Creswell & Creswell, 2017). Some of these methods require a considerable effort to apply. For instance, the qualitative analysis of video footage may require many hours of going through raw video footage, indexing, transcribing, coding, condensing, and synthesizing the data (Derry et al., 2010). Furthermore, to increase consistency and reliability, additional quality checks and duplication of activities, such as parallel coding, may be necessary (Thomas, 2006).

In this paper, we investigate how crowdsourcing can alleviate the whole theory testing lifecycle. We note that using crowdsourcing for data collection has been around for some time. In particular, crowdsourcing is becoming a very common (if not the predominant) way to deliver questionnaires in behavioural research (Bates & Lanza, 2013; Behrend et al., 2011; Jarmolowicz et al., 2012; Peer et al., 2017). In user studies, it has also started to be used as an instrument to collect user data (Kittur et al., 2008; Stewart et al., 2017), and it is also regarded as an important component of citizen science, where it supports the distributed collection of research data (Bonney et al.,

2009; Gura, 2013). However, the relationship between crowdsourcing and the whole theory testing lifecycle has not yet been systematically explored and established.

Establishing this relationship is important from both theoretical and practical perspectives. From a theoretical perspective, finding innovative ways to research humans is paramount to rigorously test existing theories and advance new theories. The existing crowdsourcing platforms also offer relatively easy-to-implement opportunities to manage the theory testing lifecycle, which in turn may foster the adoption of new methods, as well as the combination of existing methods, which may be explored in different ways. All in all, this suggests the possibility to combine more robust with more flexible approaches to research.

The possibilities brought by crowdsourcing seem even more relevant in a research landscape where problems are becoming increasingly complex and the time available for completing doctoral studies remains constant, if not shrinking (Blagojević et al., 2017). This is especially true in the Information Systems (IS) field where theory building, and therefore theory testing, have a fundamental role in shaping the identity of the field (Gregor, 2006). Furthermore, it has been reported that participation rates in scientific studies have been declining (Galea & Tracy, 2007; Van Gelder et al., 2010). Finally, as alternative dissertation formats, such as the three-paper dissertation, are becoming increasingly popular, the need for multiple theory testing efforts is likely to increase (Fong, 2017; Jump, 2015). The crowdsourcing strategy offers a new way to incentivise humans to participate in research.

To establish the role of crowdsourcing in theory testing, we pursue a number of objectives. **The first objective is to identify the conceptual elements of theory testing.** Often discussions of theory testing are embedded in those of theory building, which makes it difficult to handle theory testing as an autonomous process. We avoid this problem by clearly identifying the core elements of theory testing, independently from theory building.

The second objective is to understand how theory testing has been conducted by researchers. An understanding of the types of theory testing activities lays a foundation for a discussion on how to crowdsource them. To this end, we analyse the adopted research methodologies and methods in IS studies that were published in the “Basket of Eight” journals appearing in the last two years. By selecting the “Basket of

Eight" we ensure drawing insights from high-quality publications while still affording a reasonable level of diversity. The selected number of years balances diversity with practical feasibility.

The third objective is to develop an approach for assessment of what theory testing activities can be crowdsourced or not. To this end, we define a procedure that confronts theory testing activities against a set of criteria required by crowdsourcing.

This research contributes to IS research by offering a set of IS artefacts that can be used by IS researchers to consider how to use crowdsourcing in theory testing. The set of developed IS artefacts includes a conceptual framework of theory testing, a pattern model of theory testing activities, and a method for making decisions on what activities to crowdsource or not. In a broader perspective, **this research also contributes to theory development in the IS discipline, by leveraging the possibilities of crowdsourcing in developing innovative ways to test theory.**

The remainder of this paper is structured as follows. In the next section, we provide some theoretical background about theory building and theory testing. In Section 3 we discuss our research approach, which is based on design science research (DSR). In Section 4 we develop our first artefact: a conceptual framework of theory testing. Section 5 uses the conceptual framework to review a body of published literature with the purpose to identify how theory has been tested in practice. This section also develops our second artefact: a pattern model of theory testing activities. In Section 6 we elaborate our third artefact: a crowdsourcing template that can be used to check which theory testing activities can be crowdsourced. In Section 7 we provide some examples on how to use these artefacts. Section 8 is dedicated to assessing the research evidence based on a qualitative evaluation of the proposed artefacts. Finally, in Sections 9 and 10 we discuss the research contributions and provide some concluding remarks.

6.2 THEORETICAL BACKGROUND

An academic discipline will not have much leverage if it does not have useful theory to contribute (Niederman & March, 2015). Building a theory is therefore considered the most prestigious endeavour of any researcher. That is very much true in the IS field, which is known to be particularly focussed on theory building (Gregor, 2002, 2006), to the point of becoming an obsession (Iivari, 2007), in a field that also promotes technology design and development.

Theory helps scientists logically describe and explain a phenomenon, process or sequence of events. (Gregor, 2002, 2006) suggests a classification into five theory types summarised in Table 6.1. Type I is the most basic type of theory. It is concerned with the “what is” question, as it tends to classify or describe characteristics of individuals, groups, situations, things, or events by summarizing the common observations in these situations. These theories are especially important when there is little, or nothing known about a phenomenon. This type of theory is therefore needed for the development of other theories (Muntermann et al., 2015). Type II theory explains “how and why” something occurs. It is used as a “sensitizing tool” to see the world in certain ways. From these viewpoints, suppositions can be drawn to scaffold theory building; and they can also be used to inform practice (Gregor, 2002). Type III theory aims to predict “what will be”. These theories can predict the outcomes, results or consequences using some explanatory factors but without necessarily explaining the connections among factors or explaining the independent and dependent variables.

Some research approaches falling into this category use statistical techniques such as correlation or regression analysis to predict but not to explain phenomena (Gregor, 2002). Type IV theory is concerned with “what is”, “how”, “why” and “what will be”. A theory in this category has the capability to provide both the prediction and the scaffolding of the underlying causes of a phenomenon using theoretical constructs. To many, this is the ideal theory as it seeks to define the concepts and propositions underlying a phenomenon, while specifying a set of fundamental relations among different variables that help with predicting the phenomenon (Gregor, 2002). Finally, type V theory is concerned with “how it is done”. It is about the methodologies,

Table 6.1 Gregor's classification of theory (Gregor, 2006)

Theory Type		Description
Type I	Analysing & Describing	Describes and classifies "what is", usually using a set of dimensions associated to a phenomenon of interest.
Type II	Understanding	Explains "how" and "why" something occurs.
Type III	Predicting	Aims at predicting "what will be". Provides testable propositions.
Type IV	Explaining & Predicting	Concerns "what is", "how", "why", and "what will be". Provides testable propositions with causal explanations.
Type V	Design and Action	Concerns "how something is done". This includes methodologies, methods and tools.

methods, principles, and guidelines that shape the construction of artificial artefacts (Gregor, 2002).

From the classification in Table 6.1, type III and type IV theories are theories with testable propositions. Goode & Hatt (1952) define a testable proposition as an imaginative idea, a guess, a statement of a solvable problem or any thinking that can be put to test to determine its validity. A testable proposition can also be an hypothesis: a tentative generalization which has to be tested to determine its validity (Khan, 2011). A testable proposition should 1) refer to observable and measurable events, 2) provide a solution to a defined problem, 3) be formulated in a way that can be tested, 4) be related to the existing body of knowledge, and 5) have logical unity (Khan, 2011).

Since our goals are centred on theory testing, it derives from this classification that our research is related to theory types III and IV. The other types of theory do not have to pass the testing challenge, since propositions only need to be either identified (types I and II) or applied (type V).

The test of a theory is concerned with establishing or refuting the validity of the theory's testable propositions. A theory is tested by determining to what degree it provides or not a close fit to empirical data (Colquitt & Zapata-Phelan, 2007). These binary outcomes are the basis for codifying and accumulating scientific knowledge, as suggested by Popper's falsifiability requirement (Popper, 2014). The continuous

theory building and testing and re-testing is considered a foundation towards establishing scientific knowledge (Niederman & March, 2015).

Researchers spend time building theory; and then spend more time testing the theory (Viswesvaran & Ones, 1995). De Vaus (2013) and Bitektine (2007) developed conceptual frameworks that describe the major activities involved in the whole research lifecycle. They include important concepts such as the definition and operationalisation of theoretical constructs, data selection, definition of hypotheses, and formulation of criteria for outcome evaluation. However, a problem with these frameworks is they combine theory testing with theory building. Since we are exclusively researching theory testing, we propose a framework for discussing theory testing as an independent phenomenon. The conceptual framework will be described in detail in Section 6.4, after discussing the research approach.

6.3 RESEARCH APPROACH

We adopt the design science research (DSR) paradigm for our study (Hevner et al., 2004). DSR supports exploratory and problem-solving research, centred on the creation of innovative IS artefacts (Gregor & Hevner, 2013). The DSR paradigm aligns three different concerns: rigour, which ensures the research is grounded on a solid knowledge base; relevance, making sure the research is focussed on delivering utility to the application environment; and artefacts which materialise the design (Hevner & Chatterjee, 2010) and substantiate the theoretical contributions in the form of principles of implementation (Gregor & Jones, 2007).

We characterise our research as exploratory, oriented towards problem solving, and targeted to research practitioners, in particular IS researchers wanting to test IS theory. These characteristics are well aligned with the DSR paradigm, which emphasises balance between technical innovation and theory development (Baskerville et al., 2018).

The research activities concern the creation of several artefacts complying with rigour and relevance. Figure 6.1 overviews the defined artefacts and their relationships to rigour and relevance. We address the creation of artefact using an iterative approach (March & Smith, 1995). An artefact is developed based on a problem frame and then evaluated/justified against it. The outcomes then contribute to adjust the problem frame and to revise the artefact until a satisfying solution is obtained (Sein et al., 2011). Furthermore, the evaluation/justification of a preceding artefact provides evidence, in the form of justificatory knowledge (Gregor & Jones, 2007), informing the creation of a consequent artefact. Therefore, the research logically articulates the creation and evaluation of various artefacts, which together may be regarded as an artefact by itself, which contains an explanation of how to realise a certain function (Kroes, 2002). By combining this generic structure with the specific context of the research, we can then delineate a more detailed research approach, which is described below.

We designed three different but interrelated artefacts: conceptual framework, pattern model, and crowdsourcing template. The conceptual framework provides a frame to understand theory testing in the wider theory generation domain, highlighting which activities pertain to each domain. The pattern model describes the different theory testing activities as a system of patterns. The pattern model is essentially the synthesised output of a descriptive literature review on how theory has been tested. Our interest in building this model is not so much on articulating the conceptual parts

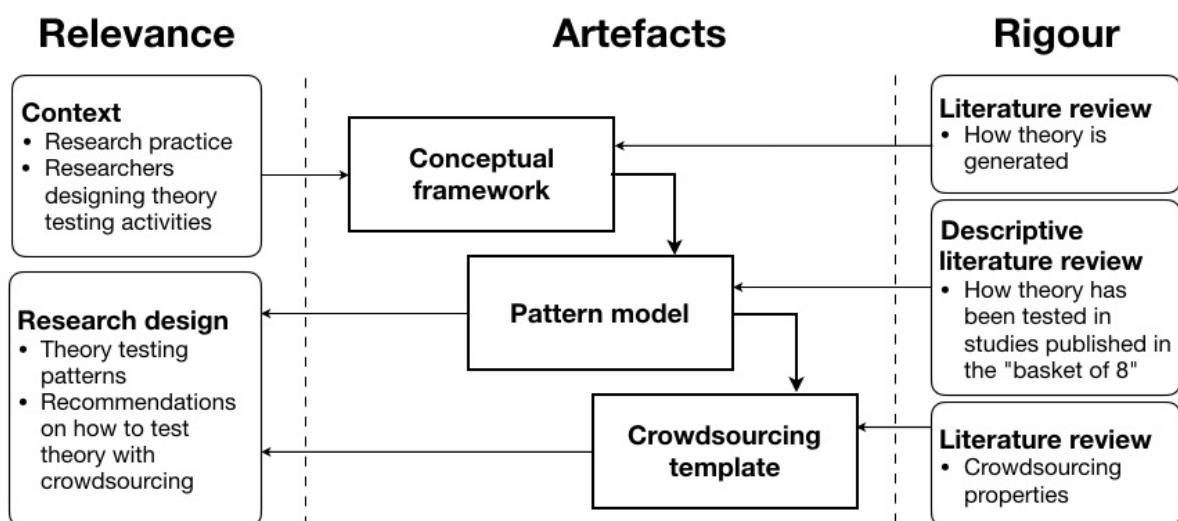


Figure 6.1 Overview of the research

of theory testing, but more on articulating the purposeful activities involved in theory testing. For that purpose, we centre the model on purposeful action. The last artefact we designed is a template to systematically check which theory testing patterns and activities can be crowdsourced using a set of criteria required by crowdsourcing.

Considering rigour in more detail, we elaborate the theory testing framework from general literature on theory building, thus providing an adequate frame from which to address the problem (Sonnenberg & vom Brocke, 2011). As there are multiple ways in which a DSR problem can be framed, an attempt to solve a DSR problem should start with a particular viewpoint (Holmström et al., 2009). In our case, we regard theory testing as the combination of preparation and action stages, where action encompasses data collection and data analysis. To rigorously build the activity model, we review the knowledge sources to extract and articulate knowledge on how theory has been tested in the IS field. This is done through a descriptive literature review of theory testing activities done in research published in the “Basket of Eight”. The outcomes of the review establish a clear link between the problem and the state-of-the-art, which is considered important to assert rigour (Gregor & Hevner, 2013; Peffers et al., 2007a). The descriptive literature review method is adopted because it helps to determine the extent to which a body of empirical knowledge supports or reveals any interpretable trends or patterns with respect to pre-existing findings (Paré et al., 2015). This is done by collecting, codifying and analysing data that reflects the frequency of relevant topics found in the literature (Paré et al., 2015). A systematic approach to the review is adopted to increase rigor and transparency (Kitchenham et al., 2009; Paré et al., 2016). We also follow the guidelines suggested by Kitchenham et al. (2009), which delineate a transparent process for synthesising the review. Finally, to rigorously build the method for assessing what theory testing activities can be crowdsourced, we analyse the crowdsourcing literature to identify the relevant properties required to apply crowdsourcing to theory testing activities. All in all, the construction of the three artefacts is anchored on three specifically targeted literature reviews.

Considering relevance in more detail, our primary goal is establishing links between our research and the research practice in the IS field. This is considered both at the beginning of the research, by specifying relevant needs, and at the end of the research, by identifying useful contributions to practice. In our case, by practitioners

we refer to IS researchers seeking to test theory. We establish close links to research practice by codifying how a relevant set of IS research has approached theory testing in the form of an activity model. We also elaborate a practical method that helps making decisions about which theory testing activities can be crowdsourced. And finally, we show how the method can be applied in practice, which is accomplished by retrospectively applying it to a set of examples.

6.3.1 Artefact Justification and Evaluation

In DSR, the assessment of artefacts can be framed according to two complementary viewpoints (Sonnenberg & vom Brocke, 2011). One is to *justify* the design of the artefacts. Another is to *evaluate* the relevance of the developed artefacts. These two viewpoints generate different types of research contributions. In the one hand, the justification of artefacts contributes to the knowledge base with prescriptions and principles of implementation, which are the foundation of design theory (Gregor & Jones, 2007). In the other hand, the evaluation of artefacts contributes to the knowledge base with expository instantiations, in the form of innovative solutions to classes of problems (Gregor & Jones, 2007).

We address both the justification and evaluation of artefacts in our research. We adopt an iterative approach to justify the artefacts, linking each artefact to the existing knowledge base, considering how the artefacts support each other in the construction of the combined artefact, and then considering how the combined artefact adds to the knowledge base (Sonnenberg & vom Brocke, 2011).

Regarding evaluation, we adopt a qualitative approach to assess the utility of combined artefact. In particular, we use the card sorting method (Prat et al., 2015), which is described next in more detail.

6.3.2 Card Sorting Method

Card sorting is a qualitative evaluation method that has been widely used in various fields such as psychology, knowledge engineering, and software engineering (Barrett & Edwards, 1995), as well as DSR (Prat et al., 2015). The method helps understanding the relationship between an artefact and the people the artefact is designed for

(Spencer & Warfel, 2004). In particular, the method can help understanding how users structure information and action mediated by the artefact (Spencer & Warfel, 2004). For instance, card sorting can be used to assess the users' needs and priorities, how they deal with information, and how they react to features provided by the artefact. Finally, through card sorting we can also gather insights on how users view the problem and the solution addressed by an artefact. This later consideration is extremely important in the DSR context, since DSR addresses wicked problems for which there are no single best problem definitions and solutions (Rittel & Webber, 1973). All in all, we use card sorting to evaluate how an artefact relates to what exists in the researchers' minds.

A card sorting exercise requires the participants to sort, group and prioritise decks of cards with pieces of information. The sorting process externalises the users' thinking and allows to analyse how they structure information (categories, priorities, relationships, groups, etc.). In this research, we use card sorting to assess artefact utility: 1) **if IS researchers find our artefacts useful in accomplishing their research goals**; and 2) **in what way the artefacts provide utility to them**.

We structured the card sorting method as a combination of card sorting exercises and interviews. All card sorting exercises were conducted in a meeting room, which had a large table where to lay out the cards. The participants were either PhD students conducting research projects in the IS field or IS faculty supervising PhD students. PhD students were selected because they represent the main target audience for this research, while IS faculty was also included because of their obvious interest and involvement in expediting PhD studies.

The participants were selected by convenience. Ten researchers participated in the evaluation. The card sorting exercises used decks of 90 cards with predefined concepts. Blank cards were also provided to the participants, so that they could add concepts if needed. The exercises were done in one-on-one sessions moderated by one of the authors (Figure 6.2). At the beginning of each individual session, the participant was informed about the purpose of the exercise, the specific artefact under evaluation, and the card sorting method. The participant was also encouraged to express her thoughts and opinions during the exercise, to talk-aloud what she was

thinking and doing, and also to consider the exercise in the scope of her research. After handing over a deck of cards to the participant, enough time was given to read through the cards. Then, the participant was asked to organise the card into categories and priorities. Along with the exercise, we gathered the participant's reactions and comments, focussing in particular on positive and negative reactions. The exercise was recorded to ease later analysis. On average, each exercise took about 30 minutes, even though some exceeded 50 minutes.

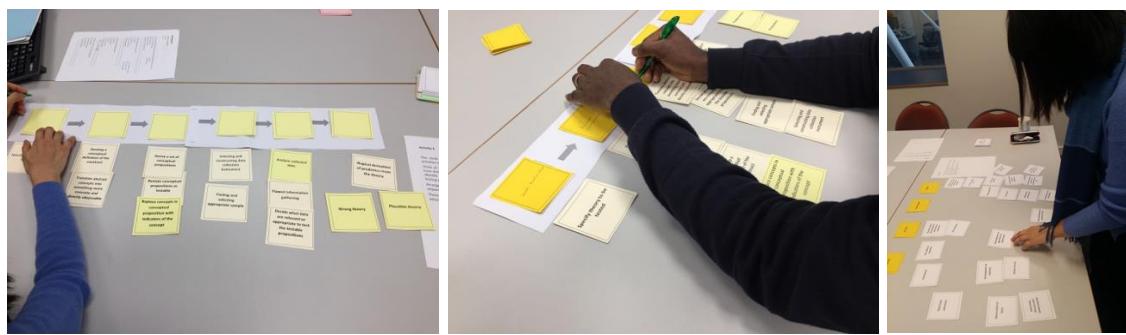


Figure 6.2 Card sorting exercises

6.4 CONCEPTUAL FRAMEWORK OF THEORY TESTING

Theory generation is an on-going process of building, observing, confirming, applying, and adapting theoretical constructs (Lynham, 2002). We regard this process as having two main activities: theory building and theory testing (Figure 6.3). Theory building, and theory testing must be carefully aligned, so that each one provides outputs reinforcing the other. That is, good conceptualisation about the world is necessary for theory testing, and good testing is required to consolidate and reinforce theory building. However, such alignment is complex, and it involves various activities related to theory building and testing, which can be designated as secondary. Within the set of secondary activities, we may consider the development of axioms, laws and propositions (Wacker, 1998), definition of domain, assumptions and constraints (Bacharach, 1989), definition of research questions and hypotheses (Agee, 2009), and definition of quantities and qualities (Gelo et al., 2008).

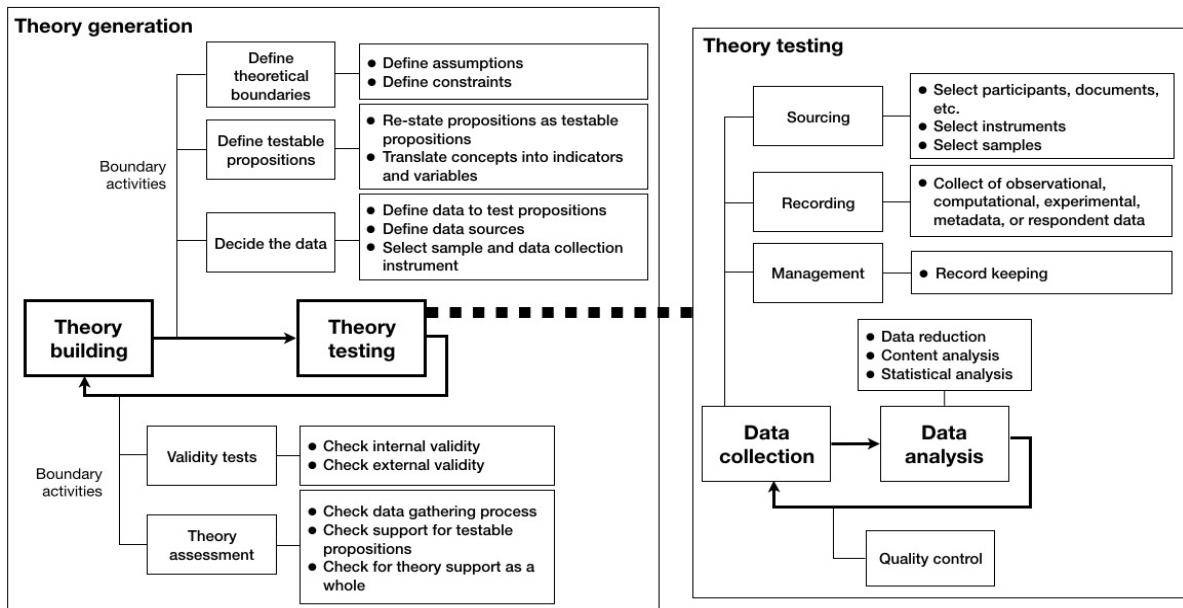


Figure 6.3 Conceptual framework of theory testing

Some secondary activities can be clearly assigned to a single primary activity, either theory building or testing. For instance, the definition of axioms and propositions clearly pertains to theory building, while the definition of a sampling plan and instrumentation clearly belong to theory testing. However, other activities exist in a grey area between both.

To address this concern, we consider the notion of boundary activity. According to Maaninen-Olsson et al. (2008), a boundary activity exists with the purpose of integrating knowledge from two different contexts. The concept was developed in the knowledge management field to understand how a community-of-practice codifies and transfers knowledge between different contexts. Here, we use the concept to discuss how certain boundary activities are necessary to integrate theory building and theory testing.

Regarding the theory generation cycle in the direction that goes from theory building to testing, we identify three important boundary activities: define theoretical boundaries, define testable propositions, and decide the data. The first boundary activity concerns setting the boundaries of the theory, identifying the value-laden assumptions taken by the researcher and the constraints to its applicability (Bacharach, 1989). The second boundary activity concerns the transformation of theoretical propositions into testable propositions. A theoretical proposition is a

conjecture about the world, which is claimed by the researcher (Wacker, 1998). On the other hand, a testable proposition is a theoretical proposition that is stated in a way that can be put to test to determine its validity (Goode & Hatt, 1952). Testable propositions make explicit which indicators and (qualitative or quantitative) variables can be used for empirically testing a conjecture about the world. Therefore, restating conceptual propositions as testable propositions, involves translating abstract concepts into observable variables, a step which is often called operationalization (De Vaus, 2013). Finally, the third boundary activity is to decide on the data, which includes defining the type of empirical data necessary to test the propositions, defining the data sources, and defining appropriate samples and collection instruments.

The direction that goes from theory testing back to theory building also involves two boundary activities. One is to check the validity of empirical data, which concerns both internal and external validity. The other is theory assessment, which involves checking the quality of the data gathering process, the support given by empirical data to each testable proposition, and the support given by empirical data to the theory as a whole, taking the set of propositions together.

We now move our attention to theory testing, increasing the level of detail. Theory testing concerns establishing to what degree a theory provides or not a close fit to empirical data, which is accomplished by determining the validity of testable propositions (Colquitt & Zapata-Phelan, 2007). Theory testing can therefore be conceptualised as a cycle with two activities: data collection and data analysis. A common perspective over data collection and analysis splits existing approaches into quantitative, qualitative and mixed methods (Creswell & Creswell, 2017). However, considering that in this research we are fundamentally interested in research activities, we adopt a more functional view over research.

Data collection is an umbrella term for a variety of secondary activities. In particular, we consider: data sourcing from participants, documents and other sources; data recording, which may include observational (e.g. counting events), computational (e.g. online processing), experimental (e.g. outputs generated from assigned tasks), metadata (collecting data about existing data), and respondent data (surveys) (Borgman et al., 2007); and management, which includes record keeping activities.

Data analysis includes three important secondary activities: data reduction (e.g. coding), content analysis and statistical analysis.

Finally, we consider one boundary activity between data collection and data analysis. This boundary activity concerns quality control of collected data, which may require adjusting data sourcing, recording and management.

Having defined a conceptual framework of theory building, it is now time to use it to investigate how theory has been tested in the IS field.

6.5 PATTERN MODEL

We now address our second research goal: to define an activity model of theory testing. To accomplish this goal, we systematically review the literature on how theory has been tested in the IS discipline. The adoption of a systematic² approach stems from the DSR principle that researchers should seek rigour by rooting the design process on the existing knowledge base (Hevner et al., 2004). Furthermore, researchers should also seek transparency by adopting well-defined and well-structured protocols (Kitchenham et al., 2009; Paré et al., 2016).

6.5.1 Literature Review

We use the “descriptive literature review” method suggested by Paré et al. (2015). This method helps to determine the extent to which a body of empirical studies supports or reveals any interpretable trends or patterns with respect to pre-existing findings (Paré et al., 2015). This is done by collecting, coding and analysing data that reflects the frequency of topics found in the literature (Kitchenham et al., 2009; Okoli & Schabram, 2010). The review procedure includes the following stages.

Select Articles. We are interested in articles that test theory to show the plausibility of such theory. A two-year period of publications in the “Basket of Eight” journals is reviewed (Table 6.2). These journals are selected because they are recognised as publishing the best research in the IS field. By covering this set, we accommodate a

²Note that being ‘systematic’ while conducting a literature review is not the same as conducting a “systematic literature review” (Pare et al. 2016).

range of research traditions in IS, while targeting the more obvious outlets for theory building.

We use a 2-year period for the review because this time frame seems sufficient to develop a comprehensive view over the topic, and also because during the review we found that we got close to a saturation point, where not much new data is collected (Fusch & Ness, 2015). The initial selection of articles considers 508 items.

Keyword selection. The next selection is obtained by keyword search, a procedure that other published reviews have also used (Niederman & March, 2019; Thuan et al., 2016). We consider articles containing any one of the following keywords: ‘theory’, ‘test’, ‘theoretical model’, ‘theoretical framework’, ‘conceptual model’, and ‘conceptual framework’. This search is done on the full body. The selection of articles based on keywords contains 506 papers, which means that only two papers are excluded from the initial set. This should not be a surprise considering that the selected outlets have a strong theoretical nature, and the selected keywords cover typical theoretical constructs.

Filter Articles. The purpose of this stage is to eliminate articles unrelated to the study. We are only interested in articles whose theory is tested. The screening procedure checks if the articles contain a theory with testable propositions (Khan, 2011; Muntermann et al., 2015), which is determined by reading the methodology section. After this stage, the pool contains 404 articles. Table 6.2 summarises the selection of articles for review.

Table 6.2. Selection of articles for review

Journal	Years	Source selection	Articles removed after keyword selection and screening	Pool of articles
MISQ	2016	29	7	22
	2017	51	15	36
JSIS	2016	19	6	13
	2017	15	5	10
JMIS	2016	42	10	32
	2017	42	6	36

JIS	2016	25	8	17
	2017	21	2	19
JAIS	2016	26	4	22
	2017	32	3	29
ISR	2016	49	11	38
	2017	46	13	33
ISJ	2016	24	5	19
	2017	23	1	22
EJIS	2016	32	4	28
	2017	32	4	28
TOTAL		508	104	404

Data extraction. This stage involves reading in detail specific sections of each article, and then coding relevant data. According to Paré et al. (2016), there are three main operational aspects to data extraction. The first one is deciding what parts of an article should be targeted. In our review, we focus on the methodology, discussion and concluding sections, as they are critical to assess the articles' theoretical contributions, where theory testing takes a central role in supporting or not the testable propositions.

Coding is the second operational aspect to consider. Coding applies the conceptual framework of theory testing (Figure 6.3) to identify theory testing activities. We are not so much interested in the conceptual parts of theory testing, but what is involved in the actual testing process. We are concerned with questions like what sources are selected, what instruments are used, what methods and techniques are used, data collection methods, and specific procedures followed during testing. We adopt a combination of open and closed codes. We use a data extraction form with predefined data fields but also open fields that allow new codes to be added if necessary. This is used to balance efficiency with openness, ensuring that the review is not confined to a small set of questions and assumptions. Figure 6.4 shows a fragment of the data extraction form.

Source	Title	Coding						Other
		Method of Data collection	Recruitment Method	Data Quality Check Method	Data Reduction Method	Content Analysis Method	Statistical Analysis Method	
MISQ 2016 starts	A TREE-BASED APPROACH FOR ADDRESSING SELFSELECTION IN IMPACT STUDIES WITH BIG DATA	Experiment * Simulation * Observation * survey	Eligibility Criteria			Performance analysis propensity score subclassification (PSS); propensity score matching (PSM)		Crowdsourcing
	AN EXPLORATORY STUDY OF THE FORMATION AND IMPACT OF ELECTRONIC SERVICE FAILURES	Survey	Invitations e-mailed to members	Used the filtering question inserted at the start of the questionnaire	Removal of ambiguous responses, coding	Content analysis (sorting, chi-square test and Crisp-Set Qualitative Comparative Analysis csQCA)	Descriptive Statistics, Post-Hoc analysis	
	ARE SOCIAL MEDIA EMANCIPATORY OR HEGEMONIC? SOCIAL EFFECTS OF MASS MEDIA DISCOURSES IN THE CASE OF THE SOPA DISCOURSE	Case study * tweets	Records from social media sites	manipulation control on task characteristics through contextual framing of survey questions	Use of keywords, description, categories, coding	Transcription, Interpretive analysis		
	CAPTURING THE COMPLEXITY OF NARRATIVE: IT USES ADAPTIVE STRUCTURATION THEORY FOR INDIVIDUALS	Survey * questionnaire	Smart phone users		Removal of responses with Missing data		PLS, Post-Hoc analysis	

Figure 6.4. Fragment of the data extraction form used in the literature review

The final operational aspect to consider is how to manage the extracted data in preparation for analysis. We use NVivo to manage the extracted data. Coding is done by the first author and checked by the other authors. To increase the reliability of the analysis (Krefting, 1991), an independent researcher codes a sample selection of articles in parallel and then the codes are compared and adjusted for clearer understanding. After data extraction, we apply the following analytical lens to the review.

6.5.2 Analytical Lens

From a DSR perspective, a pattern is a regular and intelligible form or sequence in which something is done (Alexander, 1999). It can also be seen as a generalized solution to a class of design problems (Bider et al., 2013a), which is presented in a template form (Kampffmeyer & Zschaler, 2007). According to pattern theory, professional fields such as architecture, engineering, computer science, and law face design problems for which template solutions have to be initially designed and then repeatedly used (Zigurs & Khazanchi, 2008). In such context, a pattern expresses the abstract relationship between a category of design problems and a template solution. Some vital elements associated with patterns are intent, structure and implementation (Penker, 2000). *Intent* defines the general purpose of the pattern and establishes the link between the pattern and the design problem. *Structure* defines a coherent system of knowing what to do, which is accomplished by organising a set of interrelated activities (Alexander, 1999). Finally, *implementation* considers how the pattern realises a solution to design problem.

To illustrate the operationalisation of these concepts, consider the Iterator pattern (Coplien & Schmidt, 1995). The Iterator is a generalised solution for managing

information structures (the design problem). The intent is to traverse the information structure. Information is structured as a collection of objects with bidirectional links, which allow the Iterator to move forward and backwards across the information structure. The implementation defines how the Iterator is attached to the information structure, how it moves forward and backwards, what happens when it reaches the end of the structure, etc.

We apply pattern theory to the theory testing problem. We regard intent as what gives meaning and purpose to theory testing, i.e. linking theory testing to theory building. An intent is realised by research activities, which have to be put together in a patterned way. Therefore, structure concerns both the identification of relevant activities and the arrangement of activities in a way that realises the desired theory testing intent. Implementation considers how a theory testing pattern is realized, using a set of generic rules and procedures. Finally, the combination of intent, structure and implementation defines a template solution to the theory testing problem. We note the definition of activity is relatively loose: an activity is something that has to be done by the researcher. Consequently, we may find different types of activities in a theory testing pattern. We may find activities directly involved in generating empirical data and other activities managing and supporting the generation of empirical data. Such diversity turns patterns into informal systems of knowledge (on how to conduct theory testing) rather than formalised rule systems.

Based on this reinterpretation of pattern theory in the specific context of theory testing, we now define the procedure adopted for the review:

- We start by identifying intents. An intent communicates specific theory testing meaning and purpose that link to theory building;
- We then proceed with the identification of relevant research activities realising an intent;
- The research activities are then linked in a patterned way. Links relate activities according to precedence and supply-and-demand of resources;
- Implementation details are then added to the pattern to further explain how to realize the intent.

Next, we illustrate this procedure by applying it to a concrete example.

6.5.3 Applying the Analytical Lens

This example uses an article written by Burton-Jones and Volkoff (2017), which proposes a theory of effective system use. The study adopted an inductive approach to theory building and the case study approach to theory testing. The selected case involved a community-based healthcare system.

Burton-Jones and Volkoff (2017) described theory testing as a collection of intents seeking a deep understanding of effective system use. One intent was to gather a group perception about effective system use, which was then complemented by gathering individual perceptions from selected individuals and observations of actual system use. Various data collection activities were necessary to accomplish these intents. The researchers conducted focus groups, interviews, observations, and document analysis. Furthermore, several supporting activities were also done. For instance, the researchers recruited participants from a community care unit of a regional health authority in Canada, moderated discussions in focus groups and individual interviews, selected documents for data extraction, and shadowed users.

Burton-Jones and Volkoff (2017) then analysed the collected data by transcribing the outputs from focus groups, interviews and observations, coding all data, and performing content analysis. We may therefore organise theory testing into several patterned activities:

- Select participants - Moderate discussion - Get group perception - Transcription - Coding - Content analysis
- Select participants - Moderate discussion - Get individual perceptions - Transcription - Coding - Content analysis
- Observation - Transcription - Coding - Content analysis
- Select artefacts - Select records - Data extraction - Content analysis

6.5.4 Model construction using theory testing patterns

Based on the theoretical lens explained above, we developed a model synthesising the theory testing patterns found in the review. The model is presented in Figure 5 and discussed below.

We found strong separation of concerns between data collection and data analysis, which lead to the definition of two separate categories of patterns. This separation of concerns is evidenced by the diversity of data analysis patterns that can be used with a data collection technique, and conversely the diversity of data collection patterns related to a particular data analysis technique. This separation of concerns supports the theory testing cycle defined in our conceptual framework.

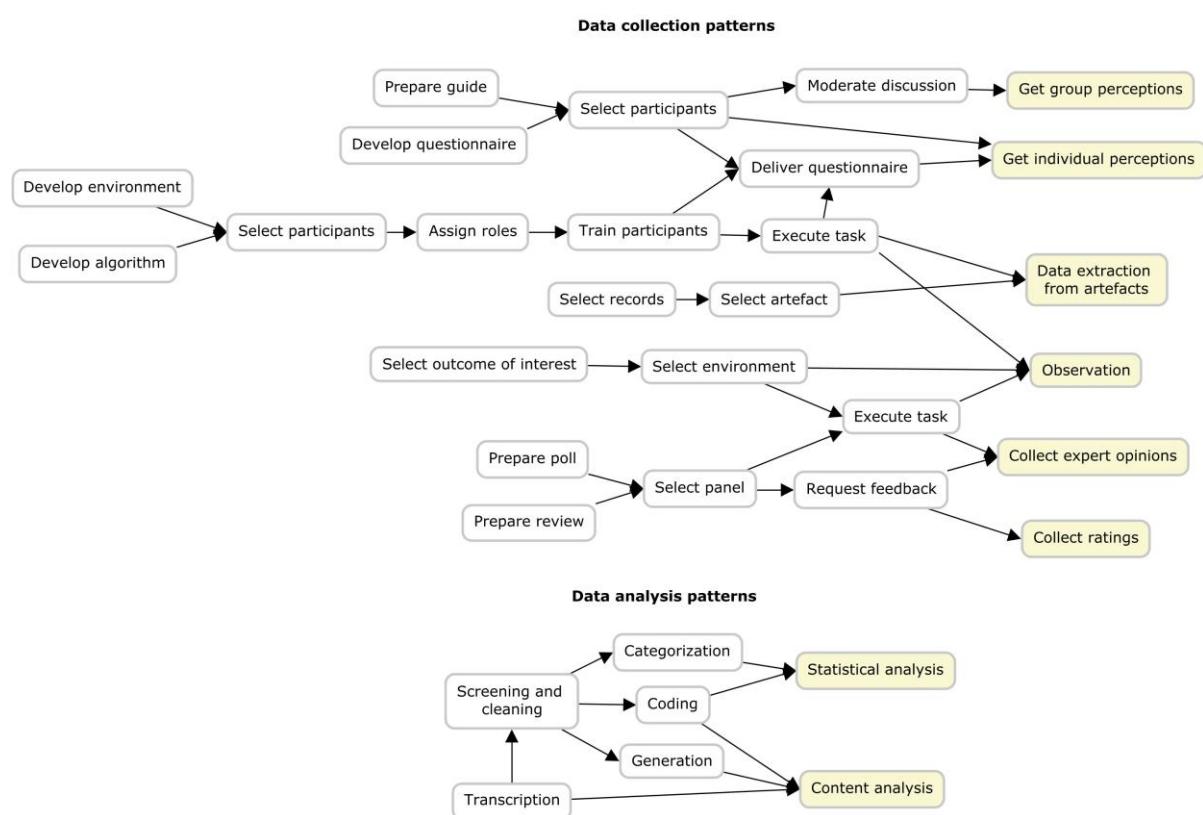


Figure 6.5. Synthesis of the literature review showing theory testing patterns

We also found a structural configuration characteristic of theory testing patterns. Pattern theory suggests that an intent is accomplished by a set of interdependent activities, which can be represented as a graph. The graph highlights input-output

dependences between activities (i.e. the output of one activity is the input of another activity). However, we found out that theory testing activities can best be represented as a tree. The top element of this tree corresponds to what we designate as a *primary activity*: an activity that fulfils a specific theory testing intent. We designate the leaves from top element as secondary activities: activities providing resources and support to primary activities. For instance, the “get group perception” pattern shown in Figure 6.5 is represented as a tree with a primary activity named “get group perception”, followed “moderate discussion”, “prepare guide” and “Select participants” secondary activities. The primary activity realises the pattern’s intent, which is to gather feedback from a group of people. The secondary activities provide resources necessary to gather feedback from the group of people, such as having a meeting guide and group moderation.

We suggest that theory testing patterns should be described by traversing from the primary activity to the secondary activities. Interestingly, by following this approach one describes theory testing activities by going backwards through the pattern. This is unlike most methodological approaches, which typically organise research activities by moving forwards (e.g. DSRM regards DSR as a sequence of activities starting with problem identification and finishing with scholarly publication (Peffers et al., 2007)).

In the review, we found six primary activities related to data collection: “get expert opinions”, “get ratings”, “get group perception”, “get individual perceptions”, “observation”, and “data extraction”. We also found two primary activities related to data analysis: “statistical analysis” and “content analysis”. Since all these patterns have a single primary activity, which realises each patterns’ intent, we name the patterns according to the primary activities.

Considering data analysis in more detail, we found in our review a myriad of approaches to data analysis, which could be classified as primary activities, e.g. cluster analysis, sociomaterial analysis, interpretive analysis, correlation analysis, econometric analysis, analysis of variance, and so forth. However, we suggest it would be very difficult to define distinct intents for each one of these. Therefore, we only identify two intents: “statistical analysis” and “content analysis”. They reflect the well-known and heavily discussed distinctions between qualitative and quantitative

analyses (Kumar, 2005). Regarding secondary activities, we also found a myriad of possibilities in the literature, which we condensed into five activities: transcription, screening and cleaning, categorization, coding, and generation. Screening and cleaning comprehend more specific activities such as filtering noise (e.g. discarding duplicates), applying selection criteria to exclude or include data (e.g. excluding non-homogeneous transactions (Lu et al., 2016), removing information that is not in English (Zhang et al., 2016)), and excluding data based on screening questions (Roberts et al., 2016). We note that coding can be applied to both qualitative and quantitative data ((Selander & Jarvenpaa, 2016) and (Li et al., 2017) are two examples), and can use various techniques, such as open and axial coding (Miles & Huberman, 1994). Finally, generation considers the generation of data from data (e.g. generating a network model from a dataset of online comments (Zhang et al., 2016)).

We emphasise that patterns are flexible in representing the dependencies between activities, which means all secondary activities are optional and not exactly a closed set. For instance, the “observation” pattern has several secondary activities such as “execute task”, “train participants” and others, but it may be instantiated as “recruit participants” and “observation”. This characteristic of patterns allows to represent different research methods in a flexible way.

We finally note that our review found ample evidence that the relationships between data collection and data analysis are manifold. For instance, there is no rationale for assuming that observational data necessarily leads to content analysis. In fact, we found examples in the literature where observational data was analysed quantitatively, qualitatively, and both (we refrain from designating the latter approach as mixed method because it brings forward the discussion on research methods, which we regard as a different concern). In fact, when synthesising our review, we attempted the exercise of linking data collection and analysis and found out examples connecting most patterns in different ways. Because of this, we do not link the two categories. We only emphasise they are logically bound through the theory building/testing model.

6.6 METHOD FOR ASSESSING WHAT THEORY TESTING ACTIVITIES CAN BE CROWDSOURCED

A crowdsourcing process can be generally described as a sequence of four steps (Thuan et al., 2016): *decision* to crowdsource, *task specification*, *task execution*, and *consolidation* of results. To be able to complete this process with success, a set of criteria must be fulfilled, which we discuss below and summarise in Figure 6.6.

Linked to the decision to crowdsource, we find the important topic of ethics. We should consider if it may be *unethical* to crowdsource. Several reasons may concur for crowdsourcing to be unethical, e.g., gathering data from inside a hospital without consent. If the strategy is accepted from an ethical perspective, we need to further consider if the ethical mandate obtained by the researcher is *transferable* or *non-transferable* to the crowd. We suggest only studies with transferable ethical mandates should be crowdsourced.

Having made the decision to crowdsource, we then focus on the task specification. A relevant criterion to consider is task complexity (Ghezzi et al., 2018; Jiang et al., 2015; Vondrick et al., 2013). A task can be classified (Antunes et al., 2012; Reason, 2008): *skills-based*, if it can be accomplished as a routine by the crowd, in an almost unconscious way; *rules-based*, if it requires following a plan, with short latitude of

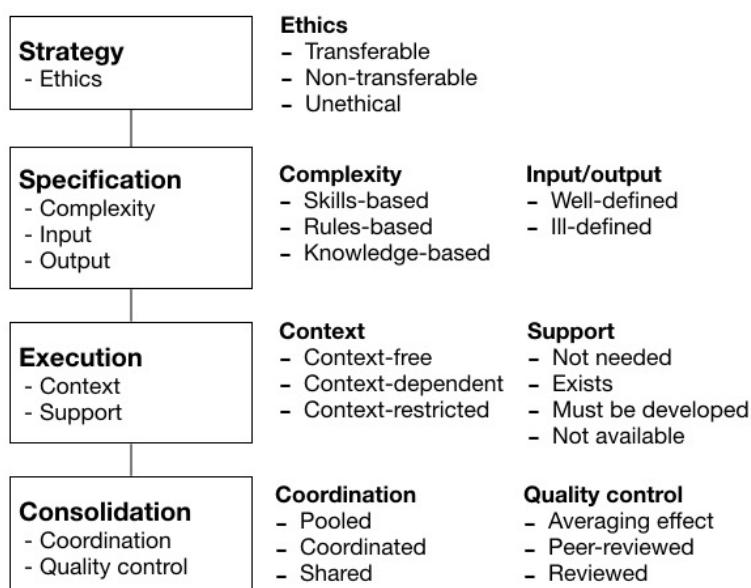


Figure 6.6. *Crowdsourcing criteria*

decision by the crowd; and *knowledge-based*, if it is novel or requires considerable problem analysis and decision-making by the crowd. We suggest that, as we move from skills-based to knowledge-based tasks, crowdsourcing becomes more difficult to achieve. This recommendation is based on research showing that the crowd prefers shorter tasks (Gadiraju et al., 2015); and also that micro-tasks enable the crowd to generate higher quality work (Cheng et al., 2015).

Another aspect to consider are the inputs and outputs required by the task. Input is the problem the crowd is asked to resolve, while output is the solution developed by the crowd, both of which can be classified as either *well-defined* (e.g. classifying photos using a predefined set of tags) or *ill-defined* (e.g. generating new research ideas) (Nakatsu et al., 2014). We contend that tasks with well-defined inputs and outputs can be more easily crowdsourced, while in the other cases they may be more difficult to crowdsource, since the crowd is more involved in defining what the task is about. This recommendation is based on research showing that the crowd prefers tasks with low associated risks (McInnis et al., 2016).

Moving on to task execution, two important criteria must be considered: context and support. Some crowdsourcing tasks are completely independent of the context in which they are executed, e.g. classifying photos according to given categories, which means they are *context-free*. Other tasks depend on a specific context to be executed, such as a physical location or an operational environment. Examples include gathering data in public libraries, driving a car, or gathering data from a social media platform. In these cases, tasks are *context-dependent*. Finally, some tasks may have to be performed in restricted contexts, which may not be easily accessible or reproducible. Examples include gathering data in private places such as military installations, coding data sets using proprietary tools, and selecting records from company databases. In these cases, we designate the tasks as *context-restricted*.

We suggest that context-free tasks are easy to crowdsource, while context-dependent tasks are difficult to crowdsource. The rationale for considering a context-dependent task difficult to crowdsource is that certain constraints have to be enforced, which make the task more difficult if not impossible to specify. In principle, we consider a context-restricted task not crowdsourceable. This is because, if it would be possible to

circumvent the restrictions, then they would be dependencies, not restrictions. Consider as an example that you plan to crowdsource data collection in a hospital. Sending the crowd to the hospital without permission is not advisable, and therefore the task is context-restricted. However, if you would get permission from the hospital, then the task would be context-dependent.

Support concerns the methods, tools and training required to perform the task. Regarding this criterion, we consider the following values: support is *not needed* when the task can be carried out by the crowd without any support (e.g. counting birds in a forest); support *exists* when training, methods and tools exist and can be used by the crowd to execute the task (e.g. a photo sharing tool is available to upload bird photos); support *must be developed*, when it is needed but must be developed by the researcher (e.g. a website must be created to classify birds in a certain way); and support is *not unavailable*, when it is needed but the researcher is either unwilling or unable to develop it. We consider that a crowdsourcing task can be easily done when support is not needed or exists, is difficult to accomplish when support must be developed, and cannot be accomplished when support is not available.

After executing the crowdsourcing task, the outputs generated by the crowd must be consolidated. Here, we have to consider what type of coordination is required to consolidate the various outputs. We consider the following values (Kevin Crowston, 2012; Malone & Crowston, 1994): the outputs may be *pooled* (e.g., individual ideas can be merged), *coordinated* (e.g., the outputs from idea generation may be fed to a selection task), or *shared* (e.g., collaborative editing). We consider that a task can be easily crowdsourced when the consolidation is pooled or coordinated but will be difficult to accomplish in cases where consolidation requires a shared approach and appropriate support must be developed. Crowdsourcing tasks cannot be accomplished when a shared approach is needed, and support is not available.

Still regarding consolidation, the researcher has also to consider the quality control of outputs generated by the crowd. Quality control may rely on the *averaging effect*, which uses multiple data sources to increase accuracy and trust (Brunt & Meidell, 2018); it can be *peer-reviewed*, where quality assurance is actually done by the crowd; and may be *reviewed* by the researcher, using either a sample or the entire data set.

We consider that, if the averaging effect or peer-reviews are viable, then a task is easy to crowdsource. The review approach has a neutral impact in relation to crowdsourcing, as it is a traditional approach to quality control.

Having discussed the crowdsourcing criteria and decision rules, we can now describe the method for assessing which theory testing activities can be crowdsourced (see Figure 6.7):

1. The method is applied at the pattern level, which means that each theory testing intent is considered separately;
2. When checking the crowdsourcing criteria, the researcher considers both the primary and secondary activities defined by the pattern corresponding to the selected intent;
3. For each patterned activity, the researcher analyses the various crowdsourcing criteria and assigns the properties that best characterise the theory testing intent and the research context;
4. For each crowdsourcing criteria, Figure 6.7 suggests four types of results: cannot crowdsource; difficult to crowdsource; easy to crowdsource; and no recommendation;
5. Figure 6.7 does not provide a specific answer to the crowdsourcing problem. It just provides suggestions derived from assigning properties to a set of theory testing activities.

We emphasise the method provides scaffolding for making an evaluative judgement on how to crowdsource a theory testing intent, and having decided to crowdsource, on which specific activities to crowdsource. A certain degree of uncertainty in the decision process is necessary because researchers have various degrees of latitude when planning their research: choosing alternative intents, patterns and activities; and choosing different levels of task complexity, types of inputs and outputs, and other

Crowdsourcing criteria			Pattern			Cannot crowdsource	Difficult to crowdsource	Easy to crowdsource
Process stage	Criteria	Properties	Activity 1	Activity 2	Activity n			
Strategy	Ethics	Transferable						X
		Non-transferable				X		
		Unethical				X		
Specification	Complexity	Skills-based						X
		Rules-based					X (1)	X (1)
		Knowledge-based				X (1)	X (1)	
	Input	Well-defined						X (2)
		Ill-defined					X	
	Output	Well-defined						X (2)
		Ill-defined					X	
Execution	Context	Free						X
		Dependent					X	
		Restricted				X		
	Support	Exist						X
		Not needed						X
		Must be developed					X	
		Unavailable				X		
Consolidation	Coordination	Pooled						X
		Coordinated						X
		Shared				X (3)	X (3)	
	Quality control	Peer-reviewed						X
		Averaging-effect						X
			Notes:			1. Decision depends on required rules or knowledge. 2. Both input and output must be well-defined. 3. Depends if support is unavailable or not.		

Figure 6.7. Template for assessing if a pattern can be crowdsourced

elements such as context, support, coordination, and quality control. In the next section we illustrate the method using some examples.

6.7 EXAMPLES

The following examples were selected to illustrate different approaches to theory testing. Crowdsourcing was not used in the selected examples. Therefore, we discuss how it could have been applied using the proposed method.

Example 1

This example uses the article written by Burton-Jones and Volkoff (2017) which was already discussed in Section 6.4.3. We apply the method to the “get group perception pattern”, which included the following activities: select participants, prepare guide, moderate discussion, and get group perception.

Strategy. Excepting “develop guide”, all other activities would require an ethical mandate to be transferred to the crowd in case they would be crowdsourced. This would be required to protect the healthcare organisation and its patients from privacy breaches.

Specification. Burton-Jones and Volkoff (2017) mentioned they had to participate in preparation sessions to understand the work context, which suggests that considerable preparation was required to conduct the focus groups. Therefore, the “moderate discussion” activity should be classified as knowledge-based and would be difficult to crowdsource.

Execution. Burton-Jones and Volkoff (2017) mentioned the focus groups involved discussions about system use, but no actual system use. Therefore, the execution context of the “moderate discussion” and “get group perception” activities was context-free and could have been crowdsourced.

Consolidation. Burton-Jones and Volkoff (2017) noted they aggregated data from several focus groups. We suggest the results could have been pooled, and therefore would be easy to crowdsource.

In summary, the method highlights that the main constraints to crowdsource this pattern was task complexity, which was knowledge-based. To crowdsource this pattern, the researcher would have to consider ways to turn the focus group into a rules-based task, e.g. by standardising the feedback generated by the participants around a set of predefined problems with system use.

Example 2

In this example we consider a study by Sanyal (2016), which studied the effects of bidding strategies on the economic performance of auctions. The study used an experimental approach to theory testing. Theory testing consisted in repeated experiments in a laboratory setting using different auction designs. Experimental data was obtained using an environment specifically developed for the experiment: a bidding interface. Sanyal (2016) noted the participants were undergraduate university students. The study adopted the “data extraction” pattern, which included the following

activities: “develop environment”, “develop task”, “execute task”, and “data extraction”. In the experiment, no specific roles were assigned to the participants, since they were all bidders. Training was also not necessary.

Strategy. Considering the artificial bidding environment that was developed, the ethical mandate to execute the experiment could be transferred to the crowd in case they were crowdsourced. Therefore, the pattern would be easy to crowdsource.

Specification. The “develop environment” activity can be classified as knowledge-based and therefore would be difficult to crowdsource. However, it has well-defined inputs and outputs, which would make it easy to crowdsource (e.g. through a competition).

We regard “select participants” as skills-based because, according to Sanyal (2016), there were no criteria for selecting the participants. The “execute task” activity can also be classified as skills-based because it was repetitive and required a single skill, which was to place bids. Specific details were not given by Sanyal (2016) about how the experimental data was extracted, but we can infer from the article that it consisted of automated usage logs. This activity can be classified as skills-based. Therefore, we suggest both “execute task” and “data extraction” would be easy to crowdsource.

“Select participants”, “execute task” and “data extraction” had clear or well-defined inputs and outputs, which makes them easy to crowdsource.

Execution. The “execute task” activity, being only focussed on bidding, was context-free. Support would have to be developed to run the experiment. However, since we have already considered that decision element in the “develop environment”, we accept that additional support would not be required. Therefore, we regard this activity as easy to crowdsource. The same argument applies to “data extraction”.

Consolidation. We can infer from Sanyal (2016) that “data extraction” was pooled, meaning individual outputs were collected and made ready for analysis, which suggests that this activity would be easy to crowdsource.

Balancing all criteria, the method suggests this pattern would be easy to crowdsource.

6.8 EVALUATION

Several card sorting exercises were set up to assess the utility of the three artefacts developed by this study

6.8.1 Utility of the conceptual framework of theory testing.

Evaluation elements used to assess the utility of the conceptual framework:

- Feedback about ease of understanding of framework concepts;
- Correct/incorrect placement of research activities in the categories defined by the framework (theory testing, theory building and boundary activities).

For this exercise, the participants were handed a deck of cards with various theory building, theory testing, and boundary activities. The participants were then told to place the cards into the correct categories while talking-aloud what they were thinking.

The exercise showed that participants with different research backgrounds performed differently. Participants working with quantitative methods could place both theory building and theory testing cards appropriately, and only had uncertainty placing boundary activities. Participants working on grounded research had more difficulties placing the theory testing cards, noting they usually collect data to build or develop theory, and not the other way around. Participants with an engineering background also had difficulties placing the theory building cards, commenting they had less experience with theory development.

All participants were uncertain about placing boundary activities, commenting that these activities could “go both ways”. This uncertainty provides further support for the notion they are best classified as boundary activities, nestled between the building and testing domains.

Overall, the exercise confirmed the usefulness of the framework in characterising theory testing for research in different backgrounds.

6.8.2 Utility of the pattern model of theory testing activities.

Evaluation elements used to assess the utility of the pattern model:

- Feedback about ease of understanding of pattern model;
- Feedback about if/how the participants find the pattern model useful to conceptualise their own research projects.

The model was presented to the participants, and instructions were given on how to use it. In a first card sorting exercise, the participants were told to go through the represented activities and pick the ones they found difficult to understand. We found that more than half of the participants fully understood the activities, but some confusion clearly remained. One participant picked out the “execute task” card, and said he did not fully understand it, as the word “execute” sounded like a computer programming phrase. Another participant picked the card labelled “assign roles”, indicating she had never used it and did not fully grasp what it meant. In general, where the participants were unclear about the specific labels appearing in the cards, their responses largely suggested a lack of familiarity with the underlying research methods.

In a follow-up to the card sorting exercise, the participants were told to analyse the patterns and classify how easy they were to understand in a five-point scale: very easy, somewhat easy, neutral, somewhat difficult, and very difficult. 30% of the participants said the model was very easy to understand, while 60% said it was somewhat easy, and 10% said it was very difficult. We gathered further feedback from the participants that said the model was very difficult to understand to know the reasons behind their answer. The main reason was related to how patterns had to be analysed back to forth (more precisely, right to left). When asked to suggest how the model could be repositioned for better understanding, the participants tried different styles but concluded it was better off how it was.

In a different card sorting exercise, the participants were told to consider their specific research project and map it using the model. Blank cards were given so the participants could write down any activities not represented in the model. Then, the participants were asked to rank the model using a five-point scale: very useful,

somewhat useful, somewhat not useful, and not at all useful. 40% considered the model very useful to their research, while 50% said it was somewhat useful, and 10% indicated it was not at all useful (Figure 6.8).

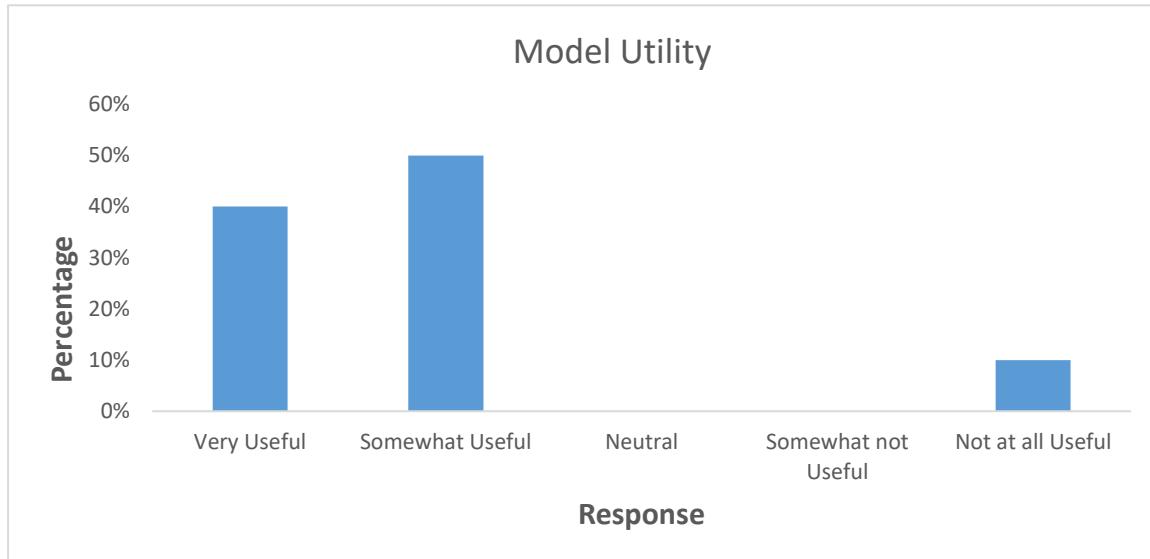


Figure 6.8. Card sorting exercise result showing Model Utility

Some participants were thrilled about the model, as it was pulled from top journals in the IS field, giving them an insight of what is expected from them if they aspire to get their articles accepted in such journals. One participant, a relatively inexperienced researcher, said the model was very useful for practitioners wanting to carry out some form of testing, as it gave them an idea about different intents, patterns and activities. The more experienced participants thought the model could benefit new researchers and felt comfortable recommending it to others. They noted that the exercise encouraged the participants, experienced or not, to think carefully about theory testing, which was seen as a positive outcome.

6.8.3 Utility of the method for assessing what theory testing activities can be crowdsourced.

Evaluation elements used to assess the utility of the method:

- Feedback about applying the method to the participants' research projects.

The final card sorting exercise was aimed at determining how the participants found the method useful to analyse how their research project could benefit from crowdsourcing. The participants picked cards with some of the activities they were considering in their projects and used them to analyse how or if those activities could be crowdsourced.

All participants were “wowed” by this exercise and felt enlightened. A participant after the exercise said she had already spent over three months on an aspect of her research, which if she had thought of crowdsourcing, or knew she could actually crowdsource, would have saved her a lot of time. Overall, 40% said the method was very useful, while 60% said it was somewhat useful (Figure 6.9).

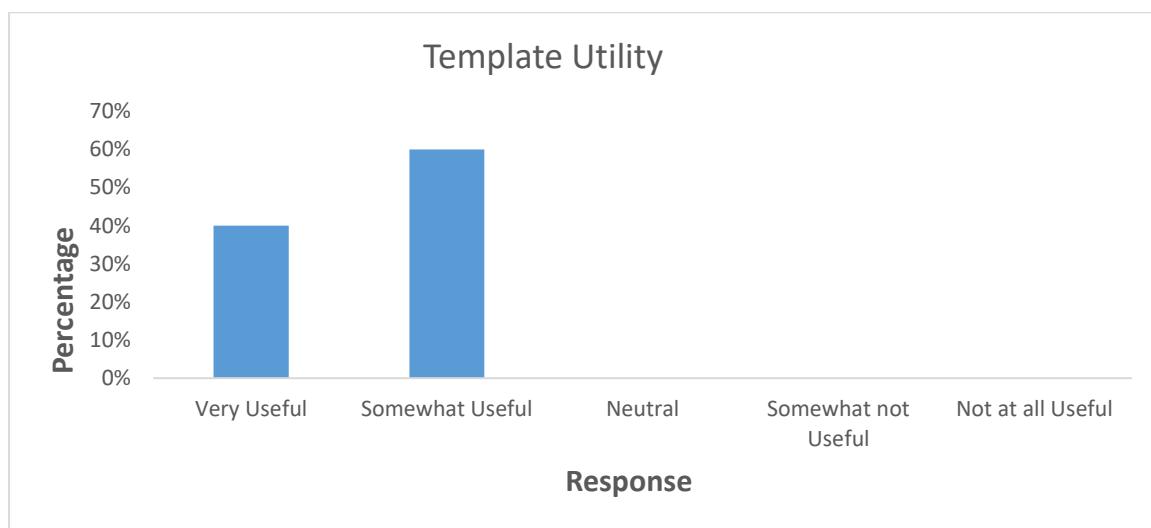


Figure 6.9. Card sorting exercise result showing Template Usability

6.9 DISCUSSION

This section is structured around the three artefacts developed by this research and the different contributions they bring to research and practice. The first artefact we developed was the conceptual framework of theory testing. This framework was developed because we had to define the perimeter of the research and the core conceptual elements of the phenomenon of interest. However, since a fundamental goal of DSR is to contribute with artefact knowledge (Gregor & Hevner, 2013), the framework also provides some interesting design contributions to research and

practice. In particular, it provides principles and technological rules that can be repeatedly used in multiple applications.

Our first undertaking consisted in delimiting the line that separates theory building from theory testing, which resulted in the two related cycles, the first one encompassing theory building and theory testing, and the second cycle disentangling theory testing into data collection and data analysis. We suggest this separation contributes to develop an operational view of theory testing as an independent phenomenon. Furthermore, the model also regards research as a recurrent activity, which may benefit from the adoption of operational principles and architectures applied to research as a system that can be designed.

Of course, research itself has for long been researched, e.g. in the fields of philosophy of science, epistemology, ontology, and methodology. The novel viewpoint proposed in this study is to articulate theory testing through design theory. Such viewpoint clearly distinguishes two different components, theory building and theory testing, as well as their boundaries and relationships. Furthermore, it also identifies a set of boundary activities that have to be considered by the researcher when aligning theory building and testing. This artefact was proved useful to survey the literature on different approaches to theory testing.

The second artefact developed by this study is a model synthesising a set of theory testing approaches. The model was developed from a review of 404 articles published in the most prestigious journals in the IS field. The model distils the different theory testing approaches into patterns, which have been divided into two groups considering data gathering and data analysis, and two other groups considering primary and secondary activities.

The model contributes to understand theory testing as a pattern system, which tries to bring together order and flexibility. Often studies combine multiple research methods and theory testing activities, and the proposed model disentangles such activities into recognisable patterns. The identified patterns cannot be understood as prescriptive. Instead, they should be understood as a flexible mechanism to articulate the operational characteristics of theory testing. Once again, the value of this artefact is strongly related to the fundamental goals of DSR: to contribute with design knowledge

which helps practitioners (in this case, researchers) to utilise principles and technological rules in their endeavours (in this case, theory building). Pattern systems reveal best practices in design. They offer a multitude of patterns that can be adopted according to different constraints and intents. Furthermore, patterns systems are conceptually open to further additions and developments. In the case of this study, such openness is particularly important since the collection of articles that were considered is just a sample of the research that has been done in the past, which is not representative of the whole body of research being undertaken today, neither expected to be undertaken in the future.

Such openness allows us to propose the pattern model as representative of successful approaches to theory testing, however accompanied with a warning that it is not intended to be seen as complete or prescriptive. The pattern model elucidates what best practices have been adopted, while at the same time providing a certain coherence about theory testing which emphasises an operational viewpoint. By adopting this pattern model, researchers may benefit from stronger links to best practices, as well as stronger transparency and coherence reporting their research. By expanding or contradicting the model, researchers may also benefit from a frame of reference against which they can claim the novelty and strengths of their research approaches.

Our final artefact is derived from a set of crowdsourcing criteria. The considered criteria offer a characterisation of crowdsourcing that is independent of specific platforms and approaches, while at the same time identifying critical elements necessary to make the decision to crowdsource, and how to crowdsource. Criteria such as complexity, support and coordination raise specific questions about research projects, which researchers must address to determine if and how theory testing activities can be crowdsourced. Once again, the adopted method provides a trade-off between the details necessary to make an informed decision and flexibility about what specific crowdsourcing technology to use. The method approaches the crowdsourcing criteria in a practical way, which relates the criteria with theory testing patterns and activities. Once again, the value of this artefact lies in its capacity to help researchers making complex decisions, in this case regarding how they can utilise crowdsourcing to their advantage.

6.9.1 Limitations of this Study

Despite the potential values described above, we should also recognise some limitations of this study. One limitation to consider is that the decision to crowdsource may extend beyond the intrinsic characteristics of crowdsourcing and theory testing activities. For instance, researchers may wish to take into consideration other contextual elements influencing their projects. That introduces a contingency factor in the decision to crowdsource that complicates further the decision-making process.

Furthermore, the proposed method takes into consideration a set of crowdsourcing criteria that may not be considered complete. Many other criteria could eventually be added, some of them addressing low-level issues such as the characteristics of specific crowdsourcing platforms. Other criteria could consider specific characteristics of the crowd, a taxonomy of crowdsourcing tasks, and different types of support provided by crowdsourcing platforms.

Finally, we should also consider that researchers may also find creative ways to overcome some of the constraints we have considered, e.g. adopting games, role playing, etc. Therefore, we suggest that the proposed method is just an initial approach to the endeavour of bringing DSR into theory building and testing, and also bringing crowdsourcing into theory testing.

6.10 CONCLUSIONS

Testing a theory can be a daunting endeavour. The complexity of today's world, as well as the sophistication of the research undertaken today, suggests researchers need to collect very large amounts of empirical data pertaining to complex environments, considering a multitude of factors, contextual elements and participants, which require new and innovative approaches. The provided collection of artefacts helps researchers systematically addressing different aspects of such endeavour. Our proposition is not just to check if data collection per se can be crowdsourced or not. Our proposition is to check which elements of the whole research project can be crowdsourced.

The combination of framework, model and method provides constructs and methods, and design principles and rules, that can be adopted by researchers in a variety of projects. Therefore, we suggest the collection of artefact contributes with operational knowledge on how to conduct research. In this case, with the added value of highlighting how researchers may benefit from the crowdsourcing strategy.

6.11 REFERENCES

- Agee, J. (2009). Developing qualitative research questions: a reflective process. *International Journal of Qualitative Studies in Education*, 22(4), 431-447.
- Aken, J. E. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of management studies*, 41(2), 219-246.
- Alexander, C. (1999). The origins of pattern theory: The future of the theory, and the generation of a living world. *IEEE software*, 16(5), 71-82.
- Amazon. (2011). *Requester Best Practices Guide*. Amazon Web Services.
- Andriole, S. J. (2010). Business impact of Web 2.0 technologies. *Communications of the ACM*, 53(12), 67-79.
- Antunes, P., Herskovic, V., Ochoa, S., & Pino, J. (2012). Structuring Dimensions for Collaborative Systems Evaluation. *ACM Computing Surveys*, 44(2), 1-28.
- Argyris, C. (1996). Actionable knowledge: Design causality in the service of consequential theory. *The Journal of Applied Behavioral Science*, 32(4), 390-406.
- Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of information technology*, 20(2), 67-87.
- Arnott, D., & Pervan, G. (2012). Design science in decision support systems research: An assessment using the Hevner, March, Park, and Ram Guidelines. *Journal of the Association for Information Systems*, 13(11), 923.
- Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of management review*, 14(4), 496-515.
- Barrett, A., & Edwards, J. (1995). Knowledge Elicitation and Knowledge Representation in a Large Domain with Multiple Experts. *Expert Systems with Applications*, 8(1), 169-176.
- Baskerville, R., Baiyere, A., , S., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems*, 19(5), 358-376.
- Bates, J., & Lanza, B. (2013). Conducting psychology student research via the Mechanical Turk crowdsourcing service. *North American Journal of Psychology*, 15(2), 385.
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior research methods*, 43(3), 800.
- Bernstein, M. S., Brandt, J., Miller, R. C., & Karger, D. R. (2011). *Crowds in two seconds: Enabling realtime crowd-powered interfaces*. Paper presented at the

- Proceedings of the 24th annual ACM symposium on User interface software and technology.
- Bhattacherjee, A. (2012). Social science research: principles, methods, and practices. *Textbooks Collection, Book 3*(http://scholarcommons.usf.edu/oa_textbooks/3).
- Bichler, M., Frank, U., Avison, D., Malaurent, J., Fettke, P., Hovorka, D., . . . Thalheim, B. (2016). Theories in business and information systems engineering. *Business & Information Systems Engineering*, 58(4), 291-319.
- Bider, I., Johannesson, P., & Perjons, E. (2013). Design science research as movement between individual and generic situation-problem-solution spaces. *Designing Organizational Systems* (pp. 35-61). Berlin, Heidelberg: Springer.
- Biemann, C. (2013). Creating a system for lexical substitutions from scratch using crowdsourcing. *Language Resources and Evaluation*, 47(1), 97-122.
- Bitektine, A. (2007). Prospective case study design: Qualitative method for deductive theory testing. *Organizational Research Methods*, 11(1), 160-180.
- Bitektine, A. (2008). Prospective case study design: qualitative method for deductive theory testing. *Organizational Research Methods*, 11(1), 160-180.
- Blagojević, V., Bojić, D., Bojović, M., Cvetačić, M., Đorđević, J., Đurđević, Đ., . . . Milutinović, V. (2017). Chapter One-A Systematic Approach to Generation of New Ideas for PhD Research in Computing. *Advances in Computers*, 104, 1-31.
- Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11), 977-984.
- Borgman, C., Wallis, J., & Enyedy, N. (2007). Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries. *International Journal on Digital Libraries*, 7(1-2), 17-30.
- Brabham, D. C. (2008). Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence: The International Journal of Research into New Media Technologies*, 14(1), 75-90. doi:10.1177/1354856507084420
- Brabham, D. C. (2008b). Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application. *First monday*, 13(6).
- Brabham, D. C. (2009). Crowdsourcing the public participation process for planning projects. *Planning Theory*, 8(3), 242-262.
- Brabham, D. C. (2010). Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. *Information, Communication & Society*, 13(8), 1122-1145.
- Brunt, L., & Meidell, E. (2018). When Are Crowdsourced Data Truthful, Accurate, and Representative. *The Journal of Business Inquiry*, 17(1), 55-71.
- Burton-Jones, A., & Volkoff, O. (2017). How can we develop contextualized theories of effective use? A demonstration in the context of community-care electronic health records. *Information Systems Research*, 28(3), 468-489.
- Campbell, J. (1990). The role of theory in industrial and organizational psychology. In Dunnette MD, Hough LM (Eds.), *Handbook of industrial and organizational psychology*, Vol. 1(2nd ed.), pp. 39-73.
- Cheng, J., Teevan, J., Iqbal, S., & Bernstein, M. (2015). Break it down: A comparison of macro-and microtasks *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 4061-4064): ACM.

- Chernova, S., DePalma, N., Morant, E., & Breazeal, C. (2011). *Crowdsourcing human-robot interaction: Application from virtual to physical worlds*. Paper presented at the RO-MAN, 2011 IEEE.
- Cleven Anne, Gubler, P., & Hüner, K. M. (2009). *Design alternatives for the evaluation of design science research artifacts*. Paper presented at the Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology.
- Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of management journal*, 50(6), 1281-1303.
- Conley, C., & Tost-Kharas, J. (2014). Crowdsourcing content analysis for managerial research. *Management Decision*, 52(4), 675-688.
- Coplien, J., & Schmidt, D. (1995). *Pattern Language of Program Design*. ACM Press/Addison-Wesley Publishing Co..
- Creswell, J., & Creswell, J. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage.
- Crowston, K. (1997). A coordination theory approach to organizational process design. *Organization Science*, 8(2), 157-175.
- Crowston, K. (2012). Amazon mechanical turk: A research tool for organizations and information systems scholars *Shaping the Future of ICT Research. Methods and Approaches* (pp. 210-221): Springer.
- Davis Jr, F. D. (1986). *A technology acceptance model for empirically testing new end-user information systems: Theory and results*. Massachusetts Institute of Technology.
- De Boer, V., Hildebrand, M., Aroyo, L., De Leenheer, P., Dijkshoorn, C., Tesfa, B., & Schreiber, G. (2012). *Nichesourcing: Harnessing the power of crowds of experts*. Paper presented at the International Conference on Knowledge Engineering and Knowledge Management.
- De Vaus, D. (2013). *Surveys in social research*: Routledge.
- Denyer, D., & Tranfield, D. J. M. D. (2006). Using qualitative research synthesis to build an actionable knowledge base. 44(2), 213-227.
- Derry, S., Pea, R., Barron, B., Engle, R., Erickson, F., Goldman, R., . . . Sherin, B. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *The Journal of the Learning Sciences*, 19(1), 3-53.
- Dicks, L. V., Walsh, J. C., & Sutherland, W. J. (2014). Organising evidence for environmental management decisions: a '4S'hierarchy. *Trends in ecology & evolution*, 29(11), 607-613.
- Dubin, R. (1976). Theory building in applied areas. *Handbook of industrial and organizational psychology*, 17, 39.
- Dubin, R. (1978). Theory building (Rev. ed.). *New Symptoms of depression in adolescents with epi*.
- Enwereuzo, I., Antunes, P., & Johnstone, D. (2017). Towards the Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: Conceptual Framework and Model. Paper presented at ACIS.
- Enwereuzo, I., Antunes, P., & Johnstone, D. (2018). On the adoption of crowdsourcing for theory testing. In *ECIS* (p. 179).
- Estelles-Arolas, E., & Gonzalez-Ladron-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189-200. doi:10.1177/0165551512437638

- Eveleigh, A. (2014). Crowding out the archivist? Locating crowdsourcing within the broader landscape of participatory archives. *Crowdsourcing our Cultural Heritage*, 211-212.
- Filatova, E. (2012). *Irony and Sarcasm: Corpus Generation and Analysis Using Crowdsourcing*. Paper presented at the LREC.
- Følstad, A. (2008). Living labs for innovation and development of information and communication technology: A literature review. *eJOV: The Electronic Journal for Virtual Organization & Networks*, 10.
- Fong, B. (2017). An exploration of changing dissertation reqs and library services to support them. *Libraries and the academy*, 17(1), 129-144.
- Ford, R. C., Richard, B., & Ciuchta, M. P. (2015). Crowdsourcing: A new way of employing non-employees? *Business Horizons*, 58(4), 377-388.
- Fusch, P., & Ness, L. (2015). Are we there yet? Data saturation in qualitative research. *The qualitative report*, 20(9), 1408.
- Gadiraju, U., Fetahu, B., & Kawase, R. (2015). Training workers for improving performance in crowdsourcing microtasks *Design for Teaching and Learning in a Networked World* (pp. 100-114). Cham: Springer.
- Galea, S., & Tracy, M. (2007). Participation rates in epidemiologic studies. *Annals of epidemiology*, 17(9), 643-653.
- Gao, H., Barbier, G., & Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems*, 26(3), 10-14.
- Geerts, S. (2009). Discovering crowdsourcing: theory, classification and directions for use. *unpublished Master of Science in Innovation Management thesis, Eindhoven University of Technology*.
- Geiger, D., Seedorf, S., Schulze, T., Nickerson, R. C., & Schader, M. (2011). *Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes*. Paper presented at the AMCIS.
- Gelo, O., Braakmann, D., & Benetka, G. (2008). Quantitative and qualitative research: Beyond the debate. *Integrative Psychological and Behavioral Science*, 42(3), 266-290.
- Gerring, J. (2006). *Case study research: Principles and practices*: Cambridge University Press.
- Ghezzi, A., Gabelloni, D., Martini, A., & Nataleccchio, A. (2018). Crowdsourcing: a review and suggestions for future research. *International Journal of Management Reviews*, 20(2), 343-363.
- Ghose, A., Ipeirotis, P. G., & Li, B. (2010). *Designing Ranking Systems for Hotels on Travel Search Engines to Enhance User Experience*. Paper presented at the ICIS.
- Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive science*, 16(3), 395-429.
- Goldkuhl, G. (2002). *Anchoring scientific abstractions—ontological and linguistic determination following socio-instrumental pragmatism*. Paper presented at the Proceedings of European Conference on Research Methods in Business, Reading.
- Goode, W., & Hatt, P. (1952). *Methods in social research*. New York: McGraw-Hill.
- Goodman, J. K., & Malkoc, S. A. (2012). Choosing here and now versus there and later: The moderating role of psychological distance on assortment size preferences. *Journal of Consumer Research*, 39(4), 751-768.

- Gregor, S. (2002). A theory of theories in information systems *Information Systems Foundations: building the theoretical base* (pp. 1-20). Canberra: Australian National University.
- Gregor, S. (2006). The nature of theory in information systems. *MIS quarterly*, 30(3), 611-642.
- Gregor, S., & Hevner, A. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 37(2), 337-355.
- Gregor, S., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association of Information Systems*, 8(5), 312-335.
- Gura, T. (2013). Citizen science: Amateur experts. *Nature*, 496(7444), 259-261.
- Heer, J., & Bostock, M. (2010). *Crowdsourcing graphical perception: using mechanical turk to assess visualization design*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Hevner, & Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol. 22): Springer Science & Business Media.
- Hevner, Von Alan, March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1), 75-105.
- Hevner, A., & Chatterjee, S. (2010b). Design science research in information systems *Design research in information systems* (pp. 9-22): Springer.
- Holley, R. (2010). Crowdsourcing: How and why should libraries do it? *D-Lib Magazine*, 16(3/4 Ma).
- Holmström, J., Ketokivi, M., & Hameri, A. (2009). Bridging practice and theory: A design science approach. *Decision Sciences*, 40(1), 65-87.
- Holsapple, C. W. (2008). DSS architecture and types *Handbook on Decision Support Systems 1* (pp. 163-189): Springer.
- Hosack, B., Hall, D., Paradice, D., & Courtney, J. (2012). A look toward the future: decision support systems research is alive and well. *Journal of the Association for Information Systems*, 13(5), 315.
- Howe, J. (2006). The rise of crowdsourcing. *Wired magazine*, 14(6), 1-4.
- Hyde, K. F. (2000). Recognising deductive processes in qualitative research. *Qualitative market research: An international journal*, 3(2), 82-90.
- Iacono, J. C., Brown, A., & Holtham, C. (2011). The use of the case study method in theory testing: the example of steel emarketplaces. *The Electronic Journal of Business Research Methods*, 9(1), 57-65.
- Iivari, J. (2007). A paradigmatic analysis of information systems as a design science. *Scandinavian journal of information systems*, 19(2), 5.
- Jarmolowicz, D., Bickel, W., Carter, A., Franck, C., & Mueller, E. (2012). Using crowdsourcing to examine relations between delay and probability discounting. *Behavioural processes*, 91(3), 308-312.
- Järvinen, P. (2007). Action research is similar to design science. *Quality & Quantity*, 41(1), 37-54.
- Jiang, L., Wagner, C., & Nardi, B. (2015). *Not Just in it for the Money: A Qualitative Investigation of Workers' Perceived Benefits of Micro-task Crowdsourcing*. Paper presented at the 48th Hawaii International Conference on System Sciences.
- Jonker, J., & Pennink, B. (2010). *The essence of research methodology: A concise guide for master and PhD students in management science*: Springer Science & Business Media.
- Jump, P. (2015). PhD: Is the doctoral thesis obsolete? *Times Higher Education*.

- Kafura, D. (1998). *Object-oriented software design and construction with C++*: Prentice-Hall, Inc.
- Kampffmeyer, H., & Zschaler, S. (2007). Finding the pattern you need: The design pattern intent ontology *International Conference on Model Driven Engineering Languages and Systems* (pp. 211-225). Berlin, Heidelberg: Springer.
- Kanazawa, S. (1999). Using laboratory experiments to test theories of corporate behavior. *Rationality and Society*, 11(4), 443-461.
- Khan, J. (2011). *Research methodology*: APH Publishing.
- Kietzmann, J. (2017). Crowdsourcing: A revised definition and introduction to new research. *Business Horizons*, 60(2), 151-153.
- Kim, J. (2012). The effect of design characteristics of mobile applications on user retention: an environmental psychology perspective.
- Kitchenham, B., Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1), 7-15.
- Kittur, A., Chi, E., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 453-456): ACM.
- Kittur, A., Smus, B., Khamkar, S., & Kraut, R. (2011). Crowdforge: Crowdsourcing complex work *Proceedings of the 24th annual ACM symposium on User interface software and technology* (pp. 43-52): ACM.
- Kleemann, F., Voß, G. G., & Rieder, K. (2008). Un (der) paid innovators: The commercial utilization of consumer work through crowdsourcing. *Science, technology & innovation studies*, 4(1), PP. 5-26.
- Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American journal of occupational therapy*, 45(3), 214-222.
- Kroes, P. (2002). Design methodology and the nature of technical artefacts. *Design studies*, 23(3), 287-302.
- Kuechler, W., & Vaishnavi, V. (2012). A framework for theory development in design science research: multiple perspectives. *Journal of the Association for Information Systems*, 13(6), 395.
- Kumar, R. (2005). *Research methodology*. Australia: Pearson.
- Lampe, C., Zube, P., Lee, J., Park, C. H., & Johnston, E. (2014). Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly*, 31(2), 317-326.
- Lee, A. S., & Hubona, G. S. (2009). A scientific basis for rigor in information systems research. *MIS quarterly*, 237-262.
- Lee, J., & Bui, T. (2000). *A template-based methodology for disaster management information systems*. In *Proceedings of the 33rd annual Hawaii international conference on system sciences* (pp. 7-pp). IEEE.
- Li, M., Tan, C., Wei, K., & Wang, K. (2017). Sequentiality of product review information provision: an information foraging perspective. *MIS Quarterly*, 41(3), 867-892.
- Lin, J., Amini, S., Hong, J. I., Sadeh, N., Lindqvist, J., & Zhang, J. (2012). *Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing*. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.
- Liu, F., Myers, M. D. J. J. o. S., & Technology, I. (2011). An analysis of the AIS basket of top journals. 13(1), 5-24.
- Lokke, A., & Sorensen, P. (2014). Theory testing using case studies. *Electronic Journal of Business Research Methods*, 12(1), 66-74.

- Lowry, P. B., D'Arcy, J., Hammer, B., & Moody, G. D. (2016). "Cargo Cult" science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels. *The Journal of Strategic Information Systems*, 25(3), 232-240.
- Lu, Y., Gupta, A., Ketter, W., & Van Heck, E. (2016). Exploring Bidder Heterogeneity in Multichannel Sequential B2B Auctions. *Management Information Systems Quarterly*, 40(3), 645-662.
- Lynham, S. (2002). The general method of theory-building research in applied disciplines. *Advances in developing human resources*, 4(3), 221-241.
- Maaninen-Olsson, E., Wismén, M., & Carlsson, S. (2008). Permanent and temporary work practices: knowledge integration and the meaning of boundary activities. *Knowledge Management Research & Practice*, 6(4), 260-273.
- MacKenzie, S., Podsakoff, P., & Podsakoff, N. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, 35(2), 293-334.
- Malan, R., & Bredemeyer, D. (2001). Defining non-functional requirements.
- Malone, T., & Crowston, K. (1994). The Interdisciplinary Study of Coordination. *ACM Computing Surveys*, 26(1), 87-119.
- March, S., & Smith, G. (1995). Design and natural science research on information technology. *Decision support systems*, 15(4), 251-266.
- Marsick, V. J. (1990). Altering the paradigm for theory building and research in human resource development. *Human Resource Development Quarterly*, 1(1), 5-24.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods*, 44(1), 1-23.
- McInnis, B., Cosley, D., Nam, C., & Leshed, G. (2016). Taking a HIT: Designing around rejection, mistrust, risk, and workers' experiences in Amazon Mechanical Turk. *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2271-2282): ACM.
- Miles, M., & Huberman, A. (1994). *Qualitative Data Analysis*. Thousand Oaks, California: SAGE.
- Miles, M., Huberman, A., & Saldana, J. (2013). Qualitative data analysis: A methods sourcebook: Thousand Oaks, CA: Sage Publications Inc.
- Mingers, J. (2003). The paucity of multimethod research: a review of the information systems literature. *Information Systems Journal*, 13(3), 233-249.
- Mueller, B., & Urbach, N. (2013). *The Why, What and How of Theories in Information Systems Research*. Paper presented at the 34th International Conference on Information Systems (ICIS).
- Mumford, E. (1985). Defining system requirements to meet business needs: a case study example. *The Computer Journal*, 28(2), 97-104.
- Muntermann, J., Nickerson, R., & Varshney, U. (2015). Towards the Development of a Taxonomic Theory. Paper presented at the 21st Americas Conference on Information Systems, Puerto Rico.
- Nakatsu, R., Grossman, E., & Iacovou, C. (2014). A taxonomy of crowdsourcing based on task complexity. *Journal of Information Science*, 40(6), 823-834.
- Nemati, H. R., Steiger, D. M., Iyer, L. S., & Herschel, R. T. J. D. S. S. (2002). Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing. 33(2), 143-161.
- Niederman, F., & March, S. (2015). Reflections on Replications. *AIS Transactions on Replication*.

- Niederman, F., & March, S. (2019). The “Theoretical Lens” Concept: We All Know What it Means, but do We All Know the Same Thing? *Communications of the Association for Information Systems*, 44.
- Nunamaker Jr, J. F., Chen, M., & Purdin, T. D. (1990). Systems development in information systems research. *Journal of management information systems*, 7(3), 89-106.
- Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts Work. Pap. Inf. Syst.*, 10, 26.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgement and Decision Making*, Vol. 5, No. 5.
- Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. *European Journal of Information Systems*, 1-16.
- Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183-199.
- Park, S., Shoemark, P., & Morency, L.-P. (2014). *Toward crowdsourcing micro-level behavior annotations: the challenges of interface, training, and generalization*. Paper presented at the Proceedings of the 19th international conference on Intelligent User Interfaces.
- Parvanta, C., Roth, Y., & Keller, H. (2013). Crowdsourcing 101 a few basics to make you the leader of the pack. *Health promotion practice*, 14(2), 163-167.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153-163.
- Peffers, K., Rothenberger, M., Tuunanen, T., & Vaezi, R. (2012). *Design science research evaluation*. Paper presented at the International Conference on Design Science Research in Information Systems.
- Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45-77.
- Penker, M. (2000). Business Modeling with UML: Business Patterns at Work: New York, NY: John Wiley & Sons.
- Piper, S. (2006). Qualitative theory testing as mixed-method research. *Journal of Research in Nursing*, 11(3), 183-193.
- Popper, K. (2014). *Conjectures and refutations: The growth of scientific knowledge*: Routledge.
- Power, D. J. (2008). Decision support systems: a historical overview. *Handbook on Decision Support Systems* 1, 121-140.
- Prat, N., Comyn-Wattiau, I., & Akoka, J. (2015). A taxonomy of evaluation methods for information systems artifacts. *Journal of Management Information Systems*, 32(3), 229-267.
- Pries-Heje, J., & Baskerville, R. (2008). The design theory nexus. *MIS quarterly*, 731-755.
- Pries-Heje, J., Baskerville, R., & Venable, J. R. (2008). *Strategies for Design Science Research Evaluation*. Paper presented at the ECIS.
- Raghunathan, S. (1999). Impact of information quality and decision-maker quality on decision quality: a theoretical model and simulation analysis. *Decision Support Systems*, 26(4), 275-286.

- Rai, A., Lang, S. S., & Welker, R. B. (2002). Assessing the validity of IS success models: An empirical test and theoretical analysis. *Information systems research*, 13(1), 50-69.
- Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. (2016). Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior*, 58, 354-360.
- Reason, J. (2008). *The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries*. Surrey, England: Ashgate.
- Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, 4(2), 155-169.
- Roberts, N., Campbell, D., & Vijayasarathy, L. (2016). Using information systems to sense opportunities for innovation: integrating postadoptive use behaviors with the dynamic managerial capability perspective. *Journal of Management Information Systems*, 33(1), 45-69.
- Ross, J., Zaldivar, A., Irani, L., & Tomlinson, B. (2009). Who are the turkers? worker demographics in amazon mechanical turk. *Department of Informatics, University of California, Irvine, USA, Tech. Rep.*
- Sanyal, P. (2016). Characteristics and Economic Consequences of Jump Bids in Combinatorial Auctions. *Information Systems Research*, 27(2), 347-364.
- Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics & Management*(1), 93-107.
- Schön, D. (1992). Designing as reflective conversation with the materials of a design situation. *Knowledge-based systems*, 5(1), 3-14.
- Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS Quarterly*, 35(1), 37-56.
- Selander, L., & Jarvenpaa, S. (2016). Digital action repertoires and transforming a social movement organization. *MIS Quarterly*, 40(2), 331-352.
- Shank, D. B. (2016). Using crowdsourcing websites for sociological research: The case of Amazon Mechanical Turk. *The American Sociologist*, 47(1), 47-55.
- Simon, H. (1996). *The Sciences of the Artificial*. Cambridge, USA: The MIT Press.
- Simon, H. A. (1960). The new science of management decision. *The Ford distinguished lectures: Vol. 3*. New York, NY, US: Harper & Brothers. <http://dx.doi.org/10.1037/13978-000>.
- Smith, D., Manesh, M. M. G., & Alshaikh, A. (2013). How can entrepreneurs motivate crowdsourcing participants? *Technology Innovation Management Review*, 3(2).
- Snoek, C. G., Freiburg, B., Oomen, J., & Ordelman, R. (2010). *Crowdsourcing rock n'roll multimedia retrieval*. Paper presented at the Proceedings of the 18th ACM international conference on Multimedia.
- Sonnenberg, C., & vom Brocke, J. (2011). *Evaluation patterns for design science research artefacts*. Paper presented at the European Design Science Symposium.
- Spencer, D. (2009). *Card sorting: Designing usable categories*: Rosenfeld Media.
- Spencer, D., & Warfel, T. (2004a). Card sorting: a definitive guide. *Boxes and arrows*. Retrieved from <https://boxesandarrows.com/Card-sorting-a-definitive-guide/>
- Steelman, Z., Hammer, B., & Limayem, M. (2014). Data Collection in the Digital Age: Innovative Alterantives to Student Samples. *Mis Quarterly*, 38(2), 355-378.

- Steelman, Z., Soror, A., Limayem, M., & Worrell, D. (2012). Obsessive compulsive tendencies as predictors of dangerous mobile phone usage. Paper presented at AMCIS.
- Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing Samples in Cognitive Science. *Trends in cognitive sciences*, 21(10), 736-748.
- Stinchcombe, A. L. (1987). *Constructing social theories*: University of Chicago Press.
- Su, A. I., Good, B. M., & van Wijnen, A. J. (2013). Gene Wiki Reviews: marrying crowdsourcing with traditional peer review. *Gene*, 531(2), 125.
- Thomas, D. (2006). A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, 27(2), 237-246.
- Thuan, N., Antunes, P., & Johnstone, D. (2016). Factors influencing the decision to crowdsource: A systematic literature review. *Information Systems Frontiers*, 18(1), 47-68.
- Thuan, N., Antunes, P., Johnstone, D., & Duy, N. (2015). *Establishing a Decision Tool for Business Process Crowdsourcing*. Paper presented at the International Conference on Future Data and Security Engineering.
- Thuan, N. H., Antunes, P., & Johnstone, D. (2016). *A Design Science Method for Emerging Decision Support Environments*. Paper presented at the Australasian Conference on Information Systems, Adelaide, Australia. *arXiv preprint arXiv:1605.04725*.
- Turner, A. M., Kirchhoff, K., & Capurro, D. (2012). Using crowdsourcing technology for testing multilingual public health promotion materials. *Journal of medical Internet research*, 14(3).
- Uchoa, A. P., Esteves, M. G. P., & De Souza, J. M. (2013). Mix4Crowds - Toward a Framework to Design Crowd Collaboration with Science. *17th IEEE International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, 61-66.
- Van Aken. (2005). Management research as a design science: Articulating the research products of mode 2 knowledge production in management. *British journal of management*, 16(1), 19-36.
- Van Aken, & Romme, A. G. (2012). A design science approach to evidence-based management. 43-57.
- Van Der Valk, T., Van Driel, J. H., & De Vos, W. (2007). Common characteristics of models in present-day scientific practice. 37(4), 469-488.
- Van Gelder, M., Bretveld, R., & Roeleveld, N. (2010). Web-based questionnaires: the future in epidemiology? *American journal of epidemiology*, 172(11), 1292-1298.
- Van Looy, A., Poels, G., & Snoeck, M. (2017). Evaluating business process maturity models. *Journal of the Association for Information Systems*, 18(6), 461.
- Venable John, Pries-Heje, J., & Baskerville, R. (2012). *A comprehensive framework for evaluation in design science research*. Paper presented at the International Conference on Design Science Research in Information Systems.
- Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, 48(4), 865-885.
- Vitak, J., Proferes, N., Shilton, K., & Ashktorab, Z. (2017). Ethics regulation in social computing research: Examining the role of Institutional Review Boards. *Journal of Empirical Research on Human Research Ethics*, 12(5), 372-382.
- Von Alan, R., Hevner, A., March, S., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1), 75-105.

- Vondrick, C., Patterson, D., & Ramanan, D. (2013). Efficiently scaling up crowdsourced video annotation. *International Journal of Computer Vision*, 101(1), 184-204.
- Vukovic, M., Laredo, J., & Rajagopal, S. (2010). *Challenges and experiences in deploying enterprise crowdsourcing service*. Paper presented at the International Conference on Web Engineering.
- Vukovic, M., Lopez, M., & Laredo, J. (2010). *Peoplecloud for the globally integrated enterprise*. Paper presented at the Service-Oriented Computing. ICSOC/ServiceWave 2009 Workshops.
- Wacker, J. (1998). A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of Operations Management*, 16(4), 361-385.
- Wand, Y., & Weber, R. J. (1995). On the deep structure of information systems. *Information Systems Journal*, 5(3), 203-223.
- Ward, M. K., & Broniarczyk, S. M. (2011). It's not me, it's you: How gift giving creates giver identity threat as a function of social closeness. *Journal of Consumer Research*, 38(1), 164-181.
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii-xxiii.
- Wei, W.C. (2009). A technology acceptance model: Mediate and moderate effect. *Asia Pacific Management Review*, 14(4), 461-476.
- West, D., & Stowell, F. J., Artech, London. (1999). Models and diagrams and their importance to information systems analysis and design. 295-311.
- Wexler, M. N. (2011). Reconfiguring the sociology of the crowd: exploring crowdsourcing. *International Journal of Sociology and Social Policy*, 31(1/2), 6-20.
- Whitla, P. (2009). Crowdsourcing and its application in marketing activities. *Contemporary Management Research*, 5(1).
- Witschey, J., Murphy-Hill, E., & Xiao, S. (2013). Conducting interview studies: Challenges, lessons learned, and open questions *1st International Workshop on Conducting Empirical Studies in Industry* (pp. 51-54): IEEE.
- Witschey, J., Murphy-Hill, E., & Xiao, S. (2013). *Conducting interview studies: Challenges, lessons learned, and open questions*. Paper presented at the Conducting Empirical Studies in Industry (CESI), 2013 1st International Workshop on.
- Wu, W., Tsai, W.-T., & Li, W. (2013). Creative software crowdsourcing: from components and algorithm development to project concept formations. *International Journal of Creative Computing*, 1(1), 57-91.
- Yang, J., Adamic, L. A., & Ackerman, M. S. (2008). *Crowdsourcing and knowledge sharing: strategic user behavior on taskcn*. Paper presented at the Proceedings of the 9th ACM conference on Electronic commerce.
- Yuen, M.C., King, I., & Leung, K.S. (2011). *A survey of crowdsourcing systems*. Paper presented at the Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on.
- Zhang, K., Bhattacharyya, S., & Ram, S. (2016). Large-Scale Network Analysis for Online Social Brand Advertising. *MIS Quarterly*, 40(4), 849-868.
- Zhao, Y., & Zhu, Q. (2012). Exploring the motivation of participants in crowdsourcing contest. Paper presented at the 33rd International Conference on Information Systems, Orlando.

- Zheng, H., Li, D., & Hou, W. (2011). Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce*, 15(4), 57-88.
- Zigurs, I., & Khazanchi, D. (2008). From profiles to patterns: A new view of task-technology fit. *Information systems management*, 25(1), 8-13.

CHAPTER 7

DESIGN AND DEVELOPMENT OF A DSS SUPPORTING THE INTEGRATION OF CROWDSOURCING IN THEORY TESTING: A DESIGN SCIENCE PERSPECTIVE

ABSTRACT

The integration of crowdsourcing in behavioral research in the IS field offers several advantages and opportunities. This paper builds on prior study, employing a design science research (DSR) paradigm to design, develop and evaluate a tool that assists researchers adopting crowdsourcing when testing theory about behavioral phenomena. The proposed tool is based on an extensive review of literature on how theory has been tested, and a pattern model that standardizes extracted concepts, activities, processes and relationships into patterns. We discuss the architecture of the proposed tool and present two prototypes, one used for knowledge articulation by representing, extracting, organizing and acting on relevant information and the other on decision making and recommendation for the tool users. Evaluation results show the applicability and utility of the tool.

Keywords: Decision support system, crowdsourcing, theory testing, design science

7.1 INTRODUCTION

Developing and testing theory about behavioural phenomena is common in the information systems (IS) field (Gregor , 2006). However, it faces many conceptual and practical challenges. Researchers must consider a diversity of variables, settings and data necessary to form robust evidence about systems and human behaviour. Researchers also have to reflect on the diversity of methodologies, methods, processes, instruments, and tools available in the IS methodological toolkit (Bhattacherjee, 2012). Furthermore, the operationalization of data collection procedures, which may involve a large number of people, systems and processes, often require long periods of trial and error (Peer et al., 2017; Witschey et al., 2013). Considering all these difficulties, a crowdsourcing strategy may seem attractive to help researchers to accomplish their theory testing objectives.

Crowdsourcing is a managerial model that relies on information technology to outsource tasks to a large number of participants using different types of incentives such as remuneration, prizes and peer-esteem (Kietzmann, 2017). Tasks can be massively distributed, done in parallel and fulfilled in short time periods. Furthermore, the task and the crowd can be easily managed through crowdsourcing platforms like Wikipedia, MTurk, Upwork, and InnoCentive.

Several researchers have already noted several advantages of adopting crowdsourcing when testing theory about behavioural phenomena. Some areas in social studies are using crowdsourcing as a de facto standard for delivering questionnaires (Bates & Lanza, 2013; Shank, 2016). Lowry et al. (2016) pointed out that crowdsourcing could increase the quality of data collection, e.g. by crowdsourcing quality control, diversifying methods and instruments, and ultimately collecting more data. Steelman et al. (2014) also highlighted several advantages of crowdsourcing in dealing with psychometrics, demographics and structural properties of data samples.

Enwereuzo et al. (2018, 2019) considered the feasibility of applying the crowdsourcing model to theory testing. In particular, the study developed a systematic procedure for checking if theory testing activities, either individually or collectively, could be crowdsourced. This study builds on prior work with the goal to *design, develop and evaluate a tool that assists researchers adopting crowdsourcing when testing theory about behavioural phenomena*. This research provides two complementary contributions: 1) an innovative decision support tool addressing a problem relevant to behavioural researchers in the IS field; and 2) results from the tool evaluation, which demonstrate the utility of the tool in resolving the research problem.

The remainder of this paper is structured as follows. Section two describes the research context of the study. Section three describes the tool development method. In Section four, we describe the design, development and evaluation of the tool. Finally, Section five highlights the research contributions and conclusions.

7.2 RESEARCH CONTEXT

Behavioural researchers in the IS field work with various types of empirical data, which they have to gather and analyse in order to determine whether the propositions articulated by theory are supported or not by the data. A variety of methods can be used by researchers. A common perspective splits methods into quantitative, qualitative and mixed categories (Creswell & Creswell, 2017). Some of these methods require a considerable effort to apply. For instance, observation often requires shadowing the activities of multiple workers in their natural work settings over long periods of time. Then, qualitative analysis of video footage will require many hours going through video footage, indexing, transcribing, coding, condensing, and synthesizing data (Derry et al., 2010). To increase consistency and reliability, quality checks and duplication of activities, such as parallel coding, are often necessary (Thomas, 2006). And in the end, the whole data gathering/analysis process may have to be repeated until reaching acceptable quality.

Investigating how crowdsourcing can alleviate the work done by behavioural researchers could have a significant impact in the IS field. In particular, expediting the theory building/testing cycle could foster new IS theory; and could also contribute to improving existing, or developing new, IS research methods.

We note that using crowdsourcing for data collection has been around for some time. In particular, crowdsourcing is becoming a very common way to deliver questionnaires to study participants (Bates & Lanza, 2013; Behrend et al., 2011; Jarmolowicz et al., 2012; Peer et al., 2017). It has also started to be used as an instrument to collect system usage logs (Kittur et al., 2008; Stewart et al., 2017). And it is already an important component of citizen science, supporting distributed collection of field data (Bonney et al., 2009; Gura, 2013). However, the relationship between crowdsourcing and theory building/testing has not yet been systematically explored and established.

7.2.1 Research Background and Goals

The whole design science research project focuses on the development of the DSS, some intermediate steps to achieve this goal has been developed in past studies. Early steps in establishing this connection were reported by Enwereuzo et al. (2018, 2019). These studies addressed three fundamental goals: 1) develop a conceptual framework of theory building/testing; 2) develop a model characterising how researchers have been testing theory; and 3) develop a procedure for checking which theory testing activities could be crowdsourced. These previous studies form the foundation for the development of this last phase of the project (see Figure 7.1). The following briefly outlines each of these three outcomes.

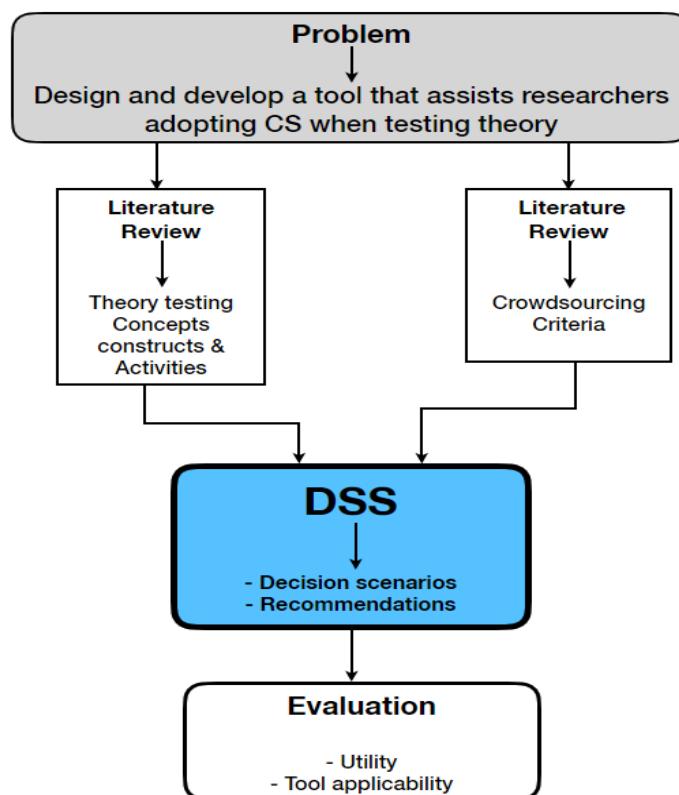


Figure 7.1 Research framework

Conceptual framework. The conceptual framework characterises the main concepts and constructs involved in theory development. The framework is significant because it characterises theory testing as a distinct component of theory building. Earlier frameworks blend theory testing with theory building. By separating theory building

from theory testing, we make the research process more amenable for crowdsourcing, since theory testing activities are the ones that could benefit most from mass distribution to the crowd.

Model. The model is a key element in understanding how crowdsourcing can actually be applied to theory testing. The model synthesises an extensive literature review on the different ways in which researchers have been testing theory within the IS discipline. The model characterises theory testing as a system consisting of various research activities interconnected by a set of patterns, which realise several research goals. The model is significant because it codifies and integrates knowledge on how theory has been tested in practice. The model is also instrumental in our pursuit of bringing crowdsourcing into theory testing because it supports researchers identifying regularities in theory testing activities, which may then be individually or collectively checked if they can be crowdsourced or not.

Procedure. The question asked was “how can a researcher decide if a theory testing activity can be crowdsourced or not?” The developed procedure systematically uses a set of crowdsourcing criteria that can be used by researchers to decide if a theory testing activity, or a set of activities, can be crowdsourced or not. The main contribution of this procedure is to give clarity and systematicity to the decision-making process.

This study builds on these outcomes in order to:

- Assist researchers establish if Crowdsourcing (CS) is a good strategy to adopt for any of the theory testing steps. This assistance could be given as a guideline or recommendation.
- Build a comprehensive and integrated view of the theory testing activities and its associated crowdsourcing attributes.
- Provide a means for effective processing and presentation of knowledge as it relates to theory testing activities and crowdsourcing.

To achieve the outlined goals, we briefly consider what type of decision support system is needed.

7.2.2 Decision Support Systems View

Decision Support Systems (DSS) represent a research area in Information Systems (IS) with a long history that can be traced back to Simon's intelligence-design-choice model developed in 1960 (Hosack et al., 2012; Simon, 1960). The main focus in this research area is centred on support and improving decision-making (Arnott & Pervan, 2005). In many cases, these systems are not meant to replace the decision makers (in our case, the researchers), but to help them extend their capabilities and to make better and more informed decisions, highlighting why the term 'support' is important (Hosack et al., 2012).

A large number of DSS have been studied and developed in IS and its related fields, and this led to several taxonomies being proposed. Five types of DSS were suggested by Power (2008), which are: data driven, model driven, knowledge driven, document driven and communication driven DSS. While Arnott and Pervan (2005) developed a seven-type taxonomy, which are: 1) personal DSS for individual managers; 2) group DSS for a group of decision makers; 3) negotiation support systems for group support systems but involves negotiation functions; 4) intelligent DSS, uses artificial intelligence; 5) knowledge DSS, provides knowledge for storage, retrieval, transfer, and application; 6) data warehousing for processing large-scale (big) data; and 7) enterprise reporting and analysis systems. Based on these taxonomies, we note that our tool is meant to support individual researchers, and as such it is a personal, model-driven DSS. It is a model driven DSS because it uses limited data and parameters, and it does not require large data bases (Power, 2008).

DSS are most often software based, and assist with decisions mostly by illustrating possible outcomes, which can be presented visually, numerically or by leading users through logical decision steps (Dicks et al., 2014). Some of these tools are complex models, mainly operated by their developers, while others have simple interfaces that can be used by non-experts. We adopt the latter description because the type of users that are targeted are relatively inexperienced users such as PhD students. We intend to develop our tool in such a way that it is user-friendly and easy to use.

7.3 RESEARCH METHOD

This study follows the Design Science Research (DSR) paradigm (Hevner & Chatterjee, 2010; Hevner et al., 2004). This paradigm was adopted because DSR concerns problem-solving (Gregor & Hevner, 2013): 1) it seeks to create and evaluate innovative, first-of-a-kind IS artefacts; 2) which solve relevant organizational problems; while 3) founding the artefact design on a rigorous knowledge basis, which simultaneously informs the artefact design with existing theory, methods and artefacts, and contributes to knowledge by delivering new theory, methods and artefacts.

In our case, the artefact design is anchored on a conceptual framework of theory testing, a pattern model of theory testing, and a procedure for checking if theory testing activities can be crowdsourced. These three artefacts can be considered *secondary artefacts*, which inform the design of a decision support tool, which is the *primary artefact* of the study. The design of an artefact based on secondary artefacts provides support and guidance (Hevner & Chatterjee, 2010; Hevner et al., 2004) and establishes a solid knowledge foundation on which to design the primary artefact.

The DSR paradigm also distinguishes two primary research activities: *build* and *evaluate* (March & Smith, 1995; Peffers et al., 2007). The build activity involves generating a preliminary problem frame: a set of requirements and design rules which then guide the artefact design (see figure 7.2). The build activity is followed by the evaluation activity, which confronts the design against the research goals. The evaluation results may suggest new problem frames, which may then lead towards new build activities, until a satisfying solution is obtained (Sein et al., 2011). Our build activity involves the design of decision support tool using a set of crowdsourcing attributes, feasibility checks as to activities that can be crowdsourced, while the evaluate activity assesses the *utility* brought by the decision support tool.

When choosing a DSR evaluation method, it is important to balance the interests of practitioners and researchers (Sonnenberg & vom Brocke, 2011). While practitioners are interested in aspects such as the applicability and usefulness of the artefact, researchers are more interested in the validity of the research process and outputs. Furthermore, the evaluation may be done *ex ante*, where artefacts are evaluated prior

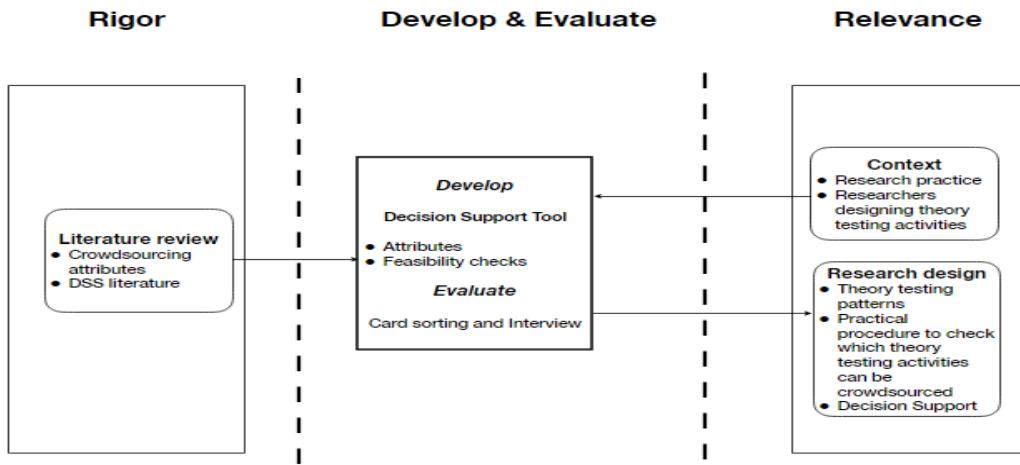


Figure 7.2. Research Framework for DSS Tool (Adapted from (Peffers et al., 2007)

to their implementation or actual construction, and ex post, where artefacts are evaluated after they have been designed and constructed (Pries-Heje et al., 2008). For these reasons, we adopt the following evaluation strategy proposed by Sonnenberg and vom Brocke (2011), which can be seen in figure 7.3.

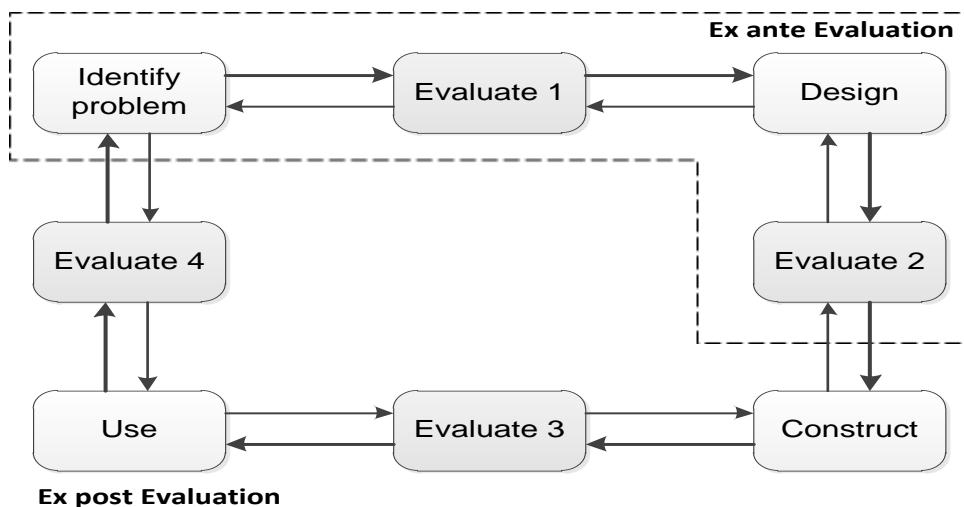


Figure 7.3. Build and Evaluation Activities (Adapted from (Sonnenberg & vom Brocke, 2011))

Problem Identification

The research problem was identified in previously conducted studies (Enwereuzo et al., 2017), which focused on highlighting theory testing as a distinct component of theory building. We disentangled the two concepts by first defining the theory

generation cycle. Theory generation is a cyclic where theory is consecutively built, checked against empirical data and adjusted to reflect the real world (Lynham, 2002). Then we aligned theory building and testing in such a way that each one provides outputs reinforcing the other by creating boundary activities.

Evaluation 1

After problem identification activity, there is an evaluation done. This evaluation serves the purpose of ensuring that a meaningful DSR problem is selected and formulated. It also helps to demonstrate whether the envisioned design problem is important for practice, and represents a research gap (Sonnenberg & vom Brocke, 2011). The card sorting method and interviews were used for this evaluation (Enwereuzo et al., 2017). The outcome of the experiment supported the identified problem, which showed that PhD students were not aware of the theory testing activities, and there was so much mix-up of the two concepts – theory building and theory testing.

Design

Based on the result of the first evaluation, changes or revisions were made to the problem identification activity, and then the design of the artefact was done (Enwereuzo et al., 2017). The designed and developed artefact accomplished the following requirements:

- Highlight theory testing steps
- Provide theory testing pathways and activities to help with decision making
- Highlight CS requirements necessary for theory testing
- Give recommendations as to what theory testing activity(ies) can be crowdsourced

Evaluation 2

The second evaluation done served the purpose of showing that the artefact design incorporates the solution to the stated problem (Sonnenberg & vom Brocke, 2011). The evaluation method used were card sorting and semi-structured interviews. These methods were selected because of the targeted audience for the design which are researchers, young and upcoming researchers.

Construct

Feedback was gotten from the design stage, which enabled changes to be made to the design. A CS template was constructed based on the design (Enwereuzo et al., 2018). This template was constructed based on previously developed artefacts and the results of their respective evaluations.

Evaluation 3

This evaluation is carried out to initially demonstrate if the artefact performed well and how well it performed. Inferences on the utility of the constructed template were made because it initiates and informs subsequent artefact design activity. Evaluation method used were card sorting and interview. The result of the evaluation gave us an insight on how users perceived the artefact, especially about its utility and ease of use. Feedback and comments were obtained, and some changes were made.

Use

This is the last phase of the cycle, this stage focuses on instantiation, which can be in form of a software. We would be focusing on the development and use of the proposed software. This development is as a result of the constructed template and its evaluation.

Evaluation 4

This form of evaluation involves real users using real systems to accomplish real tasks in real settings (Pries-Heje et al., 2008). Evaluation criteria for this study would focus on applicability, understandability, utility and impact of artefact on users (Prat et al., 2015). This evaluation would involve real users using real systems to accomplish real tasks in real settings (Pries-Heje et al., 2008). The evaluation method chosen are card sorting, observation and interviews. This method is used in this study due to the success of subsequent ones used in previous study and the amount of useful feedback received.

7.4 DSS ARTEFACT

The DSS is based on the three secondary artefacts, which are the conceptual model, the pattern model and the CS template. The pattern model and the CS template feeds the DSS and serves as its knowledge base. The DSS basically allows users to play with the model using a set of restrictions and requirements.

7.4.1 Tool's Conceptual Design

The conceptual design of the tool has three main components as shown in Figure 7.4, based on a framework proposed by Holsapple (2008): GUI (Graphic User Interface), Problem processing component (PPC) and the knowledge component (KC). The GUI component is responsible for the interaction between the users (researchers) and the tool. Inputs and parameters are accepted by this component. These inputs are then processed by the PPC. Input flows are also controlled by the PPC by choosing and adapting what elements the GUI presents. The PPC acquires the required knowledge from the KC, processes the knowledge, and returns the output to the user.

The KC consists of procedural knowledge extracted from the conceptual model (Holsapple, 2008), which characterizes steps for theory testing. It is a step-wise specification of what to do to accomplish theory testing. We also have the descriptive knowledge obtained from the pattern model – this provides the definitions and descriptions of concepts that must be considered in the process of decision-making. It also includes relationships among concepts and activities. Reasoning knowledge provides the CS requirements and criteria constraining these elements. Using the KC, researchers can make decisions as to what pathway to select and its related activities and recommendations as to what activity can be CS is provided.

7.4.2 Tool Development

The tool development consists of two phases: template-based DSS and the user interface-based DSS. Lee and Bui (2000) suggests that templates are useful tools for representing, extracting, organizing and acting on relevant information. Therefore, the domain knowledge articulation is demonstrated through the template, of which knowledge from the pattern model is transferred into computer-based formulation.

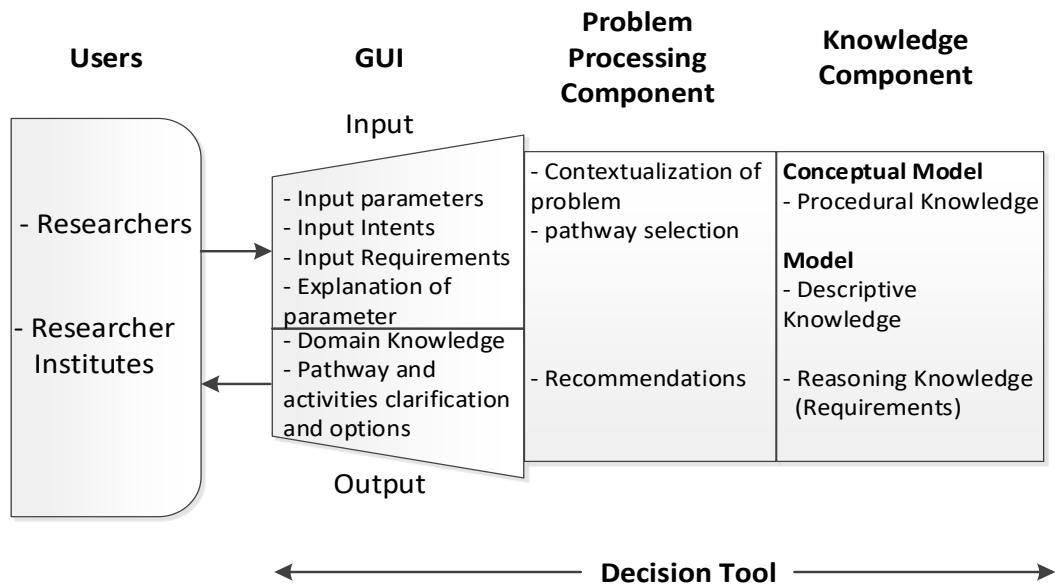


Figure 7.4. Tool's Conceptual Design (adapted from (Holsapple, 2008))

7.4.2.1 Template-based Prototype

Templates were used because of the numerous advantages it provides as a preliminary stage in design and development of a tool. By providing a checklist of relevant information to look for, templates support proactive processing of information. They make it easier to compare multiple cases and infer generalizations from them. They support the design of a system, by providing information needed for enacting or automating the process (Lee & Bui, 2000).

We develop a set of templates (See Figure 7.5) to systematically analyse how to adopt crowdsourcing for the various theory testing patterns previously defined. The templates allow for systematic checking of the properties of the crowdsourcing process (defined in the next section) while at the same time addressing the specific theory testing activities defined for a particular pattern. A template may suggest three types of results: cannot crowdsource the activity; difficult to crowdsource the activity; and easy to crowdsource the activity.

			Activity 1	Activity 2	Activity n	Intent	Cannot crowdsource	Difficult to crowdsource	Easy to crowdsource	
Strategy	Ethics	Transferable		Y		Proceed			X	
		Non-transferable		Y		Stop	X			
		Unethical		Y		Stop	X			
Specification	Complexity	Skills-based	N	Y	N	Proceed with activities with "Y"			X	
		Rules-based	Y	N	Y	Proceed with activities with "N"		X (1)	X (1)	
		Knowledge-based	Y	N	Y	Proceed with activities with "N"	X (1)	X (1)		
	Input	Well-defined		Y		Proceed			X (2)	
		III-defined		Y		Modify else stop	X			
		Well-defined		Y		Proceed			X (2)	
Execution	Context	III-defined		Y		Modify else stop	X			
		Free		Y		Proceed			X	
		Dependent		Y		Modify if possible	X			
	Support	Restricted		Y		Consider alternative else stop				
		Exist		Y		Proceed			X	
		Not needed		Y		Proceed			X	
Consolidation	Coordination	Must be developed		Y		Proceed if possible		X		
		Unavailable		Y		Consider alternative else stop	X			
		Pooled		Y		Proceed			X	
	Quality control	Coordinated		Y		Proceed			X	
		Shared		Y		Stop if support is unavailable, else proceed	X (3)	X (3)		
		Peer-reviewed		Y		Proceed			X	
Notes:										
1. Decision depends on required rules or knowledge.										
2. Both input and output must be well-defined.										
3. Depends if support is unavailable or not.										

Figure 7.5. Template-based tool on theory testing and CS

Crowdsourcing Criteria

The crowdsourcing process can generally be described as a sequence of four steps (figure 7.6). The first step considers the adoption of crowdsourcing at the strategic level. Then, we consider the specification and execution of tasks. Finally, the outcomes from the different tasks must be consolidated. To be able to complete this process with success, several properties must be fulfilled.

Linked to the adoption of crowdsourcing at the strategic level, we find the important topic of ethics. First, we should consider if it may be *unethical* to crowdsource. A variety of situations may suggest crowdsourcing to be unethical, e.g., gathering data from inside a hospital without consent. If the strategy is accepted from an ethical perspective, it would typically involve acquiring ethical approval when necessary, or sticking to ethical guidelines especially as relates to research. We need to further consider if the ethical mandate, such as ethical approval obtained by the researcher is *transferable* or *non-transferable* to the crowd. We suggest that both unethical goals and non-transferable ethical mandates should not be crowdsourced.

After considering the strategy, we then have to examine the specification of one or more crowdsourcing tasks. One critical property to consider at this level is task complexity (Ghezzi et al., 2018; Jiang et al., 2015; Vondrick et al., 2013). A task can be classified as (Antunes et al., 2012; Reason, 2008): *skills-based*, if it can be accomplished as a routine by the crowd, in an almost unconscious way; *rules-based*, if it requires following a plan, with some latitude of decision-making by the crowd; and *knowledge-based*, if it is novel, requiring considerable problem analysis and solution-based decision-making by the crowd. We suggest that, as we move from skills-based to knowledge-based tasks, crowdsourcing becomes more difficult to achieve. This recommendation is based on research showing that the crowd prefers shorter tasks (Gadiraju et al., 2015); and also that micro tasks - which are tasks that have been broken down into smaller tasks, enable the crowd to generate higher quality work (Cheng et al., 2015).

Another facet to analyze are the inputs and outputs required by tasks. Input is the problem the crowd is asked to resolve, while output is the solution developed by the crowd, both of which can be classified as either *well-defined* (e.g. classifying photos using a predefined set of tags) or *ill-defined* (e.g. generating new research ideas) (Nakatsu et al., 2014). We contend that tasks with well-defined inputs and outputs can be more easily crowdsourced. Ill-defined tasks may be more difficult to crowdsource, since the crowd is more involved in defining what the task is about, which may lead to unwanted outcomes. This recommendation is based on research showing that the crowd prefers tasks with low associated risks (McInnis et al., 2016), and should be considered here especially when such tasks are related to theory testing.

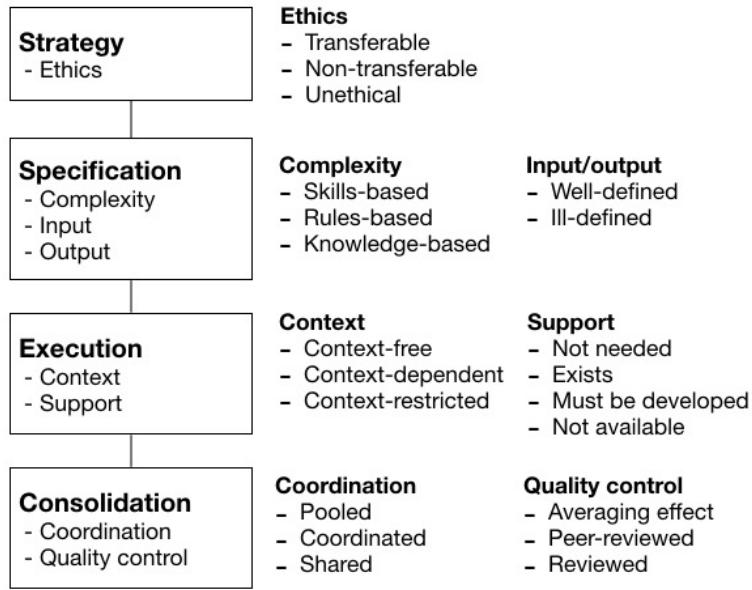


Figure 7.6. Crowdsourcing criteria

Regarding execution, two important properties must be considered: context and support. Some tasks are completely independent of the context in which they are executed, e.g. classifying photos according to given categories, which means they are *context-free*. Other tasks depend on a specific context to be executed, such as a physical location, operational environment, or virtual place. Examples include gathering data in public libraries, driving a car, or gathering data from commonly used social media. In these cases, tasks are *context-dependent*. Finally, some tasks may have to be performed in restricted contexts, which may not be easily accessible or reproducible. Examples include gathering data in private places such as military installations, coding data sets using proprietary tools, and selecting records from company databases. In these cases, we designate the tasks as *context-restricted*.

We suggest that context-free tasks are easy to crowdsource, while context-dependent tasks are difficult to crowdsource. The rationale for considering a context-dependent task difficult to crowdsource is that certain constraints have to be enforced, which make the task more difficult if not impossible to specify. In principle, we consider a context-restricted task not crowdsourceable. This is because, if it would be possible to circumvent the restrictions, then they would be dependencies, not restrictions. Consider, as an example, that you plan to crowdsource data collection in a hospital. Sending the crowd to the hospital without permission is not advisable, and therefore

the task is context-restricted. However, if you would get permission from the hospital, then the task would be context-dependent.

Support concerns the methods, tools and training required to perform the task. Regarding this criterion, we consider the following values: support is *not needed* when the task can be carried out by the crowd without any support (e.g. counting birds in a forest); support *exists* when training, methods and tools exist and can be used by the crowd to execute the task (e.g. a photo sharing tool is available to upload bird photos); support *must be developed*, when it is needed but must be developed by the researcher (e.g. a website must be created to classify birds in a certain way); and support is *not unavailable*, when it is needed but the researcher is either unwilling or unable to provide it. We consider that crowdsourcing can be easily done when support is not needed or exists, crowdsourcing is difficult to accomplish when support must be developed, and crowdsourcing cannot be accomplished when support is necessary but not available.

After executing the tasks, their outputs must be consolidated. Here, we have to consider what type of coordination is required to consolidate the various task outputs. We consider the following values (Kevin Crowston, 2012; Malone & Crowston, 1994): the task outputs may be *pooled* together (e.g. individual ideas can be merged), the outputs from different tasks may be *coordinated* (e.g. the outputs from idea generation may be fed to a selection task), or the outputs from different tasks may be *shared* (e.g. collaborative editing). We consider that crowdsourcing can be easily done when the consolidation is pooled or coordinated. Crowdsourcing will be difficult to accomplish in cases where consolidation requires a shared approach, and appropriate support must be developed. Crowdsourcing cannot be accomplished when a shared approach is needed, and support is not available.

Still regarding consolidation, the researcher has to consider quality control. Considering the characteristics of crowdsourcing, it seems natural that the outputs generated by the crowd should be subject to quality control. Quality control may rely on the *averaging effect*, which uses multiple data sources to increase accuracy and trust (Brunt & Meidell, 2018). It can also be *peer-reviewed*, where quality assurance is done by the crowd. And finally, the outputs may be *reviewed* by the researcher, using

either a sample or the entire data set. We consider that, if the averaging effect or peer-reviews are viable, then the task is easy to crowdsource. The review approach has a neutral impact in relation to crowdsourcing, as it is a traditional approach to quality control.

7.4.2.2 User Interface-based Prototype

The user interface-based prototype was developed as an improved and revised version of the template-based prototype. This prototype was implemented using Visual Basic for Applications (VBA) Excel, providing wider access to the knowledge base. VBA is the programming language used on an excel interface. It enables building of user-defined functions (UDFs), automating processes and assessing windows API through dynamic-link libraries (DLLs).

The prototype consists of two decision functions (Figure 7.7), the left-hand side provides the list of theory testing activities, which the user has an option of selecting the appropriate ones relating to the user's needs, of which when an activity is selected, it creates or generates options on the right-hand side. The right-hand side is option-based generated based on the crowdsourcing criteria.

Figure 7.7. User Interface-Based Prototype

This gives different options to the user, to assist in decision making as to if the selected activity could be crowdsourced or not, and the level of difficulty if it can be crowdsourced. There is a generate report button, that when clicked generates a visual report to the user, which the user can then decide if the chosen activity is worth crowdsourcing or not. There is also an option that allows the user to receive the report through email as an attachment.

7.5 EVALUATION

The evaluation strategy adopted for this study includes card sorting, interviews and observation. Card sorting is a qualitative evaluation method that has been widely used in various fields such as psychology, knowledge engineering, and software engineering (Barrett & Edwards, 1995), as well as DSR (Prat et al., 2015). Insights on how users view the problem and the solution addressed by an artefact can be gotten through card sorting. This is extremely important in the DSR context, since DSR addresses wicked problems for which there are no single best problem definitions and solutions (Rittel & Webber, 1973). We use card sorting to evaluate how our developed artefact relates to what exists in the researchers' minds. Our focus was evaluating the idea, what the tool is meant to achieve, and the process and not on the efficiency or effectiveness of the software itself. We structured the card sorting approach as a combination of card sorting exercises, interviews and observation. They were conceived to acquire expository knowledge regarding the utility, understandability and applicability of the artefact by:

E1: Establishing how useful the participants find the DSS in making a decision based on their research project.

E2: Testing user's understanding of the how the DSS process works and what it is meant to achieve;

E3: Testing the applicability of the DSS to their study, and research in general;

Setting and participants. The card sorting exercises were conducted in a meeting room, which had a large table where to lay out the cards and several computers for the exercise. PhD students conducting research projects in the IS field were the

participants used. PhD students were selected because they represent the main target audience for the DSS. 11 were selected by convenience.

Materials. The card sorting materials involved sets of 95 cards with words on them and a unique identifier number for recording purposes. Blank cards were also provided to write on if needed.

Procedure. The exercises were done in one-on-one sessions moderated by the first author (Figure 7.8). At the beginning of each session, the purpose of the exercises, the artefacts under evaluation, and the card sorting technique was expounded to the participant. The participants were then encouraged to express their thoughts and opinions during the exercises, to provide detailed feedback about what they were thinking and doing. The exercises were then operationalised according to the following steps.



Figure 7.8. Card sorting exercise

The purpose of each exercise was explained before handing over a deck of cards to the participant. The participant was then given some time to read through the cards for familiarization of contents. The first exercise contained a deck of cards with theory testing activities, which the participant should group in three categories: Cannot CS, difficult to CS and easy to CS. The participant is then given sometime to use the developed tool based on their research project. The second exercise and subsequent exercises, interview and observations were aimed at getting participants reactions, and feedback regarding how beneficial the DSS was in making decisions if a chosen theory testing activity can be crowdsourced.

Along with information about the card sorts, we also gathered the participant's reactions and comments made throughout the exercises, focusing on the positive and negative reactions. The exercises were audio recorded and captured to ease later analysis. The moderator also observed the impact of the tool on the users, focusing on their reactions, and facial expressions, notes were taken. All participants completed the entire procedure and each session took an average of 40 minutes.

7.5.1 Usefulness of the DSS in Decision Making (E1)

This exercise was aimed at determining if the participants found the DSS useful in deciding to CS the selected theory testing activity based on their research project. Participants were told to sort the cards into three categories: Cannot CS, difficult to CS and easy to CS. The cards contained different theory testing activities. The cards had unique numbers, making it easy for analysis. The participants were then asked to use the tool to determine if such activities were placed in the right category based on the recommendation of the tool. The result suggests that 20% had little mismatch, because they had prior knowledge of CS, while 80% had a lot of mismatch because they concept of CS was new to them. Using the tool helped in categorizing the mismatch activity (Pries-Heje & Baskerville) in its right category, especially based on their research project. From the second exercise conducted, 40% of the participants said the tool was very useful, while 50% said it was somewhat useful, the last 10% said it was somewhat not useful.

Interviews were done to understand the reason behind their choices, especially for those that considered it somewhat not useful. Reason behind their response was due to the fact of being at the tail-end of their study, and have used other methods to acquire data, so the tool is not beneficial to them now but would be beneficial for other studies. Generally, the participants felt it was useful and advised that the tool be used at the beginning of one's study, to help with decision making.

Observation showed that the participants were very keen to see the result of their selection, and what the tool will recommend at the end. The "wow" expression could be seen when the report page of the tool was generated.

Ease of use

We evaluated the ease of use, focusing more on the process it took for the tool to actual help in making a decision. We added this evaluation because ease of use could affect how users see the usefulness of the tool, it serves as a moderating factor (Wei, 2009).

The result suggests that, 30% of the sample said that it was very easy to use the tool and understand the procedures, while 60% said it was somewhat easy, another 10% responded that it was not easy at all to use and understand. We interviewed the participants that said it wasn't easy to use and understand to know the reason behind their response, and most of them said it was easy to use and understand because the moderator was there to clarify somethings they don't understand, but if the moderator wasn't there, that would make it less easy. This feedback was well noted and crucial for further development of the software, since this was just a prototype. Generally, most of the participants found it easy to use and understand.

7.5.2 Understanding of the Intent of the DSS (E2)

This step evaluated the degree to which the artefact can be comprehended, if the participants had a clear understanding of the intent of the tool, that is, what it is meant to do. Participants were given cards with various words that implied understandability and were told to pick out those that they feel expressed their level of understanding. Based on content analysis, we found that all the participants fully understood the intent of the tool, some words like "I get the idea, I get the point, I comprehend etc., were used to express their opinions. From the interviews, participants gave reasons why they gave such responses, most of which were they being able to use the tool, and the tool helping them with making decisions, which they would naturally not have been able to make. Also, some of the CS criteria, which the tool highlighted, they would not have thought about them, but were important to consider before deciding to CS.

7.5.3 Applicability of the DSS (E3)

The last exercise performed by the participants was aimed at determining how applicable this tool is to research in general, if the participants found the tool applicable

to use in their research project, next project or for other new researchers. Using content analysis, we found that all the participants found it applicable, not necessarily to their immediate project, as some of them were rounding up, but to research in general. Words like “relevant, I can relate to it, suitable, etc.”, were used to describe their opinion.

Some of the participants were thrilled about the tool, as such wanted to discuss their project more, and how they could use CS in some of their testing activity to reduce the time it would cost them if they were to do that alone. The exercise encouraged the participants, experienced or not, to think carefully about testing activities that they thought could not be CS, how that could possibly be done, and to also consider the CS criteria, and how that affects the outcome of their decisions. Finally, the participants expressed their general feelings and opinion of the tool and exercise, which were mainly that, it was simple, great, good, interesting, etc. There was also considerable positive feedback about the card sorting exercises, as it provided an effective opportunity to compare their thoughts with the use of the tool.

7.6 DISCUSSION AND CONCLUSION

Gathering and analysing data to test a theory can be a daunting endeavour. The complexity of today’s world, as well as the sophistication of the research undertaken today, suggests researchers need to collect very large amounts of empirical data pertaining to complex environments, considering a multitude of factors, contextual elements and stakeholders, which require new, innovative approaches.

We developed the tool consisting of three main components: GUI, information processing component, and knowledge component. This was utilized in two prototype implementations: the template-based prototype and the user-interface based prototype. While the template-based prototype was used for knowledge articulation by representing, extracting, organizing and acting on relevant information, the DSS was targeted for decision making and recommendation for researchers using the tool. Therefore, the two prototypes make complementary contributions to research and practice.

The developed template and DSS help researchers to systematically check if crowdsourcing can be applied to a variety of theory testing patterns. So, our proposition is not just to check if data collection per se can be crowdsourced or not. Our proposition is to check if the patterned activities related to theory testing can be crowdsourced or not. The developed artefact provides constructs and methods, and design principles and rules that may be adopted by researchers in a variety of situations. It also contributes with operational knowledge highlighting how researchers may benefit from the crowdsourcing strategy.

Implementing these prototypes from a Design Science perspective represents a form of evaluation for the pattern model, which demonstrates its applicability (Peffers et al., 2012). Since the prototype helps with decision making, and can be used as a research tool, it contributes to crowdsourcing research.

Despite the potential values described above, we should also recognize some limitations of this research. One limitation to consider is that the decision to crowdsource may extend beyond the intrinsic characteristics of theory testing activities. For instance, researchers may wish to take into consideration other contextual elements influencing the research, such as institutional polices, culture etc. That introduces a contingency factor in the decision to crowdsource that complicates the decision-making process.

Furthermore, our procedure takes into consideration a set of properties that cannot be considered complete. Many other properties could eventually be added, some of them addressing operational issues such as the characteristics of specific platforms used to crowdsource. Other properties could consider characteristics of the crowd and a more comprehensive taxonomy of crowdsourcing tasks. Another issue to consider, which is illustrated in our examples, is that the development of task support may be considered along with the automation of some theory testing activities. However, automation was not considered in our research.

Finally, we should also consider that researchers may also find creative ways to overcome some of the constraints addressed by the crowdsourcing process, e.g. adopting games, role playing, etc. Therefore, we suggest that both the templates and the prototype described in this paper are just initial approaches to the endeavour of

bringing design science into theory building and theory testing, and also bringing crowdsourcing into theory testing. These limitations can serve as directions for future research.

7.7 REFERENCES

- Antunes, P., Herskovic, V., Ochoa, S., & Pino, J. (2012). Structuring Dimensions for Collaborative Systems Evaluation. *ACM Computing Surveys*, 44(2).
- Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of information technology*, 20(2), 67-87.
- Barrett, A., & Edwards, J. (1995). Knowledge Elicitation and Knowledge Representation in a Large Domain with Multiple Experts. *Expert Systems with Applications*, 8(1), 169-176.
- Bates, J., & Lanza, B. (2013). Conducting psychology student research via the Mechanical Turk crowdsourcing service. *North American Journal of Psychology*, 15(2), 385.
- Behrend, T., Sharek, D., Meade, A., & Wiebe, E. (2011). The viability of crowdsourcing for survey research. *Behavior research methods*, 43(3), 800.
- Bhattacherjee, A. (2012). Social science research: principles, methods, and practices.
- Bonney, R., Cooper, C., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K., & Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11), 977-984.
- Brunt, L., & Meidell, E. (2018). When Are Crowdsourced Data Truthful, Accurate, and Representative. *The Journal of Business Inquiry*, 17(1), 55-71.
- Cheng, J., Teevan, J., Iqbal, S., & Bernstein, M. (2015). Break it down: A comparison of macro-and microtasks *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 4061-4064): ACM.
- Creswell, J., & Creswell, J. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage.
- Crowston, K. (2012). Amazon mechanical turk: A research tool for organizations and information systems scholars *Shaping the Future of ICT Research. Methods and Approaches* (pp. 210-221): Springer.
- Derry, S., Pea, R., Barron, B., Engle, R., Erickson, F., Goldman, R., . . . Sherin, B. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *The Journal of the Learning Sciences*, 19(1), 3-53.
- Dicks, L. V., Walsh, J. C., & Sutherland, W. J. (2014). Organising evidence for environmental management decisions: a '4S'hierarchy. *Trends in ecology & evolution*, 29(11), 607-613.
- Enwereuzo, I., Antunes, P., & Johnstone, D. (2017). Towards the Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: Conceptual Framework and Model. Paper presented at ECIS.
- Gadiraju, U., Fetahu, B., & Kawase, R. (2015). Training workers for improving performance in crowdsourcing microtasks *Design for Teaching and Learning in a Networked World* (pp. 100-114). Cham: Springer.

- Ghezzi, A., Gabelloni, D., Martini, A., & Natalicchio, A. (2018). Crowdsourcing: a review and suggestions for future research. *International Journal of Management Reviews*, 20(2), 343-363.
- Gregor, S. (2006). The nature of theory in information systems. *MIS quarterly*, 611-642.
- Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 37(2), 337-355.
- Gura, T. (2013). Citizen science: Amateur experts. *Nature*, 496(7444), 259-261.
- Hevner, A., & Chatterjee, S. (2010a). *Design research in information systems: theory and practice* (Vol. 22): Springer Science & Business Media.
- Hevner, A., & Chatterjee, S. (2010b). Design science research in information systems *Design research in information systems* (pp. 9-22): Springer.
- Hevner, A., March, S. T., Park, J., & Ram, S. J. M. q. (2004). Design science in information systems research. 28(1), 75-105.
- Holsapple, C. W. (2008). DSS architecture and types *Handbook on Decision Support Systems* 1 (pp. 163-189): Springer.
- Hosack, B., Hall, D., Paradice, D., & Courtney, J. F. (2012). A look toward the future: decision support systems research is alive and well. *Journal of the Association for Information Systems*, 13(5), 315.
- Jarmolowicz, D., Bickel, W., Carter, A., Franck, C., & Mueller, E. (2012). Using crowdsourcing to examine relations between delay and probability discounting. *Behavioural processes*, 91(3), 308-312.
- Jiang, L., Wagner, C., & Nardi, B. (2015). *Not Just in it for the Money: A Qualitative Investigation of Workers' Perceived Benefits of Micro-task Crowdsourcing*. Paper presented at the 48th Hawaii International Conference on System Sciences.
- Kietzmann, J. (2017). Crowdsourcing: A revised definition and introduction to new research. *Business Horizons*, 60(2), 151-153.
- Kittur, A., Chi, E., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 453-456): ACM.
- Lee, J., & Bui, T. (2000). *A template-based methodology for disaster management information systems*. Paper presented at the System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on.
- Lowry, P. B., D'Arcy, J., Hammer, B., & Moody, G. D. (2016). "Cargo Cult" science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels. *The Journal of Strategic Information Systems*, 25(3), 232-240.
- Lynham, S. (2002). The general method of theory-building research in applied disciplines. *Advances in developing human resources*, 4(3), 221-241.
- Malone, T., & Crowston, K. (1994). The Interdisciplinary Study of Coordination. *ACM Computing Surveys*, 26(1), 87-119.
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251-266.
- McInnis, B., Cosley, D., Nam, C., & Leshed, G. (2016). Taking a HIT: Designing around rejection, mistrust, risk, and workers' experiences in Amazon Mechanical Turk *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2271-2282): ACM.
- Nakatsu, R., Grossman, E., & Iacovou, C. (2014). A taxonomy of crowdsourcing based on task complexity. *Journal of Information Science*, 40(6), 823-834.

- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153-163.
- Peffers, K., Rothenberger, M., Tuunanen, T., & Vaezi, R. (2012). *Design science research evaluation*. Paper presented at the International Conference on Design Science Research in Information Systems.
- Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45-77.
- Power, D. J. (2008). Decision support systems: a historical overview. *Handbook on Decision Support Systems* 1, 121-140.
- Prat, N., Comyn-Wattiau, I., & Akoka, J. (2015). A taxonomy of evaluation methods for information systems artifacts. *Journal of Management Information Systems*, 32(3), 229-267.
- Pries-Heje, J., & Baskerville, R. (2008). The design theory nexus. *MIS quarterly*, 731-755.
- Pries-Heje, J., Baskerville, R., & Venable, J. R. (2008). *Strategies for Design Science Research Evaluation*. Paper presented at the ECIS.
- Reason, J. (2008). *The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries*. Surrey, England: Ashgate.
- Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, 4(2), 155-169.
- Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS quarterly*, 37-56.
- Shank, D. B. (2016). Using crowdsourcing websites for sociological research: The case of Amazon Mechanical Turk. *The American Sociologist*, 47(1), 47-55.
- Simon, H. A. (1960). The new science of management decision. *The Ford distinguished lectures: Vol. 3*. New York, NY, US: Harper & Brothers. <http://dx.doi.org/10.1037/13978-000>.
- Sonnenberg, C., & vom Brocke, J. (2011). *Evaluation patterns for design science research artefacts*. Paper presented at the European Design Science Symposium.
- Steelman, Z., Hammer, B., & Limayem, M. (2014). Data Collection in the Digital Age: Innovative Alterantives to Student Samples. *Mis Quarterly*, 38(2), 355-378.
- Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing Samples in Cognitive Science. *Trends in cognitive sciences*, 21(10), 736-748.
- Thomas, D. (2006). A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, 27(2), 237-246.
- Vondrick, C., Patterson, D., & Ramanan, D. (2013). Efficiently scaling up crowdsourced video annotation. *International Journal of Computer Vision*, 101(1), 184-204.
- Wei, W.-C. (2009). A technology acceptance model: Mediate and moderate effect. *Asia Pacific Management Review*, 14(4), 461-476.
- Witschey, J., Murphy-Hill, E., & Xiao, S. (2013). Conducting interview studies: Challenges, lessons learned, and open questions *1st International Workshop on Conducting Empirical Studies in Industry* (pp. 51-54): IEEE.

Chapter 8

8 DISCUSSION AND CONCLUSIONS

This chapter discusses the research outcome considering its main contributions, limitations and concluding remarks. The chapter starts with the discussion section, which covers the contributed artefacts, research contributions and limitations, while the concluding remarks and the future work are discussed in the latter section of this chapter.

8.1 DISCUSSION

This research adopts the design science research paradigm in exploring the adoption of crowdsourcing in theory testing. The DSR framework in figure 1.1 guided the research process across the whole thesis. We highlight the systematic and transparent characteristics of the process. The systematic characteristic is reflected through our justification of every research step used in the process. In particular, the systematic procedure of analysing and synthesizing a body of knowledge (Paré et al., 2015; Paré et al., 2016). The transparency characteristic is reflected in all the paper, where the research activities and decisions were explained as explicitly as possible.

We adapt the SCOA method developed by Thuan et al. (2016a) for DSS development in emerging areas. We consider the incorporation of CS in theory testing as an emerging area, as much research has not been done on this domain. This method is highly applicable to the development of our DSS, as the decision-support tasks involves first consolidating domain knowledge of both theory testing and CS for better decision-making (Nemati et al., 2002). It also suggests the role of design science as it emphasizes a rigorous approach to the advancement of knowledge on design and development. Design science is considered to be an appropriate paradigm to address the development of innovative artefacts and to address unstructured issues, which is also the major target of DSS research in emerging areas, as is our focus. The SCOA method represents a heuristic way that grounds experience and expertise knowledge in low-theoretical DSS domains.

8.1.1 Contributed Artefacts

The primary outcome of this research consists of various artefacts. The artefacts examined the concepts, relationships and attributes of the domain of the study (theory testing and crowdsourcing), aligning to the continuum from analysis to synthesis designs suggested and adapted from Miles et al. (2013). Figure 8.1 illustrates this continuum. The first artefact explored and conceptualized theory testing, which resulted in a framework. The second artefact extended this conceptualization by organizing the concepts, and using such concepts to scope the literature, of which a pattern model was developed. The final artefact instantiated a decision tool founded on previous artefacts, and then was evaluated to confirm the utility of the tool.

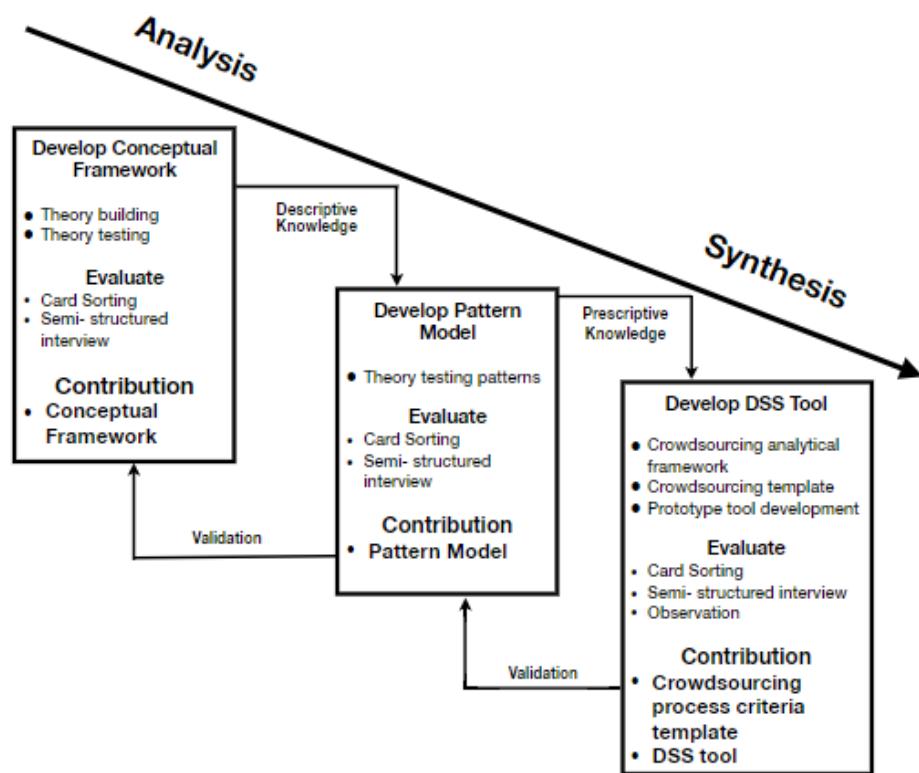


Figure 8.1. Interrelated yet different artefact outcomes

Seen together, the artefacts form three levels of abstraction, (depicted in Table 8.1). The conceptual framework presents abstract building blocks for the research, by laying a foundation of the domain area. The pattern model acts upon these building blocks for further literature scoping, breaking down the domain area into elements, activities, processes, data and their relationships. The decision tool operationalises the elements

of the pattern model into decision templates and contextual recommendations. Considering the three levels of abstraction independently, it is feasible for different researchers to focus any of these levels and gain independent understanding of theory testing with crowdsourcing.

It is also important to note that it is possible to trace back the decision tool through the research stages (see figure 8.2). Simply put, the operationalized knowledge in the decision tool can be rightly traced back to the pattern model elements, which can be mapped to the components of the conceptual framework, and in turn traced back to the various knowledge sources. The systematic approach brought by design science research makes this traceability possible. We systematically structure the research

Knowledge Base	Outcome Artefacts	Contribution Types
More abstract  More specific	Conceptual framework	Conceptual model
	Pattern Model	Process model
	Decision tool	Instantiation

Figure 8.2. Research design products (adapted from (Baskerville et al., 2018; Gregor & Hevner, 2013))

activities similar to the evidence-based strategy in design science research (Denyer & Tranfield, 2006; Van Aken, 2005; Van Aken & Romme, 2012) to explicitly justify and present key decisions and developments made.

We highlight the innovativeness of the artefacts. The development of the conceptual framework displays innovativeness by conceptualizing the theory development cycle. This form of conceptualization has not been expressed this way, showing its uniqueness while providing clarity. Innovativeness is also seen in the pattern model, as it integrates the whole theory testing process into a pattern, through methodological exploration of literature. The development of the decision tool is the first-of-its-kind in the theory testing domain. The design, development and introduction of these artefacts

into the research community also adds prescriptive knowledge contributions (Baskerville et al., 2018).

To ensure that all research questions are answered, we structure and summarize our result in a tabular format (see table 8.1). The structure of the table has four columns, the first column shows the developmental stages of the research, while the second and third column presents the research activities and outcomes respectively. The last column highlights the research questions being answered.

Table 8.1. Summary of research activity and results answering research questions

Stages	Research development stage	Research Activity	Research Outcome	RQ
1	Literature review	Theory, Theory building and Theory testing (TT) Knowledge base <ul style="list-style-type: none"> Identified different types of theory Analysed different types of theory, identified theories that can be operationalized 	<ul style="list-style-type: none"> Coverage of the domain concepts and relationships Clarity of the domain and selection of theories that can be operationalized 	RQ1
2	Artefact Development	Conceptual Model <ul style="list-style-type: none"> Contextualize the conceptual elements of theory testing Contextualize theory building and theory testing process Distil theory testing from building Identify boundary conditions 	<ul style="list-style-type: none"> Coverage of theory testing steps and process Re-enforcing the relationship between theory building and TT Contextualize and conceptualizing of theory generation cycle Integrated knowledge from two concepts (theory building and testing) to generate boundary activities 	RQ1
3	Literature review	Identify challenges researchers face doing research with human subjects in theory testing <ul style="list-style-type: none"> Synthesized relevant literature to identify challenges faced by theory testing researchers Conducted a problem frame assessment to get feedback from participants (Researchers) 	<ul style="list-style-type: none"> Identified challenges Prioritized the challenges in the order of the most challenging to least challenging Validation of research problem 	RQ1
		Preliminary source (MISQ 10-year period) <ul style="list-style-type: none"> Identified and analysed 248 knowledge sources 	<ul style="list-style-type: none"> Identified 52 different activities 	RQ1 and RQ2

		<ul style="list-style-type: none"> • Synthesized TT processes • Synthesized various TT elements, activities and relationships <p>Extended Knowledge base source (“Basket of 8” IS Journals – 2-year period)</p> <ul style="list-style-type: none"> • Identified and analysed 404 knowledge sources • Synthesized TT processes • Synthesized various TT elements, activities and relationships 		
4	Artefact Development	<p>Pattern Model</p> <ul style="list-style-type: none"> • Pattern capture <ul style="list-style-type: none"> ◦ Analysed the knowledge sources in detail ◦ Identified TT elements – concepts, activities, relationships • Pattern organization <ul style="list-style-type: none"> ◦ Synthesized the pattern elements ◦ Organized the elements into patterns interlinking relationships 	<ul style="list-style-type: none"> • Merged closely related identified activities into groups, obtained 32 TT activities <ul style="list-style-type: none"> • 1st Model <ul style="list-style-type: none"> ◦ Segmented processes and elements into patterns having– Intents segment and structure segment, and into two categories – data gathering and data analysis ◦ Developed 1st phase of the pattern model • 2nd Model <ul style="list-style-type: none"> ◦ Extended and re-modelled the previously developed model for simplicity and clearer understanding. ◦ Categorised activities into two – primary activities and secondary activities ◦ Developed pattern model with interrelated activities 	
5	Literature review	<p>Crowdsourcing Knowledge Base</p> <ul style="list-style-type: none"> • Identified and synthesized relevant literature • Identified CS requirements, elements and criteria 	<ul style="list-style-type: none"> • Obtained 4 CS processing steps and 7 criteria 	RQ2

		<ul style="list-style-type: none"> Synthesized the result and identified those relevant to TT 		
6	Instantiated artefact Development	<p>Decision Tool</p> <ul style="list-style-type: none"> Based on the pattern model and crowdsourcing criteria Developed two prototypes <ul style="list-style-type: none"> The first was used for knowledge articulation by representing, extracting and organizing relevant information and to obtain feedback. The second was developed based on feedback, for decision support and recommendations. 	<ul style="list-style-type: none"> A decision tool with the main aim of supporting TT researchers with the decision to CS any theory testing activity. 	RQ2

8.1.2 Contributions to Theory

This dissertation being a design science endeavour, contributes knowledge. The research was about exploratory knowledge, which is anchored on a set of artefacts. We discuss three main contributions in the following sections.

8.1.2.1 The significance of the Conceptual Framework

The conceptual framework plays an important role in guiding IS research. It gives a critical, holistic perspective of the research by setting out concepts and explanations needed to make sense of a domain area and defining boundaries. Our framework serves as an abstract representation of the theory testing domain. It serves as a means of structuring the interconnection between theory building and theory testing.

We conceptualize the theory testing process as a cycle based on a set of building blocks. We identified two main activities – data collection and data analysis each with its own secondary activities.

The innovativeness in addressing the formative validity of theory development, uniquely contributes to the knowledge base by giving clarity on the interconnection of theory building and theory testing. The conceptualization achieved three things: 1) clearly distinguishes the two key components of theory development: theory building and theory testing; 2) highlights how these two components are tightly interconnected in an iterative cycle designed to learn more about (and perhaps improve) the theory; and 3) highlights the boundary activities that lie between these two components.

8.1.2.2 The Significance of the Pattern Model

Having introduced the main concepts of theory testing, the dissertation proposes a pattern model that offers knowledge structure around this concept. The pattern model provides various benefits in the theory testing domain. We consider these benefits from three main research perspectives: IS, design science and DSS.

Models have played an important role in representing the IS domain (Wand & Weber, 1995; West & Stowell, 1999). The pattern model aids problem and solution understanding and represents the connection between the problem and solution components. It structures the domain by presenting relational links to various theory testing activities, **offering a scaffold for understanding basic processes and concept of theory testing.** We can classify the representation of the pattern model in three aspects: clarity and coverage (Van Der Valk et al., 2007).

The pattern model has high clarity contributing to the understanding of the domain. It defines not only the domain concepts but also activities and relationships, which increases understanding in the domain. Using various knowledge sources, the pattern model brings out clarity and a holistic view of the domain. The pattern model also has a high coverage of the domain concepts, activities and relationships. Adopting the grounded approach, in which elements freely emerged from the various knowledge sources, the pattern model offers diversity and broad coverage of the domain.

The role of models is highlighted in design science, which is considered as a significant paradigm in IS research. Design science suggests the contributions of models for building knowledge bases (Hevner et al., 2004). The pattern model contributes to the theory testing knowledge base. It builds the knowledge base through

structuring key concepts, elements, activities and their relationships based on exploration of knowledge sources, from which knowledge can be inferred. Furthermore, the knowledge base role of the pattern model has been clearly seen when the pattern model formed the basis for the construction of the tool. According to Hevner and Chatterjee (2010), founding artefact construction is one of the distinct characteristics of knowledge bases. Having said this, we note that the knowledge offered by the pattern model is not only limited to the construction of the tool but can be used in various methodological context for the understanding of the domain, and for further construction and remodelling in various fields.

We finally consider the pattern model from the DSS (decision support system) perspective. The pattern model played a very important role in the development of the DSS. It served as a means of acquiring needed concepts and elements needed in the development of the DSS, providing various decision alternatives. The pattern model is one of the main building blocks of the DSS, without which the DSS will not function. It provides decision alternatives which was operationalized in the decision tool. These roles suggest the value of the pattern model.

8.1.2.3 The Significance of the Decision Tool

Decision tools can play important roles in decision making. They may help overcome barriers to good decision making by providing necessary decision support, carrying out in-depth analysis and suggesting possible courses of action or proffer recommendations (Asemi et al., 2011).

The tool which is an experiment, is an instantiated artefact, according to Gregor & Hevner (2013). As an instantiated artefact, the tool operationalizes the pattern model into computer-based prototypes based on identified CS criteria. This is a first-of-a-kind tool. Although we note that CS has been used in some part of theory testing but has not been considered in the whole theory testing cycle. Our tool extends the incorporation of CS in theory testing by providing support for the whole theory testing cycle. The assessment done, highlights the need for such DSS, and its importance. The actual development of this tool is a major contribution to the theory testing domain.

The successful construction of the tool proves that it is possible to integrate and adopt crowdsourcing in the theory testing process and can actually be operationalized. Furthermore, the tool which was developed based on the pattern model, has demonstrated the feasibility of the pattern model, that is, the pattern model can be implemented and operationalized. We also consider that the decision tool serves as an instantiation artefact (Hevner et al., 2004), providing a means to support its users in making a decision on adopting crowdsourcing for various theory testing activities.

Considering utility, the findings suggest that using the tool may lead to better decision making and may create an awareness of the effects of different CS criteria before making a final decision. We note that the tool does not make a decision for the user but supports the user in decision making. The card sorting exercise, interviews and observation results suggest that the tool is beneficial and provides additional information for making informed decisions. When using the tool, it is also found that participants have a positive perception towards ease of use, while giving some good suggestions for improvement.

In general, the conceptual framework, pattern model and decision tool are innovative artefacts that we contribute to the theory testing domain. We note that these artefacts can be used either as separate artefacts, or as a set of artefacts in capturing theory testing knowledge from abstract to concrete as illustrated in figure 8.1. Therefore, whether separately or as a set, their use is beneficial and guides researchers, especially upcoming researchers on the theory testing process. Overall, the importance of the dissertation relies not only on theoretical efforts, but also on having empirical and qualitative evidence as well.

8.1.3 Contributions to Practice

Our study contributes to research practice. Researchers can use the conceptual framework and pattern model as blueprint for conceptualizing, analysing and structuring their theory testing process. The pattern model provides methodological contributions by summarizing the theory testing process of various researchers into a pattern. Together this gives a clearer picture of the theory testing process.

Another practical contribution comes from the proposed CS criteria framework and the set of decision templates. They support decision makers to evaluate whether crowdsourcing is an appropriate strategy to incorporate in each theory testing activity. The templates and CS criteria framework guides decision makers on what criteria should be considered when considering the use of crowdsourcing in theory testing. As a result, researchers can use these templates and tool as a practical guide and means to measure the possibility of adopting crowdsourcing for any of the theory testing activity.

Finally, the study provides a computer-based tool for incorporating crowdsourcing in theory testing. The tool structures concepts, elements, activities and relationships, which supports researchers in their decision-making process. This practical support is highlighted through the evaluation done, where the results show that the tool can help in decision making and give beneficial recommendations. This tool can also be beneficial to platform developers, especially research-based platforms, by examining the tool, platform developers can integrate these basic CS criteria and support to assist their users.

8.1.4 Limitations of the Research

This dissertation inevitably has some limitations when viewed through a critical lens. First, we highlight that the method has not been thoroughly evaluated, as this study focused on exploring the design and establishment of incorporating CS in the theory testing process. Evaluations done were focused on generating valuable feedback from participants to help improve and design the artefacts. Therefore, further evaluation using other forms of evaluation method is needed to increase external validity and to validate the utility of the developed artefacts.

Another limitation relates to the development of the decision tool. The tool was targeted at a level of evaluation and demonstration as a proof of concept. Future research could implement the tool by applying software engineering methods for better product, thereby improving the usability of the tool.

8.1.5 Future Work

This dissertation creates paths for further exploration to the general research field on theory testing and adoption of crowdsourcing in theory testing.

Future research should aim to move the knowledge provided by the artefacts built in this study forward to a higher level of abstraction, which according to Gregor & Hevner (2013), can be done by generalizing the proposed artefacts. For instance, the models and tool can be applied in different contexts and research fields. Therefore, future research should further apply these artefacts and knowledge to various research fields that develop and test theory, other than IS discipline, which will show its application principles.

Further experiments with the tool could be carried out to validate the utility of the tool. This would increase its external validity. Such experiments could also be conducted in other fields that test theory outside the IS discipline, creating avenue for other discipline to benefit from this tool. Another area for further work is on the improvement of the decision tool. We note that the tool has some work to be done on it, to make it better in its performance and outlook, increasing efficiency and usability. Which when done can be evaluated based on such criteria.

Finally, the evaluation method used, card sorting is not yet popular in IS evaluations, and as such might not seem like an appropriate method of evaluation. In the course of this study, our use of card sorting has created an enlightenment of the numerous benefits card sorting brings. Developing this evaluation method and adding it to the various established evaluation strategies in DSR would be beneficial to the research community.

8.2 CONCLUDING REMARKS

The pivotal importance of theory has been continuously emphasized in the IS discipline since its inception (Mueller & Urbach, 2013). A good theory is one that is plausible, having a quality criteria of being falsifiable. To determine a theory's falsifiability, it has to be tested.

Since theory testing is of such importance, this study considers the challenges faced by IS theory testing researchers, especially when human participation is involved. Some of these challenges were highlighted in the study. We note that some recommendations and support have been suggested by researchers, nonetheless, this study explores a new and innovative way to tackle some of these challenges by considering the crowdsourcing strategy.

Crowdsourcing could be a useful strategy to adopt for theory testing because of the numerous benefits it offers. CS is a powerful tool that expands the reach of researchers, improving external validity (Ramsey et al., 2016). CS permits recruitment of geographically and culturally diverse participants, enhancing external validity (Ramsey et al., 2016). Research outputs produced through CS have quality and are reliable (Shank, 2016). Due to lack of face-to-face contact with participants, some experimenter bias and social desirability can be mitigated (Paolacci et al., 2010), cheap and fast online participants can also be gotten through CS (Shank, 2016). This research therefore focused on the feasibility of adopting the CS strategy in theory testing.

Our research has led us to consider the viability of testing IS theories using crowdsourcing. In carrying out this research we have brought to light the underlying concepts of theory building and theory testing in the IS domain. We discovered through a preliminary assessment that some researchers were 1) not very familiar with theory testing concepts and its associated activities and 2) in some cases, getting participants for testing activities was a challenge.

We set out to address these issues. First, we developed a conceptual model on theory testing. Drawing out on wide literature reviews on the patterns or activities used in theory testing across the IS domain. Second, we identified the concepts and building blocks of crowdsourcing and operationalized these concepts to help address the problem of getting participants for an IS theory testing undertaking. In addressing these problem-solving research objectives, we used design science as our methodology. This is because design science is focused on the iterative design of IS artefacts, models and frameworks, and supports exploratory, problem-solving research.

In writing this thesis, we've opted to present our research findings in published papers (table 8.2). Although the papers are related, they are independent pieces of work, we note that sometimes they intercept, which is more like an improvement to previous work, while sometimes they leave gap, of which the next paper fills up.

In paper 1, we undertook the problem of understanding how to align theory testing with crowdsourcing and given the different types of theory, what type can be operationalized. Our findings provided some insights into theory testing patterns, various pathways researchers can take to achieve the same theory testing goal and what pathways can be fully or partially crowdsourced.

In paper 2, we looked at problems associated with distinguishing theory building component from theory testing, identification of development method to be used in the entire study and explored the idea of developing a DSS to incorporate CS in the theory testing process while also exploring the challenges researchers face and acquiring user requirements and priorities. Our results extended the development methods of SCOA and DSS DSR and identified the problems associated with distinguishing between theory building and theory testing components. We provided some insights to PhD researchers on how to incorporate CS in the theory testing process.

Paper 3 presented us with a new challenge around the complexities of theory testing process and the development of a framework to help with the understanding and process of theory testing. Our result showed the development of an operational view to theory testing as an independent phenomenon and how it contributes to research and practice. We also provided patterns that can be used when testing a theory which in turn contributes to design science knowledge.

In Paper 4, we addressed the question of how we develop a Decision support system (Sein et al.) based upon the findings of our previous papers. In solving this research problem, we provided insights to how researchers can benefit from the crowdsourcing strategy and provided the design principles that can be adopted by researchers in different situations. Paper 5 addressed developing and accessing the steps to the creation of the DSS. As a result, we developed artefacts and presented various possibilities of leveraging on crowdsourcing in developing innovative ways to test or validate theories in IS. And finally, paper 6 gave us a novel artefact as a DSS

contribution to IS research by addressing the problem of developing a DSS prototype for decision to crowdsource theory testing in IS.

Table 8.2. Summary of research objectives and contributions

Papers	Identified problem	Research Objectives (RO)	Research Questions (RQ)	What has been done	Contributions
Paper 1	<ul style="list-style-type: none"> • How to align theory testing with crowdsourcing 	<ul style="list-style-type: none"> • Identify, condense and make sense of the conceptual elements of theory testing (RO1) • Undertake a systematic review of how theories in IS domain have been tested (RO2) 	• RQ1	<ul style="list-style-type: none"> • Preliminary identification and analysis of theory testing steps and elements from various literature • Identification of theory types that can be operationalized and using that knowledge for the literature review. • Preliminary descriptive literature review (MISQ – 10 years) • Initial pattern model development • Identification of some crowdsourcing requirements and concepts 	<ul style="list-style-type: none"> • Insights into theory testing patterns offering decision alternatives researchers can take to achieve same goal. • Some insight as to the various patterns that can be fully/partially crowdsourced
Paper 2	<ul style="list-style-type: none"> • Problems associated with distinguishing theory building from theory testing. • Identification of development method to be used in the entire study. 	<ul style="list-style-type: none"> • Identify conceptual elements of theory testing (RO1) • Initial consideration of developing a DSS (RO4) 	• RQ1	<ul style="list-style-type: none"> • Identification and adaptation of the development method • Initial evaluations to get a clear picture of the identified problems 	<ul style="list-style-type: none"> • Identifying and confirming the problems associated with distinguishing theory building component from theory testing. Hence providing some insight to the underexposure of new researchers (PhD) to theory

	<ul style="list-style-type: none"> Initial idea of developing a DSS to incorporate CS in the theory testing process. Acquiring user requirements and priorities. 				<ul style="list-style-type: none"> testing, when compared to theory building Insights to researchers need of a DSS to incorporate CS in the theory testing process and priority features
Paper 3	<ul style="list-style-type: none"> Challenges faced by IS researchers when testing theory Complexities of the theory testing processes and the development of a framework to help alleviate that. How to make the theory testing process and activities easy to understand 	<ul style="list-style-type: none"> Identify conceptual elements of theory testing (RQ1) Undertake a systematic review of how theories in IS domain have been tested (RQ2) Profile different ways in which theories have been tested using a set of patterns (RQ3) 	RQ1	<ul style="list-style-type: none"> Development of a conceptual framework Development of pattern model 	<ul style="list-style-type: none"> Operational view to theory testing as an independent phenomenon. Insights to different patterns that can be adopted when testing a theory.

Paper 4	<ul style="list-style-type: none"> • Acquiring crowdsourcing requirements and features necessary for the development of the DSS 	<ul style="list-style-type: none"> • Develop an approach for assessment of what theory testing activity can be crowdsourced or not (RO4) 	• RQ2	<ul style="list-style-type: none"> • Identification of CS requirements, features and attributes necessary for theory testing • Crowdsourcing analytic framework for the assessment of what theory testing activity can be crowdsourced or not, and the level of difficulty involved. 	<ul style="list-style-type: none"> • Insights to how researchers can benefit from the crowdsourcing strategy • Design principles and rules that can be adopted by researchers in a wide range of situations.
Paper 5	<ul style="list-style-type: none"> • Acquiring crowdsourcing process criteria necessary for the development of the DSS • Implementation and integration of CS criteria with theory testing pattern into decision templates 	<ul style="list-style-type: none"> • Identify conceptual elements of theory testing (RO1) • Undertake a systematic review of how theories in IS domain have been tested (RO2) • Profile different ways in which theories have been tested using a set of patterns (RO3) • Identify crowdsourcing criteria necessary to make the decision to crowdsource (RO4) 	• RQ2	<ul style="list-style-type: none"> • Consolidation of the previously developed artefact • using the building blocks identified in the conceptual model as a foundation and knowledge base for the next artefact development • Identification and synthesizing of crowdsourcing criteria necessary to make the decision to crowdsource • Development of crowdsourcing templates based on the identified criteria • Evaluations 	<ul style="list-style-type: none"> • Insights to how crowdsourcing can alleviate the whole theory testing lifecycle • Leveraging the possibilities of crowdsourcing in developing innovative ways to validate theory contributes to theory development • Developed artefacts contributes to IS research
Paper 6	<ul style="list-style-type: none"> • Development of a DSS prototype for 	<ul style="list-style-type: none"> • Develop a DSS that helps IS researchers making decisions 	• RQ2	<ul style="list-style-type: none"> • Development of a decision tool based on the developed templates using user's 	<ul style="list-style-type: none"> • Development of a novel artefact as a

	decision to CS theory testing activities	about how to integrate CS in testing IS theory (RO5)		requirement and identified CS criteria	DSS contributes to IS research
--	--	--	--	--	--------------------------------

In conclusion of this dissertation, it is clear that the crowdsourcing strategy could be adopted in the theory testing process. This can serve as a support to mitigate some of the challenges researchers face during when testing their theory. By establishing this support strategy, researchers can take full advantage of the strategy, not just for surveys, but other aspects of theory testing. The developed artefacts provide solid knowledge that researchers, especially upcoming ones can take advantage of.

8.3 REFERENCES

- Asemi, A., Safari, A., & Zavareh, A. A. (2011). The role of management information system (MIS) and Decision support system (DSS) for manager's decision making process. *International Journal of Business and Management*, 6(7), 164-173.
- Baskerville, R., Baiyere, A., S., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems*, 19(5), 358-376.
- Denyer, D., & Tranfield, D. J. M. D. (2006). Using qualitative research synthesis to build an actionable knowledge base. 44(2), 213-227.
- Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 37(2).
- Hevner, & Chatterjee. (2010). Design science research in information systems *Design research in information systems* (pp. 9-22): Springer.
- Hevner, A., March, S. T., Park, J., & Ram, S. J. (2004). Design science in information systems research. 28(1), 75-105.
- Miles, M., Huberman, A., & Saldana, J. (2013). Qualitative data analysis: A methods sourcebook: Thousand Oaks, CA: Sage Publications Inc.
- Mueller, B., & Urbach, N. (2013). *The Why, What and How of Theories in Information Systems Research*. Paper presented at the 34th International Conference on Information Systems (ICIS).
- Nemati, H. R., Steiger, D. M., Iyer, L. S., & Herschel, R. (2002). Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing. 33(2), 143-161.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgement and Decision Making*, Vol. 5, No. 5.
- Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. 25(6), 493-508.
- Paré, G., Trudel, M.C., Jaana, M., Kitsiou, S., & Management. (2015). Synthesizing information systems knowledge: A typology of literature reviews. 52(2), 183-199.

- Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. (2016). Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior*, 58, 354-360.
- Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS quarterly*, 37-56.
- Shank, D. B. (2016). Using crowdsourcing websites for sociological research: The case of Amazon Mechanical Turk. *The American Sociologist*, 47(1), 47-55.
- Thuan, N. H., Antunes, P., & Johnstone, D. (2016). A Design Science Method for Emerging Decision Support Environments. *arXiv preprint arXiv:1605.04725*.
- Van Aken. (2005). Management research as a design science: Articulating the research products of mode 2 knowledge production in management. *British journal of management*, 16(1), 19-36.
- Van Aken, & Romme, A. G. (2012). A design science approach to evidence-based management. 43-57.
- Van Der Valk, T., Van Driel, J. H., & De Vos, W. (2007). Common characteristics of models in present-day scientific practice. 37(4), 469-488.
- Wand, Y., & Weber, R. (1995). On the deep structure of information systems. *Information Systems Journal*, 5(3), 203-223.
- West, D., & Stowell, F., Artech, London. (1999). Models and diagrams and their importance to information systems analysis and design. 295-311.

REFERENCES

- Adzeh, K. J. (2014). The challenge of conducting a quantitative business research: Analysis of issues with survey design, sampling, validity, and reliability. *Research Gate*.
- Agee, J. (2009). Developing qualitative research questions: a reflective process. *International Journal of Qualitative Studies in Education*, 22(4), 431-447.
- Aken, J. E. v. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of management studies*, 41(2), 219-246.
- Alexander, C. (1999). The origins of pattern theory: The future of the theory, and the generation of a living world. *IEEE software*, 16(5), 71-82.
- Amazon. (2011). *Requester Best Practices Guide*. Amazon Web Services.
- Andriole, S. J. (2010). Business impact of Web 2.0 technologies. *Communications of the ACM*, 53(12), 67-79.
- Antunes, P., Herskovic, V., Ochoa, S., & Pino, J. (2012). Structuring Dimensions for Collaborative Systems Evaluation. *ACM Computing Surveys*, 44(2), 1-28.
- Argyris, C. (1996). Actionable knowledge: Design causality in the service of consequential theory. *The Journal of Applied Behavioral Science*, 32(4), 390-406.
- Arnott, D., & Pervan, G. (2005). A critical analysis of decision support systems research. *Journal of information technology*, 20(2), 67-87.
- Arnott, D., & Pervan, G. (2012). Design science in decision support systems research: An assessment using the Hevner, March, Park, and Ram Guidelines. *Journal of the Association for Information Systems*, 13(11), 923.
- Asemi, A., Safari, A., & Zavareh, A. A. (2011). The role of management information system (MIS) and Decision support system (DSS) for manager's decision making process. *International Journal of Business and Management*, 6(7), 164-173.
- Ashton, S. J. N. R. (2014). Researcher or nurse? Difficulties of undertaking semi-structured interviews on sensitive topics. 22(1).
- Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of management review*, 14(4), 496-515.
- Barrett, A., & Edwards, J. (1995). Knowledge Elicitation and Knowledge Representation in a Large Domain with Multiple Experts. *Expert Systems with Applications*, 8(1), 169-176.
- Baskerville, R., Baiyere, A., , S., Hevner, A., & Rossi, M. (2018). Design Science Research Contributions: Finding a Balance between Artifact and Theory. *Journal of the Association for Information Systems*, 19(5), 358-376.
- Bates, J., & Lanza, B. (2013). Conducting psychology student research via the Mechanical Turk crowdsourcing service. *North American Journal of Psychology*, 15(2), 385.
- Behrend, T. S., Sharek, D. J., Meade, A. W., & Wiebe, E. N. (2011). The viability of crowdsourcing for survey research. *Behavior research methods*, 43(3), 800.
- Bernstein, M. S., Brandt, J., Miller, R. C., & Karger, D. R. (2011). *Crowds in two seconds: Enabling realtime crowd-powered interfaces*. Paper presented at the Proceedings of the 24th annual ACM symposium on User interface software and technology.

- Bhattacherjee, A. (2012). Social science research: principles, methods, and practices. *Textbooks Collection, Book 3*(http://scholarcommons.usf.edu/oa_textbooks/3).
- Bichler, M., Frank, U., Avison, D., Malaurent, J., Fettke, P., Hovorka, D., . . . Thalheim, B. (2016). Theories in business and information systems engineering. *Business & Information Systems Engineering*, 58(4), 291-319.
- Bider, I., Johannesson, P., & Perjons, E. (2013). Design science research as movement between individual and generic situation-problem-solution spaces *Designing Organizational Systems* (pp. 35-61). Berlin, Heidelberg: Springer.
- Biemann, C. (2013). Creating a system for lexical substitutions from scratch using crowdsourcing. *Language Resources and Evaluation*, 47(1), 97-122.
- Bitektine, A. (2008). Prospective case study design: qualitative method for deductive theory testing. *Organizational Research Methods*, 11(1), 160-180.
- Blagojević, V., Bojić, D., Bojović, M., Cvetanović, M., Đorđević, J., Đurđević, Đ., . . . Milutinović, V. (2017). Chapter One-A Systematic Approach to Generation of New Ideas for PhD Research in Computing. *Advances in Computers*, 104, 1-31.
- Bonney, R., Cooper, C. B., Dickinson, J., Kelling, S., Phillips, T., Rosenberg, K. V., & Shirk, J. (2009). Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11), 977-984.
- Borgman, C., Wallis, J., & Enyedy, N. (2007). Little science confronts the data deluge: habitat ecology, embedded sensor networks, and digital libraries. *International Journal on Digital Libraries*, 7(1-2), 17-30.
- Brabham. (2008). Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. *Convergence: The International Journal of Research into New Media Technologies*, 14(1), 75-90. doi:10.1177/1354856507084420
- Brabham, D. C. (2008b). Moving the crowd at iStockphoto: The composition of the crowd and motivations for participation in a crowdsourcing application. *First monday*, 13(6).
- Brabham, D. C. (2009). Crowdsourcing the public participation process for planning projects. *Planning Theory*, 8(3), 242-262.
- Brabham, D. C. (2010). Moving the crowd at Threadless: Motivations for participation in a crowdsourcing application. *Information, Communication & Society*, 13(8), 1122-1145.
- Brunt, L., & Meidell, E. (2018). When Are Crowdsourced Data Truthful, Accurate, and Representative. *The Journal of Business Inquiry*, 17(1), 55-71.
- Burton-Jones, A., & Volkoff, O. (2017). How can we develop contextualized theories of effective use? A demonstration in the context of community-care electronic health records. *Information Systems Research*, 28(3), 468-489.
- Callison-Burch, C. (2009). *Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk*. Paper presented at the Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1.
- Campbell, J. (1990). The role of theory in industrial and organizational psychology. In Dunnette MD, Hough LM (Eds.), *Handbook of industrial and organizational psychology*, Vol. 1(2nd ed.), pp. 39-73.
- Cheng, J., Teevan, J., Iqbal, S., & Bernstein, M. (2015). Break it down: A comparison of macro-and microtasks *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 4061-4064): ACM.

- Chernova, S., DePalma, N., Morant, E., & Breazeal, C. (2011). *Crowdsourcing human-robot interaction: Application from virtual to physical worlds*. Paper presented at the RO-MAN, 2011 IEEE.
- Cleven Anne, Gubler, P., & Hüner, K. M. (2009). *Design alternatives for the evaluation of design science research artifacts*. Paper presented at the Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology.
- Colquitt, J., & Zapata-Phelan, C. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal*, 50(6), 1281-1303.
- Conley, C., & Tost-Kharas, J. (2014). Crowdsourcing content analysis for managerial research. *Management Decision*, 52(4), 675-688.
- Coplien, J., & Schmidt, D. (1995). *Pattern Language of Program Design*. ACM Press/Addison-Wesley Publishing Co..
- Creswell, J., & Creswell, J. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*: Sage.
- Crowston, K. (1997). A coordination theory approach to organizational process design. *Organization Science*, 8(2), 157-175.
- Crowston, K. (2012). Amazon mechanical turk: A research tool for organizations and information systems scholars *Shaping the Future of ICT Research. Methods and Approaches* (pp. 210-221): Springer.
- De Boer, V., Hildebrand, M., Aroyo, L., De Leenheer, P., Dijkshoorn, C., Tesfa, B., & Schreiber, G. (2012). *Nichesourcing: Harnessing the power of crowds of experts*. Paper presented at the International Conference on Knowledge Engineering and Knowledge Management.
- De Vaus, D. (2013). *Surveys in social research*: Routledge.
- Denyer, D., & Tranfield, D. J. M. D. (2006). Using qualitative research synthesis to build an actionable knowledge base. 44(2), 213-227.
- Derry, S., Pea, R., Barron, B., Engle, R., Erickson, F., Goldman, R., . . . Sherin, B. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *The Journal of the Learning Sciences*, 19(1), 3-53.
- Dicks, L. V., Walsh, J. C., & Sutherland, W. J. (2014). Organising evidence for environmental management decisions: a '4S'hierarchy. *Trends in ecology & evolution*, 29(11), 607-613.
- Enwereuzo, I., Antunes, P., & Johnstone, D. (2017). Towards the Development of a DSS Supporting the Integration of Crowdsourcing in Theory Testing: Conceptual Framework and Model. Paper presented at ACIS.
- Enwereuzo, I., Antunes, P., Johnstone, D., & Mary, T. (2018). *On the Adoption of Crowdsourcing for Theory Testing*. Paper presented at the ECIS Proceedings.
- Enwereuzo, I., Antunes, P., & Johnstone, D. (2019). Patterns of Testing Theory with Human Subjects: A Dsign Science Perspective. Paper presented at the Americas Conference on Information systems (AMCIS).
- Estelles-Arolas, E., & Gonzalez-Ladron-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189-200. doi:10.1177/0165551512437638
- Eveleigh, A. (2014). Crowding out the archivist? Locating crowdsourcing within the broader landscape of participatory archives. *Crowdsourcing our Cultural Heritage*, 211-212.

- Filatova, E. (2012). *Irony and Sarcasm: Corpus Generation and Analysis Using CrowdsourceHelping*. Paper presented at the LREC.
- Følstad, A. (2008). Living labs for innovation and development of information and communication technology: A literature review. *eJOV: The Electronic Journal for Virtual Organization & Networks*, 10.
- Fong, B. (2017). An exploration of changing dissertation reqs and library services to support them. *Libraries and the academy*, 17(1), 129-144.
- Ford, R. C., Richard, B., & Ciuchta, M. P. (2015). Crowdsourcing: A new way of employing non-employees? *Business Horizons*, 58(4), 377-388.
- Fusch, P., & Ness, L. (2015). Are we there yet? Data saturation in qualitative research. *The qualitative report*, 20(9), 1408.
- Gadiraju, U., Fetahu, B., & Kawase, R. (2015). Training workers for improving performance in crowdsourcing microtasks *Design for Teaching and Learning in a Networked World* (pp. 100-114). Cham: Springer.
- Galea, S., & Tracy, M. (2007). Participation rates in epidemiologic studies. *Annals of epidemiology*, 17(9), 643-653.
- Gao, H., Barbier, G., & Goolsby, R. (2011). Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems*, 26(3), 10-14.
- Geerts, S. (2009). Discovering crowdsourcing: theory, classification and directions for use. *unpublished Master of Science in Innovation Management thesis, Eindhoven University of Technology*.
- Geiger, D., Seedorf, S., Schulze, T., Nickerson, R. C., & Schader, M. (2011). *Managing the Crowd: Towards a Taxonomy of Crowdsourcing Processes*. Paper presented at the AMCIS.
- Gelo, O., Braakmann, D., & Benetka, G. (2008). Quantitative and qualitative research: Beyond the debate. *Integrative Psychological and Behavioral Science*, 42(3), 266-290.
- Gerring, J. (2006). *Case study research: Principles and practices*: Cambridge University Press.
- Ghezzi, A., Gabelloni, D., Martini, A., & Nataleccchio, A. (2018). Crowdsourcing: a review and suggestions for future research. *International Journal of Management Reviews*, 20(2), 343-363.
- Ghose, A., Ipeirotis, P. G., & Li, B. (2010). *Designing Ranking Systems for Hotels on Travel Search Engines to Enhance User Experience*. Paper presented at the ICIS.
- Goel, V., & Pirolli, P. (1992). The structure of design problem spaces. *Cognitive science*, 16(3), 395-429.
- Goldkuhl, G. (2002). *Anchoring scientific abstractions—ontological and linguistic determination following socio-instrumental pragmatism*. Paper presented at the Proceedings of European Conference on Research Methods in Business, Reading.
- Goode, W., & Hatt, P. (1952). *Methods in social research*. New York: McGraw-Hill.
- Goodman, J. K., & Malkoc, S. A. (2012). Choosing here and now versus there and later: The moderating role of psychological distance on assortment size preferences. *Journal of Consumer Research*, 39(4), 751-768.
- Green, J., & Thorogood, N. (2018). *Qualitative methods for health research*: Sage.
- Gregor, & Hevner. (2013). Positioning and presenting design science research for maximum impact. *MIS quarterly*, 337-355.

- Gregor, S. (2002). A theory of theories in information systems *Information Systems Foundations: building the theoretical base* (pp. 1-20). Canberra: Australian National University.
- Gregor, S. (2006). The nature of theory in information systems. *MIS quarterly*, 30(3), 611-642.
- Gregor, S., & Jones, D. (2007). The Anatomy of a Design Theory. *Journal of the Association of Information Systems*, 8(5), 312-335.
- Gura, T. (2013). Citizen science: Amateur experts. *Nature*, 496(7444), 259-261.
- Heer, J., & Bostock, M. (2010). *Crowdsourcing graphical perception: using mechanical turk to assess visualization design*. Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.
- Hevner, & Chatterjee. (2010). Design science research in information systems *Design research in information systems* (pp. 9-22): Springer.
- Hevner, & Chatterjee, S. (2010). *Design research in information systems: theory and practice* (Vol. 22): Springer Science & Business Media.
- Hevner, Von Alan, March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1), 75-105.
- Holden, R. J., Scott, A. M. M., Hoonakker, P. L., Hundt, A. S., & Carayon, P. J. Q. o. L. R. (2015). Data collection challenges in community settings: Insights from two field studies of patients with chronic disease. 24(5), 1043-1055.
- Holley, R. (2010). Crowdsourcing: How and why should libraries do it? *D-Lib Magazine*, 16(3/4 Ma).
- Holmström, J., Ketokivi, M., & Hameri, A. (2009). Bridging practice and theory: A design science approach. *Decision Sciences*, 40(1), 65-87.
- Holsapple, C. W. (2008). DSS architecture and types *Handbook on Decision Support Systems 1* (pp. 163-189): Springer.
- Hosack, B., Hall, D., Paradice, D., & Courtney, J. (2012). A look toward the future: decision support systems research is alive and well. *Journal of the Association for Information Systems*, 13(5), 315.
- Howe, J. (2006). The rise of crowdsourcing. *Wired magazine*, 14(6), 1-4.
- Hyde, K. F. (2000). Recognising deductive processes in qualitative research. *Qualitative market research: An international journal*, 3(2), 82-90.
- Iivari, J. (2007). A paradigmatic analysis of information systems as a design science. *Scandinavian journal of information systems*, 19(2), 5.
- Israel, B. A., Krieger, J., Vlahov, D., Ciske, S., Foley, M., Fortin, P., . . . Palermo, A.-g. J. J. o. U. H. (2006). Challenges and facilitating factors in sustaining community-based participatory research partnerships: lessons learned from the Detroit, New York City and Seattle Urban Research Centers. 83(6), 1022-1040.
- Jarmolowicz, D., Bickel, W., Carter, A., Franck, C., & Mueller, E. (2012). Using crowdsourcing to examine relations between delay and probability discounting. *Behavioural processes*, 91(3), 308-312.
- Järvinen, P. (2007). Action research is similar to design science. *Quality & Quantity*, 41(1), 37-54.
- Jiang, L., Wagner, C., & Nardi, B. (2015). *Not Just in it for the Money: A Qualitative Investigation of Workers' Perceived Benefits of Micro-task Crowdsourcing*. Paper presented at the 48th Hawaii International Conference on System Sciences.
- Jonker, J., & Pennink, B. (2010). *The essence of research methodology: A concise guide for master and PhD students in management science*: Springer Science & Business Media.

- Jump, P. (2015). PhD: Is the doctoral thesis obsolete? *Times Higher Education*.
- Kafura, D. (1998). *Object-oriented software design and construction with C++*: Prentice-Hall, Inc.
- Kampffmeyer, H., & Zschaler, S. (2007). Finding the pattern you need: The design pattern intent ontology *International Conference on Model Driven Engineering Languages and Systems* (pp. 211-225). Berlin, Heidelberg: Springer.
- Khan, J. A. (2011). *Research methodology*. APH Publishing.
- Kietzmann, J. (2017). Crowdsourcing: A revised definition and introduction to new research. *Business Horizons*, 60(2), 151-153.
- Kim, J. (2012). The effect of design characteristics of mobile applications on user retention: an environmental psychology perspective. Paper presented at AMCIS.
- Kitchenham, B., Brereton, O., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology*, 51(1), 7-15.
- Kittur, A., Chi, E., & Suh, B. (2008). Crowdsourcing user studies with Mechanical Turk *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 453-456): ACM.
- Kittur, A., Smus, B., Khamkar, S., & Kraut, R. (2011). Crowdforge: Crowdsourcing complex work *Proceedings of the 24th annual ACM symposium on User interface software and technology* (pp. 43-52): ACM.
- Kleemann, F., Voß, G. G., & Rieder, K. (2008). Un (der) paid innovators: The commercial utilization of consumer work through crowdsourcing. *Science, technology & innovation studies*, 4(1), PP. 5-26.
- Krefting, L. (1991). Rigor in qualitative research: The assessment of trustworthiness. *American journal of occupational therapy*, 45(3), 214-222.
- Kroes, P. (2002). Design methodology and the nature of technical artefacts. *Design studies*, 23(3), 287-302.
- Kuechler, W., & Vaishnavi, V. (2012). A framework for theory development in design science research: multiple perspectives. *Journal of the Association for Information systems*, 13(6), 395.
- Kumar, R. (2005). *Research methodology*. Australia: Pearson.
- Lampe, C., Zube, P., Lee, J., Park, C. H., & Johnston, E. (2014). Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. *Government Information Quarterly*, 31(2), 317-326.
- Lee, A. S., & Hubona, G. S. (2009). A scientific basis for rigor in information systems research. *MIS quarterly*, 237-262.
- Lee, J., & Bui, T. (2000). *A template-based methodology for disaster management information systems*. In *Proceedings of the 33rd annual Hawaii international conference on system sciences* (pp. 7-pp). IEEE.
- Li, M., Tan, C., Wei, K., & Wang, K. (2017). Sequentiality of product review information provision: an information foraging perspective. *MIS Quarterly*, 41(3), 867-892.
- Lin, J., Amini, S., Hong, J. I., Sadeh, N., Lindqvist, J., & Zhang, J. (2012). *Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing*. Paper presented at the Proceedings of the 2012 ACM Conference on Ubiquitous Computing.
- Lokke, A., & Sorensen, P. (2014). Theory testing using case studies. *Electronic Journal of Business Research Methods*, 12(1), 66-74.
- Lowry, P. B., D'Arcy, J., Hammer, B., & Moody, G. D. (2016). "Cargo Cult" science in traditional organization and information systems survey research: A case for

- using nontraditional methods of data collection, including Mechanical Turk and online panels. *The Journal of Strategic Information Systems*, 25(3), 232-240.
- Lu, Y., Gupta, A., Ketter, W., & Van Heck, E. (2016). Exploring Bidder Heterogeneity in Multichannel Sequential B2B Auctions. *Management Information Systems Quarterly*, 40(3), 645-662.
- Lynham, S. (2002). The general method of theory-building research in applied disciplines. *Advances in developing human resources*, 4(3), 221-241.
- Maaninen-Olsson, E., Wismén, M., & Carlsson, S. (2008). Permanent and temporary work practices: knowledge integration and the meaning of boundary activities. *Knowledge Management Research & Practice*, 6(4), 260-273.
- MacKenzie, S., Podsakoff, P., & Podsakoff, N. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS quarterly*, 35(2), 293-334.
- Malan, R., & Bredemeyer, D. (2001). Defining non-functional requirements.
- Malone, T., & Crowston, K. (1994). The Interdisciplinary Study of Coordination. *ACM Computing Surveys*, 26(1), 87-119.
- March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251-266.
- Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods*, 44(1), 1-23.
- McInnis, B., Cosley, D., Nam, C., & Leshed, G. (2016). Taking a HIT: Designing around rejection, mistrust, risk, and workers' experiences in Amazon Mechanical Turk. *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2271-2282): ACM.
- Miles, M., & Huberman, A. (1994). *Qualitative Data Analysis*. Thousand Oaks, California: SAGE.
- Miles, M., Huberman, A., & Saldana, J. (2013). Qualitative data analysis: A methods sourcebook: Thousand Oaks, CA: Sage Publications Inc.
- Mingers, J. (2003). The paucity of multimethod research: a review of the information systems literature. *Information Systems Journal*, 13(3), 233-249.
- Morris, R. J. J. o. M. (1994). Computerized content analysis in management research: A demonstration of advantages & limitations. 20(4), 903-931.
- Mueller, B., & Urbach, N. (2013). *The Why, What and How of Theories in Information Systems Research*. Paper presented at the 34th International Conference on Information Systems (ICIS).
- Mumford, E. (1985). Defining system requirements to meet business needs: a case study example. *The Computer Journal*, 28(2), 97-104.
- Munro, M., Sogoni, M., Mashiyi, M., Shigadhlala, O., & Tsotsi, N. (2009). Fieldwork experiences and challenges in data collection: The Eastern Cape Province of South Africa. *A paper presented at the 57th Session of the International Statistical Institute*.
- Muntermann, J., Nickerson, R., & Varshney, U. (2015). Towards the Development of a Taxonomic Theory. Paper presented at the 21st Americas Conference on Information Systems, Puerto Rico.
- Nakatsu, R., Grossman, E., & Iacovou, C. (2014). A taxonomy of crowdsourcing based on task complexity. *Journal of Information Science*, 40(6), 823-834.
- Nel, E., Rich, E., Morojele, N., Harker Burnhams, N., Petersen Williams, P., Parry, C. J. D. E., Prevention, & Policy. (2017). Data collection challenges experienced while conducting the international alcohol control study (IAC) in Tshwane, South Africa. 24(5), 376-383.

- Nemati, H. R., Steiger, D. M., Iyer, L. S., & Herschel, R. (2002). Knowledge warehouse: an architectural integration of knowledge management, decision support, artificial intelligence and data warehousing. *33*(2), 143-161.
- Niederman, F., & March, S. (2015). Reflections on Replications. *AIS Transactions on Replication Research*, 1(1), 7.
- Niederman, F., & March, S. (2019). The “Theoretical Lens” Concept: We All Know What it Means, but do We All Know the Same Thing? *Communications of the Association for Information Systems*, 44.
- Nunamaker Jr, J. F., Chen, M., & Purdin, T. D. (1990). Systems development in information systems research. *Journal of management information systems*, 7(3), 89-106.
- Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts Work. Pap. Inf. Syst*, 10, 26.
- Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on amazon mechanical turk. *Judgement and Decision Making*, Vol. 5, No. 5.
- Paré, G., Tate, M., Johnstone, D., & Kitsiou, S. (2016). Contextualizing the twin concepts of systematicity and transparency in information systems literature reviews. *European Journal of Information Systems*, 25(6), 1-16.
- Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management*, 52(2), 183-199.
- Park, S., Shoemark, P., & Morency, L.-P. (2014). *Toward crowdsourcing micro-level behavior annotations: the challenges of interface, training, and generalization*. Paper presented at the Proceedings of the 19th international conference on Intelligent User Interfaces.
- Parvanta, C., Roth, Y., & Keller, H. (2013). Crowdsourcing 101 a few basics to make you the leader of the pack. *Health promotion practice*, 14(2), 163-167.
- Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70, 153-163.
- Peffers, K., Rothenberger, M., Tuunanen, T., & Vaezi, R. (2012). *Design science research evaluation*. Paper presented at the International Conference on Design Science Research in Information Systems.
- Peffers, K., Tuunanen, T., Rothenberger, M., & Chatterjee, S. (2007a). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45-77.
- Penker, M. (2000). Business Modeling with UML: Business Patterns at Work: New York, NY: John Wiley & Sons.
- Pierce, C. S., Scherra, E., & Care, H. (2012). The challenges of data collection in rural dwelling samples. 4(2), 25-30.
- Piper, S. (2006). Qualitative theory testing as mixed-method research. *Journal of Research in Nursing*, 11(3), 183-193.
- Popper, K. (2014). *Conjectures and refutations: The growth of scientific knowledge*: Routledge.
- Power, D. J. (2008). Decision support systems: a historical overview. *Handbook on Decision Support Systems* 1, 121-140.
- Prat, N., Comyn-Wattiau, I., & Akoka, J. (2015). A taxonomy of evaluation methods for information systems artifacts. *Journal of Management Information Systems*, 32(3), 229-267.

- Pries-Heje, J., & Baskerville, R. (2008). The design theory nexus. *MIS quarterly*, 731-755.
- Pries-Heje, J., Baskerville, R., & Venable, J. R. (2008). *Strategies for Design Science Research Evaluation*. Paper presented at the ECIS.
- Ramsey, S. R., Thompson, K. L., McKenzie, M., & Rosenbaum, A. (2016). Psychological research in the internet age: The quality of web-based data. *Computers in Human Behavior*, 58, 354-360.
- Reason, J. (2008). *The Human Contribution: Unsafe Acts, Accidents and Heroic Recoveries*. Surrey, England: Ashgate.
- Rittel, H., & Webber, M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, 4(2), 155-169.
- Roberts, N., Campbell, D., & Vijayasarathy, L. (2016). Using information systems to sense opportunities for innovation: integrating postadoptive use behaviors with the dynamic managerial capability perspective. *Journal of Management Information Systems*, 33(1), 45-69.
- Ross, J., Zaldivar, A., Irani, L., & Tomlinson, B. (2009). Who are the turkers? worker demographics in amazon mechanical turk. *Department of Informatics, University of California, Irvine, USA, Tech. Rep.*
- Sanyal, P. (2016). Characteristics and Economic Consequences of Jump Bids in Combinatorial Auctions. *Information Systems Research*, 27(2), 347-364.
- Sapateiro, C., Antunes , P., Johnstone, D., & Pino, J. (2016). Gathering big data for teamwork evaluation with microworlds. *Cluster Computing*, 20(2), 1637-1659.
- Sarkar, M. (2014). Challenges in Conducting Educational Research: The Case of a Developing Country. *Contemporary approaches to Research in Mathematics, Science, Health and Environmental Education*.
- Schenk, E., & Guittard, C. (2011). Towards a characterization of crowdsourcing practices. *Journal of Innovation Economics & Management*(1), 93-107.
- Schön, D. (1992). Designing as reflective conversation with the materials of a design situation. *Knowledge-based systems*, 5(1), 3-14.
- Schön, D. (1983). *The reflective practitioner: How professionals think in action*. New York: Basic Books.
- Sein, M., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action design research. *MIS quarterly*, 35(1), 37-56.
- Selander, L., & Jarvenpaa, S. (2016). Digital action repertoires and transforming a social movement organization. *MIS Quarterly*, 40(2), 331-352.
- Shank, D. B. (2016). Using crowdsourcing websites for sociological research: The case of Amazon Mechanical Turk. *The American Sociologist*, 47(1), 47-55.
- Simon, H. (1996). *The Sciences of the Artificial*. Cambridge, USA: The MIT Press.
- Simon, H. A. (1960). The new science of management decision. *The Ford distinguished lectures: Vol. 3*. New York, NY, US: Harper & Brothers. <http://dx.doi.org/10.1037/13978-000>.
- Smith, D., Manesh, M. M. G., & Alshaikh, A. (2013). How can entrepreneurs motivate crowdsourcing participants? *Technology Innovation Management Review*, 3(2).
- Snoek, C. G., Freiburg, B., Oomen, J., & Ordelman, R. (2010). *Crowdsourcing rock n'roll multimedia retrieval*. Paper presented at the Proceedings of the 18th ACM international conference on Multimedia.
- Sonnenberg, C., & vom Brocke, J. (2011). *Evaluation patterns for design science research artefacts*. Paper presented at the European Design Science Symposium.

- Spencer, D. (2009). *Card sorting: Designing usable categories*: Rosenfeld Media.
- Spencer, D., & Warfel, T. (2004). Card sorting: a definitive guide. *Boxes and arrows*. Retrieved from <https://boxesandarrows.com/Card-sorting-a-definitive-guide/>
- Steelman, Z., Hammer, B., & Limayem, M. (2014). Data Collection in the Digital Age: Innovative Alterantives to Student Samples. *Mis Quarterly*, 38(2), 355-378.
- Steelman, Z., Soror, A., Limayem, M., & Worrell, D. (2012). Obsessive compulsive tendencies as predictors of dangerous mobile phone usage. Paper presented at AMCIS.
- Stewart, N., Chandler, J., & Paolacci, G. (2017). Crowdsourcing Samples in Cognitive Science. *Trends in cognitive sciences*, 21(10), 736-748.
- Stinchcombe, A. L. (1987). *Constructing social theories*: University of Chicago Press.
- Su, A. I., Good, B. M., & van Wijnen, A. J. (2013). Gene Wiki Reviews: marrying crowdsourcing with traditional peer review. *Gene*, 531(2), 125.
- Thomas, D. (2006). A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation*, 27(2), 237-246.
- Thuan, N., Antunes, P., & Johnstone, D. (2016). Factors influencing the decision to crowdsource: A systematic literature review. *Information Systems Frontiers*, 18(1), 47-68.
- Thuan, N., Antunes, P., Johnstone, D., & Duy, N. (2015). *Establishing a Decision Tool for Business Process Crowdsourcing*. Paper presented at the International Conference on Future Data and Security Engineering.
- Thuan, N. H., Antunes, P., & Johnstone, D. (2016b). *A Design Science Method for Emerging Decision Support Environments*. Paper presented at the Australasian Conference on Information Systems, Adelaide, Australia.
- Thuan, N. H., Drechsler, A., & Antunes, P. (2019). Construction of Design Science Research Questions. *Communications of the Association for Information Systems*, 44, pp-pp.
- Turner, A. M., Kirchhoff, K., & Capurro, D. (2012). Using crowdsourcing technology for testing multilingual public health promotion materials. *Journal of medical Internet research*, 14(3).
- Van Aken. (2005). Management research as a design science: Articulating the research products of mode 2 knowledge production in management. *British journal of management*, 16(1), 19-36.
- Van Aken, & Romme, A. G. (2012). A design science approach to evidence-based management. 43-57.
- Van Der Valk, T., Van Driel, J. H., & De Vos, W. (2007). Common characteristics of models in present-day scientific practice. 37(4), 469-488.
- Van Gelder, M., Bretveld, R., & Roeleveld, N. (2010). Web-based questionnaires: the future in epidemiology? *American journal of epidemiology*, 172(11), 1292-1298.
- Van Looy, A., Poels, G., & Snoeck, M. (2017). Evaluating business process maturity models. *Journal of the Association for Information Systems*, 18(6), 461.
- Venable John, Pries-Heje, J., & Baskerville, R. (2012). *A comprehensive framework for evaluation in design science research*. Paper presented at the International Conference on Design Science Research in Information Systems.
- Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, 48(4), 865-885.
- Vitak, J., Proferes, N., Shilton, K., & Ashktorab, Z. (2017). Ethics regulation in social computing research: Examining the role of Institutional Review Boards. *Journal of Empirical Research on Human Research Ethics*, 12(5), 372-382.

- Von Alan, R., Hevner, A., March, S., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS quarterly*, 28(1), 75-105.
- Vondrick, C., Patterson, D., & Ramanan, D. (2013). Efficiently scaling up crowdsourced video annotation. *International Journal of Computer Vision*, 101(1), 184-204.
- Vukovic, M., Laredo, J., & Rajagopal, S. (2010). *Challenges and experiences in deploying enterprise crowdsourcing service*. Paper presented at the International Conference on Web Engineering.
- Vukovic, M., Lopez, M., & Laredo, J. (2010). *Peoplecloud for the globally integrated enterprise*. Paper presented at the Service-Oriented Computing. ICSOC/ServiceWave 2009 Workshops.
- Wacker, J. (1998). A definition of theory: research guidelines for different theory-building research methods in operations management. *Journal of Operations Management*, 16(4), 361-385.
- Wand, Y., & Weber, R. (1995). On the deep structure of information systems. *Information Systems Journal*, 5(3), 203-223.
- Ward, M. K., & Broniarczyk, S. M. (2011). It's not me, it's you: How gift giving creates giver identity threat as a function of social closeness. *Journal of Consumer Research*, 38(1), 164-181.
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, xiii-xxiii.
- Wei, W.-C. (2009). A technology acceptance model: Mediate and moderate effect. *Asia Pacific Management Review*, 14(4), 461-476.
- West, D., & Stowell, F. J. S. M. f. B. P. I., Artech, London. (1999). Models and diagrams and their importance to information systems analysis and design. 295-311.
- Wexler, M. N. (2011). Reconfiguring the sociology of the crowd: exploring crowdsourcing. *International Journal of Sociology and Social Policy*, 31(1/2), 6-20.
- Whitla, P. (2009). Crowdsourcing and its application in marketing activities. *Contemporary Management Research*, 5(1).
- Witschey, J., Murphy-Hill, E., & Xiao, S. (2013). Conducting interview studies: Challenges, lessons learned, and open questions *1st International Workshop on Conducting Empirical Studies in Industry* (pp. 51-54): IEEE.
- Wu, W., Tsai, W.T., & Li, W. (2013). Creative software crowdsourcing: from components and algorithm development to project concept formations. *International Journal of Creative Computing*, 1(1), 57-91.
- Yang, J., Adamic, L. A., & Ackerman, M. S. (2008). *Crowdsourcing and knowledge sharing: strategic user behavior on taskcn*. Paper presented at the Proceedings of the 9th ACM conference on Electronic commerce.
- Yuen, M.C., King, I., & Leung, K.S. (2011). *A survey of crowdsourcing systems*. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing* (pp. 766-773). IEEE.
- Zhang, K., Bhattacharyya, S., & Ram, S. (2016). Large-Scale Network Analysis for Online Social Brand Advertising. *MIS Quarterly*, 40(4), 849-868.
- Zhao, Y., & Zhu, Q. (2012). Exploring the motivation of participants in crowdsourcing contest. Paper presented at the 33rd International Conference on Information Systems, Orlando.

Zheng, H., Li, D., & Hou, W. (2011). Task design, motivation, and participation in crowdsourcing contests. *International Journal of Electronic Commerce*, 15(4), 57-88.

Zigurs, I., & Khazanchi, D. (2008). From profiles to patterns: A new view of task-technology fit. *Information systems management*, 25(1), 8-13.

APPENDICES

Appendix A. Theory testing Coding forms

This appendix presents some of the coding forms used for the study. They include reviewed literature, extracted data and analysis done on those literature.

Form A

Intents	Structure								
	Data Gathering				Data Analysis				
Case Study	Site Selection	Structured Interview	Perceptions reported		Cluster analysis				
					Socio material analysis				
					Pro-fit analysis				
					Transcription	Theme identifications	Sub-theme identification	coding	
	Online forum		Perceptions reported			Coding	Category identification		
						Content Analysis			
						Cluster analysis			

						Socio material analysis			
						Pro-fit analysis			
		Collection of posts from participants				Transcription	Theme identifications	Sub-theme identification	coding
							Coding	Category identification	
							Content Analysis		
							Cluster analysis		
	Training	Perceptions reported				Transcription	Socio material analysis		
							Pro-fit analysis		
							Theme identifications	Sub-theme identification	coding
							Coding	Category identification	
		Survey link sent to participants					Content Analysis		
							Analysis of variance (ANOVA)		
							Correlation analysis		

						Structural equation model analysis				
						Descriptive statistics				
Records	Development of Algorithm for experiment	Training	Survey link sent to participants			Analysis of variance (ANOVA)				
						Correlation analysis				
	Study for iPhone users	Observational data				Structural equation model analysis				
						Descriptive statistics				
						Content analysis				
						Descriptive statistics				
						Cluster analysis				
Data from database, archival data, downloads, posts, logs	Site Selection	Structured Interview	Perceptions reported			Socio material analysis				
						Pro-fit analysis				
						Transcription	Theme identifications	Sub-theme identification	coding	

							Coding	Category identification	
							Content Analysis		
						Cluster analysis			
						Socio material analysis			
						Pro-fit analysis			
							Theme identifications	Sub-theme identification	coding
							Coding	Category identification	
							Content Analysis		
						Cluster analysis			
						Socio material analysis			
						Pro-fit analysis			
							Theme identifications	Sub-theme identification	coding
							Coding	Category identification	
							Content Analysis		

						Analysis of variance (ANOVA)			
						Correlation analysis			
						Structural equation model analysis			
						Descriptive statistics			
						Econometric analysis			
						Descriptive statistics			
						Correlation analysis			
						Survival analysis			
						Panel vector autoregression analysis (PVAR)			
Experiment	Development of Algorithm for experiment	Training	Perceptions reported			Cluster analysis			
						Socio material analysis			
						Pro-fit analysis			
						Transcription	Theme identifications	Sub-theme identification	coding

						Coding	Category identification
						Content Analysis	
						Analysis of variance (ANOVA)	
						Correlation analysis	
						Structural equation model analysis	
						Descriptive statistics	
						Content analysis	
						Descriptive statistics	
						Cluster analysis	
						Socio material analysis	
						Pro-fit analysis	
						Theme identifications	Sub-theme identification
						Coding	Category identification
Data from database, archival data, downloads, posts, logs	Site Selection	Structured Interview	Perceptions reported		Transcription		coding

							Content Analysis		
Online forum	Collection of posts from participants	Perceptions reported	Transcription	Cluster analysis					
				Socio material analysis					
				Pro-fit analysis					
				Theme identifications	Sub-theme identification	coding			
				Coding	Category identification				
				Content Analysis					
Training	Perceptions reported	Transcription	Cluster analysis						
			Socio material analysis						
			Pro-fit analysis						
			Theme identifications	Sub-theme identification	coding				
			Coding	Category identification					
			Content Analysis						
	Survey link sent to participants		Analysis of variance (ANOVA)						

					Correlation analysis			
					Structural equation model analysis			
					Descriptive statistics			
					Econometric analysis			
					Descriptive statistics			
					Correlation analysis			
					Survival analysis			
					Panel vector autoregression analysis (PVAR)			
Experimental design	Simulation	Assigning roles to participants with experimental condition	Questionnaire distributed for information collection		Analysis of variance (ANOVA)			
					Correlation analysis			
					Structural equation model analysis			
					Confirmatory factor analysis			

						Path least square analysis			
Training on assigned activity	Questionnaire distributed for information collection	Structural equation model analysis	Confirmatory factor analysis	Path least square analysis	Descriptive statistics	Correlation analysis	Multivariate analysis (MANCOVA)	Confirmatory factor analysis	ANCOVA
Perceptions reported		Regression analysis	Path least square analysis	Regression analysis	LISTREL	Path least square analysis	Regression analysis	Regression analysis	Regression analysis

						Post-Hoc analysis			
						Analysis of variance (ANOVA)			
						Correlation analysis			
						Structural equation model analysis			
						Confirmatory factor analysis			
						Path least square analysis			
						Analysis of variance (ANOVA)			
						Correlation analysis			
						Structural equation model analysis			
						Confirmatory factor analysis			
						Path least square analysis			
						Descriptive statistics			
						Correlation analysis			

						Multivariate analysis (MANCOVA)			
						Confirmatory factor analysis			
						ANCOVA			
						LISTREL			
						Path least square analysis			
						Regression analysis			
						Post-Hoc analysis			
	Assigning roles to participants with experimental condition					Analysis of variance (ANOVA)			
						Correlation analysis			
						Structural equation model analysis			
						Confirmatory factor analysis			
						Path least square analysis			
						Questionnaire distributed for information collection			
		Training on assigned activity				Analysis of variance (ANOVA)			
						Correlation analysis			

						Structural equation model analysis			
						Confirmatory factor analysis			
						Path least square analysis			
						Descriptive statistics			
						Correlation analysis			
						Multivariate analysis (MANCOVA)			
						Confirmatory factor analysis			
						ANCOVA			
						LISTREL			
						Path least square analysis			
						Regression analysis			
						Post-Hoc analysis			
Lab experiment	Development of websites	Assigning roles to participants with			Questionnaire distributed for information collection		Analysis of variance (ANOVA)		
							Correlation analysis		

			experimental condition				Structural equation model analysis		
							Confirmatory factor analysis		
							Path least square analysis		
							Analysis of variance (ANOVA)		
							Correlation analysis		
							Structural equation model analysis		
							Confirmatory factor analysis		
							Path least square analysis		
							Descriptive statistics		
							Correlation analysis		

						Multivariate analysis (MANCOVA)		
						Confirmatory factor analysis		
						ANCOVA		
						LISTREL		
						Path least square analysis		
						Regression analysis		
						Post-Hoc analysis		
Reviews	participants evaluate reviews	Perceptions reported				Descriptive statistics		
						Correlation analysis		
						Multivariate analysis (MANCOVA)		
						Confirmatory factor analysis		
						ANCOVA		
						LISTREL		

							Path least square analysis		
							Regression analysis		
							Post-Hoc analysis		
Survey	Training on assigned activity	Questionnaire distributed for information collection					Analysis of variance (ANOVA)		
							Correlation analysis		
							Structural equation model analysis		
							Confirmatory factor analysis		
							Path least square analysis		
	Distribution of survey materials	Training on assigned activity	Questionnaire distributed for information collection				Analysis of variance (ANOVA)		
							Correlation analysis		
							Structural equation model analysis		

							Confirmatory factor analysis		
							Path least square analysis		
							Descriptive statistics		
							Correlation analysis		
							Multivariate analysis (MANCOVA)		
							Confirmatory factor analysis		
							ANCOVA		
							LISTREL		
							Path least square analysis		
							Regression analysis		
							Post-Hoc analysis		
Online survey	Survey link sent to participants	Participants contacted after a period of time	Perceptions reported				Path least square analysis		
							Rating done/ Reassessment		
							Rating aggregation		

						Descriptive statistics		
						Correlation analysis		
						Multivariate analysis (MANCOVA)		
						Confirmatory factor analysis		
						ANCOVA		
						LISTREL		
						Path least square analysis		
						Regression analysis		
						Post-Hoc analysis		
Delphi Study	Survey link sent to participants	Participants contacted after a period of time	Perceptions reported			Descriptive statistics		
						Correlation analysis		
						Multivariate analysis (MANCOVA)		
						Confirmatory factor analysis		

							ANCOVA		
							LISTREL		
							Path least square analysis		
							Regression analysis		
							Post-Hoc analysis		
			Rating done/ Reassessment				Path least square analysis		
							Rating aggregation		

Form B

Intents	Structure					
	Data Gathering					Data Analysis
Case Study	Site Selection	Structured Interview	Perceptions reported			
		Online forum	Collection of posts from participants			
		Training	Perceptions reported			
			Survey link sent to participants			
Records	Development of Algorithm for experiment	Training	Survey link sent to participants			
	Data from database, archival data, downloads, posts, logs	Study for iPhone users	Observational data			
		Site Selection	Structured Interview	Perceptions reported		
			Online forum	Collection of posts from participants		
		Training	Perceptions reported			

				Survey link sent to participants		
		Data extracted over a period of time				
Experiment	Development of Algorithm for experiment	Training	Perceptions reported			
			Survey link sent to participants			
		Study for iPhone users	Observational data			
	Data from database, archival data, downloads, posts, logs	Site Selection	Structured Interview	Perceptions reported		
			Online forum	Collection of posts from participants		
			Training	Perceptions reported		
				Survey link sent to participants		
		Data extracted over a period of time				
	Experimental design	Simulation	Assigning roles to participants with experimental condition	Questionnaire distributed for information collection		

				Training on assigned activity	Questionnaire distributed for information collection	
				Perceptions reported		
Virtual environment created for experiment	Assigning roles to participants with experimental condition	Assigning roles to participants with experimental condition	Questionnaire distributed for information collection			
				Training on assigned activity	Questionnaire distributed for information collection	
				Perceptions reported		
Assigning roles to participants with experimental condition	Assigning roles to participants with experimental condition	Assigning roles to participants with experimental condition	Questionnaire distributed for information collection			
				Training on assigned activity	Questionnaire distributed for information collection	
				Perceptions reported		

				Questionnaire distributed for information collection		
Survey	Lab experiment	Development of websites	Assigning roles to participants with experimental condition	Questionnaire distributed for information collection		
				Training on assigned activity	Questionnaire distributed for information collection	
			Perceptions reported			
	Reviews		participants evaluate reviews	Perceptions reported		
	Training on assigned activity	Questionnaire distributed for information collection				
	Distribution of survey materials	Training on assigned activity	Questionnaire distributed for information collection			
	Online survey	Survey link sent to participants	Participants contacted after a period of time	Perceptions reported		
				Rating done/ Reassessment		

Delphi Study	Survey link sent to participants	Participants contacted after a period of time	Perceptions reported			
			Rating done/ Reassessment			

Form C

Detail of Coding on Theory testing using Crowdsourcing								
General information			Data Collection				Data Analysis	
Journal	Year	Title	Method of Data collection	Recruitment Method	Data Quality Check Method	Data Reduction Method	Content Analysis Method	Statistical Analysis Method
	2016	A TREE-BASED APPROACH FOR ADDRESSING SELFSELECTION IN IMPACT STUDIES WITH BIG DATA	Experiment * Simulation * Observation * survey	Eligibility Criteria			Performance analysis propensity score sub-classification (PSS) propensity score matching (PSM)	
	2016	AN EXPLORATORY STUDY OF THE FORMATION AND IMPACT OF ELECTRONIC SERVICE FAILURES	Survey	Invitations e-mailed to members	Used the filtering question inserted at the start of the questionnaire	Removal of ambiguous responses, coding	Content analysis (sorting, chi-square test and Crisp-Set Qualitative Comparative Analysis csQCA)	Descriptive Statistics, Post-Hoc analysis
	2016	ARE SOCIAL MEDIA EMANCIPATORY OR HEGEMONIC? SOCIETAL EFFECTS OF MASS MEDIA DIGITIZATION IN THE CASE OF THE SOPA DISCOURSE	Case study * tweets	Records from social media sites	manipulation control on task characteristics through contextual framing of survey questions	Use of Keywords, description, categories, coding	Transcription, Interpretive analysis	

	2016	CAPTURING THE COMPLEXITY OF MALLEABLE IT USE: ADAPTIVE STRUCTURATION THEORY FOR INDIVIDUALS	Survey * questionnaire	Smart phone users		Removal of responses with Missing data		PLS, Post-Hoc analysis	
	2016	COMBATING INFANT MORTALITY IN RURAL INDIA: EVIDENCE FROM A FIELD STUDY OF EHEALTH KIOSK IMPLEMENTATIONS	Survey *Training * questionnaire, * semi-structured interview	Use of sampling frame				hierarchical linear modelling, descriptive statistics, correlation	
	2016	COMPREHENSIBLE PREDICTIVE MODELS FOR BUSINESS PROCESSES	Experiment * Simulation * Logs *		Data filter based on keyword(complete)		Content analysis		
	2016	CONTRACT DESIGN CHOICES AND THE BALANCE OF EX ANTE AND EX POST TRANSACTION COSTS IN SOFTWARE DEVELOPMENT OUTSOURCING	Record * archival data					descriptive statistics, regression analysis	
	2016	CROWD-SQUARED: AMPLIFYING THE PREDICTIVE POWER OF SEARCH TREND DATA	Experiment * development of website (online word assoc.) * Assigning tasks (Crowdsourcing)	CS platform	Repetition of a given task		Comparative analysis	descriptive statistics, correlation analysis	CS, platform selection, recruitment (open to all)

	2016	DIGITAL ACTION REPERTOIRES AND TRANSFORMING A SOCIAL MOVEMENT ORGANIZATION	Case study * interviews * observation	Use of sampling frame		Coding	Content analysis (theme identification)		
	2016	EXPLORING BIDDER HETEROGENEITY IN MULTICHANNEL SEQUENTIAL B2B AUCTION	Records * logs					Cluster analysis	
	2016	FREE VERSUS FOR-A-FEE: THE IMPACT OF A PAYWALL ON THE PATTERN AND EFFECTIVENESS OF WORD- OF-MOUTH VIA SOCIAL MEDIA	Case study * tweets * online survey				Case study * tweets	descriptive statistic	
	2016	LARGE-SCALE NETWORK ANALYSIS FOR ONLINE SOCIAL BRAND ADVERTISING	Experiment * development of algorithm *	launch advertisements on a social platform (e.g., Facebook)	(Data Cleansing) Removal of posts not in English, users who made very few comments were not included, designed a set of rules to remove fake users and their corresponding activities		Network Analysis, performance comparison	descriptive statistic	

	2016	MOBILE APP ANALYTICS: A MULTIPLE DISCRETE-CONTINUOUS CHOICE FRAMEWORK	Experiment * development of App * Assigning tasks	stratified sampling				ANOVA, Correlation	
	2016	PIRATES IN THE LAB: USING INCENTIVIZED CHOICE EXPERIMENTS TO EXPLORE PREFERENCE FOR (UN)AUTHORIZED CONTENT	Experiment * training * Assigning tasks based on treatment* questionnaire	Invitation through the ORSEE Internet recruitment system Another study - recruited from a nation-wide survey panel	Manipulation check			Descriptive	
	2016	SENIOR EXECUTIVES' IT MANAGEMENT RESPONSIBILITIES: SERIOUS IT-RELATED DEFICIENCIES AND CEO/CFO TURNOVER	Records * financial reports, databases					exploratory factor analysis, Descriptive statistics and univariate tests, regression analyses	

	2016	THE COMPENSATORY INTERACTION BETWEEN USER CAPABILITIES AND TECHNOLOGY CAPABILITIES IN INFLUENCING TASK PERFORMANCE: AN EMPIRICAL ASSESSMENT IN TELEMEDICINE CONSULTATIONS	Survey * Online survey, Experiment, * laboratory experiment * Experimental design * assigning task based on experimental condition	Purposeful sampling technique, we contacted individuals based on two criteria, were contacted directly and invited to participate in the online survey - students that have completed a course	answered questions to test the manipulations	Removal of outliers		confirmatory factor analysis, regression analyses, Post Hoc Analysis, ANOVA	
	2016	THE DUALITY OF EMPOWERMENT AND MARGINALIZATION IN MICROTASK CROWDSOURCING: GIVING VOICE TO THE LESS POWERFUL THROUGH VALUE SENSITIVE DESIGN	case study * semi-structured interview/ questions Another study - Case study * online forum, reporting of perception	CS platform		Coding	open coding, analytical categories, data display matrices, themes identification	Descriptive Statistics and Correlation, Sensitivity Analysis	CS (m-Turk as case study) - used for data collection, but not for analysis
	2016	TOWARD A BETTER MEASURE OF BUSINESS PROXIMITY: TOPIC MODELING FOR INDUSTRY INTELLIGENCE	Records * logs			Random selection of sample (desired size)		Correlation Analysis	

	2016	TOWARD A DIGITAL ATTRIBUTION MODEL: MEASURING THE IMPACT OF DISPLAY ADVERTISING ON ONLINE CONSUMER BEHAVIOR	Records, * observational data			Time frame		Descriptive Statistics	
	2016	TV'S DIRTY LITTLE SECRET: THE NEGATIVE EFFECT OF POPULAR TV ON ONLINE AUCTION SALES	Record * transaction data, click data, data set			Time frame, price range		descriptive statistics, correlation analysis	
MISQ 2016 ends	2016	WHEN DOES REPOSITORY KMS USE LIFT PERFORMANCE? THE ROLE OF ALTERNATIVE KNOWLEDGE SOURCES AND TASK ENVIRONMENTS	Record * archived data, Survey			Removal of uncompleted responses		descriptive statistics, correlation analysis	
	2017	USING FORUM AND SEARCH DATA FOR SALES PREDICTION OF HIGH-INVOLVEMENT PROJECTS	Record * Sales Data, logs, forum data, website			Time frame		PLS, Post-Hoc analysis, Descriptive Statistics and Correlations, regression, Econometric Analysis	

	2017	TRUST AND THE STRENGTH OF TIES IN ONLINE SOCIAL NETWORKS: AN EXPLORATORY FIELD EXPERIMENT	Experiment *Dev of website for experiment, training (instruction sent via email, further instruction via YouTube video), random assigning to roles, questionnaire	Recruitment done via a hybrid online–offline snowball sampling method, an initial e-mail was sent, sign-up required, then added to subject pool, incentive given	Quiz	incomplete responses		OLS regression, Descriptive statistics	
	2017	HOW IS YOUR USER FEELING? INFERRING EMOTION THROUGH HUMAN–COMPUTER INTERACTION DEVICES	Experiment * Dev of webpage, training, assigning task, observation, perception reported via survey	CS platform				Descriptive Statistics	CS only used for performing the task
	2017	GROWING ON STEROIDS: RAPIDLY SCALING THE USER BASE OF DIGITAL VENTURES THROUGH DIGITAL INNOVATION	Case study *Archival data, Interviews, observation				Mapping, open coding, axial coding, clustering, selective, coding		
	2017	DESIGN AND EVALUATION OF AUTO-ID ENABLED SHOPPING ASSISTANCE ARTIFACTS IN CUSTOMERS' MOBILE PHONES: TWO RETAIL STORE LABORATORY EXPERIMENTS	Experiment *Experimental design, Dev of webpage, assigning task, perception reported via survey	recruited from a mailing list, participants were located, e-mail sent to customers Another study - recruited through a subject pool at a public university, incentive				Descriptive Statistics	

	2017	AN EMPIRICAL ANALYSIS OF INTELLECTUAL PROPERTY RIGHTS SHARING IN SOFTWARE DEVELOPMENT OUTSOURCING	Record * logs, database			Time frame, Missing information, based on some conditions		descriptive statistics, correlation analysis, regression	
	2017	FROM MONOLOGUE TO DIALOGUE: PERFORMATIVE OBJECTS TO PROMOTE COLLECTIVE MINDFULNESS IN COMPUTER-MEDIATED TEAM DISCUSSIONS	Experiment *Experimental design, training, assigning task, perception reported via questionnaire			failed to follow instructions removed from the analysis		PLS, Descriptive Statistics	
	2017	CYBERCRIME DETERRENCE AND INTERNATIONAL LEGISLATION: EVIDENCE FROM DISTRIBUTED DENIAL OF SERVICE ATTACKS	Record * logs, database					Descriptive Statistics, Correlations	
	2017	WHEN DO IT SECURITY INVESTMENTS MATTER? ACCOUNTING FOR THE INFLUENCE OF INSTITUTIONAL FACTORS IN THE CONTEXT OF HEALTHCARE DATA BREACHES	Record * logs, database			Time frame		Descriptive Statistics, Correlations	
	2017	USER COMPENSATION AS A DATA BREACH RECOVERY ACTION: AN INVESTIGATION OF THE SONY PLAYSTATION NETWORK BREACH	Survey * Online survey, (CS)		follow-up data collection			Descriptive Statistics and Correlations, Response Surface Analysis	CS (used for data collection)

	2017	SEQUENTIALITY OF PRODUCT REVIEW INFORMATION PROVISION: AN INFORMATION FORAGING PERSPECTIVE	Experiment, * laboratory experiment * Experimental design * Training *assigning task based on experimental condition	recruited through posts on the university online forum		Coding	Coding	MANOVA	
	2017	OPERATIONAL IT FAILURES, IT VALUE DESTRUCTION, AND BOARD-LEVEL IT GOVERNANCE CHANGES	Record * logs, database					Descriptive Statistics, Econometric analysis	
	2017	ON THE ROLE OF FAIRNESS AND SOCIAL DISTANCE IN	Experiment, * laboratory experiment * Experimental design * Training *assigning task based on experimental condition	used a standard random procedure algorithm for selection				ANOVA	
	2017	KNOWLEDGE MANAGEMENT SYSTEM USE AND JOB	Survey * Online survey, semi-structured interviews					Descriptive Statistics and Correlations	
	2017	GROUNDED THEORY METHODOLOGY IN	Case study *Interviews,				Classification		
	2017	EXPLAINING POST-IMPLEMENTATION EMPLOYEE SYSTEMUSE AND JOB PERFORMANCE: IMPACTS OF THE CONTENT	Experiment, * * Training *assigning task based on experimental condition			Coding	Theme identification, coding	Descriptive Statistics and Correlations, PLS	

	2017	THROUGH THE EYES OF OTHERS: HOW ONLOOKERS	Case study *semi-structured Interviews, observation		cut-off window		Coding		
	2017	THE IMPACT OF HEALTH INFORMATION SHARING ON	Record * logs, database					descriptive statistics, econometric , Difference in Difference Analysis	
	2017	SOCIAL NETWORK INTEGRATION AND USER CONTENT GENERATION: EVIDENCE FROM NATURAL EXPERIMENTS	Record * reviews					Descriptive Statistics, Correlation, Econometric , difference-in-differences	
	2017	SOCIAL MEDIA AFFORDANCES FOR CONNECTIVE ACTION:	Record * microblog posts (tweets)			Time frame, use of hashtags	cluster analysis, motifs analysis		
	2017	INFORMATION TECHNOLOGY INVESTMENTS AND FIRM RISK	Record * database			Time frame		Econometric , Descriptive Statistics, OLS regressions	
	2017	EXTRACTING REPRESENTATIVE INFORMATION ON	Experiment* Training *assigning task based on experimental condition* questionnaire					PLS, Descriptive Statistics	
	2017	ANTECEDENTS OF INFORMATION SYSTEMS SOURCING	Record * database			missing data		Descriptive Statistics and	

								Correlations, PLS	
MISQ 2017 ends	2017	A NOMOLOGICAL NETWORK OF KNOWLEDGE MANAGEMENT SYSTEM USE: ANTECEDENTS	Case study *semi- structured Interviews, survey				transcription, coding, Interpretive analysis		
	2016	The role of Transaction Cost Economics in Information	Record * Literature			Coding		correlation, ANOVA	
	2016	The effects of collecting and connecting activities on knowledge	Survey * questionnaire * semi-structured interviews			missing data, coding		confirmato ry factor analyses, EFA, structural equation model	
	2016	Social media network behavior: A study of user passion and	Survey * online survey	online via e- mail,				PLS-SEM	
	2016	Open source project success: Resource access, flow,	Record * database, archival data			Coding		Descriptive statistics	

	2016	Is SAM still alive? A bibliometric and interpretive mapping	Record * database, archival data, bibliographic data			Coding, based on citation	Open coding, category identification, interpretive analyses		
	2016	Governing innovation in U.S. state government: An ecosystem	Record * database, archival data			Coding		exploratory factor analysis	
	2016	Expectable use: An important facet of IT usage	case study * semi-structured interview/ e-questionnaires			Removal of outliers, coding	Coding, category identification, interpretive analyses	PLS	
	2016	Do shareholders favor business analytics announcements?	Record * database, archival data			insufficient information	cluster	OLS regression, Post hoc analysis	
	2016	Does mutuality matter? Examining the bilateral nature	Survey * online questionnaire	Randomly selected from database, invitation via email		insufficient data quality		Descriptive statistics, CFA	
	2016	Do different kinds of trust matter? An examination of the three	Experiment * experimental design *simulation *assigning task based on experimental	recruited participants through flyers, advertisements				PLS, Post-hoc analysis	

			condition* questionnaire						
	2016	Creating agile organizations through IT: The influence of (CS)	Survey * online questionnaire	CS platform	using established scales, randomizing the appearance of questions, and using different scaling for some measures, providing warnings, breaking up the survey, providing attention traps, recording the time spent on the survey, screening based on their IP addresses, language, and geographic locations, preventing “ballotstuffing	coding		PLS	CS used to get data

JSIS 2016 ends	2016	A communicative-tension model of change-induced collective	case study * semi-structured interview* observation *document * emails			coding			
	2017	Towards analysing the rationale of information security noncompliance: Devising a Value-Based Compliance analysis	case study * semi-structured interview* observation				Transcription, coding, categorization, Compliance analysis		
	2017	The paradoxical effects of legal intervention over unethical	case study * semi-structured interview				transcription, code-based content analysis (open coding, axial coding, and selective coding), category aggregation		
	2017	Strategy, Resource Orchestration and E-commerce Enabled	case study * semi-structured interview *magazine, newspapers			Transcription, coding,	visual mapping strategy, continuous comparison strategy, theme generation		

	2017	Performance implications of information-value offering in e-service systems: Examining the resource-based perspective	Survey * online questionnaire			discarding incomplete response		confirmatory factor analysis, PLS	
	2017	Solvers' participation in crowdsourcing platforms: Examining (CS)	Survey * online questionnaire *archival data	platform selection, invitation by emails through the internal messaging tool	non-response bias check			PLS, Descriptive statistics, Post hoc analysis	CS used for data collection
	2017	Distinguishing the effects of B2B information quality, system	Experiment, * laboratory experiment * Experimental design * Training *assigning task based on experimental condition * questionnaire		open-ended questions, different scale headers for different questions		transcription, theme identification	PLS	
	2017	Conflict resolution in business services outsourcing	Survey * questionnaire * interview			coding	transcription, coding		
	2017	Closing the loop: Empirical evidence for a positive feedback	Record * database, archival data			Time frame		Descriptive statistics, correlation	

	2017	Exploring the tension between transparency and datafication	Report *reports, website content, media articles			Coding	Classification, coding		
JSIS 2017 ends	2017		Report *research articles			based on set criteria, coding	Categorization, grouping, coding, comparison analysis		Article on theory testing ...
	2016	Using Information Systems to Sense Opportunities	Survey * questionnaire * interview	used Qualtrics	included two screening questions	excluding incomplete responses		wave analysis, confirmatory factor analysis	
	2016	Untangling a Web of Lies: Exploring Automated	Record * emails, archival data			removal of automated e-mails	text analysis, coding		
	2016	Understanding Information Systems Integration	Record * cases, archival data			coding, lacking sufficient information	sorting, categorization, coding, comparison, clustering, integration		

	2016	Trust Development in Globally Distributed Collaboration: A Case of U.S. and Chinese Mixed	Experiment, * Experimental design * Training *assigning task based on experimental condition * questionnaire * interview * documentation (home assignment)				transcription, categorized themes, Review and discussion	EFA and CFA, ANOVA	
	2016	The Effects of IT-Enabled Cognitive Stimulation	Experiment, * Experimental design * Training *assigning task based on experimental condition * questionnaire *			outlier, incomplete data			
	2016	The Determinants and Impacts of Aesthetics in	Experiment, * Experimental design * Training *assigning task based on experimental condition * questionnaire *			criterion application		PLS	
	2016	Technology Evaluation and Imitation: Do They Have Differential or Dichotomous Effects on ERP	case study * semi-structured interview *questionnaire	randomly selected from the company database	extrapolation procedure			regression analysis	

	2016	Sharing Knowledge in Social Q&A Sites: The	Survey *Online survey	To recruit participants, a link to the survey was posted on the home page of the social Q&A website under investigation				MANOVA	
	2016	Problematic Use of Social Networking Sites:	Survey *Online survey					CFA, Post hoc Analyses	
	2016	Not as Smart as We Think: A Study of Collective	Experiment, * Experimental design * Training *assigning task based on experimental condition * questionnaire *					CFA	
	2016	More Than Meets the Eye: How Oculometric Behaviors Evolve Over the Course of Automated	Experiment, * Experimental design * Training *assigning task based on experimental condition * questionnaire *		completion of an automated deception detection interview used to identify smugglers				
	2016	Influentials, Imitables, or Susceptibles? Virality	Record * database						

	2016	Influence of Firm's Recovery Endeavors upon	Experiment, * Experimental design * Training *assigning task based on experimental condition * questionnaire *	sample frame, e-mail invitation, notified that participation was voluntary		did not meet the age criterion		LISREL	
	2016	Impact of Network Structure on Malware Propagation: A Growth Curve Perspective	Record * database			invalid user IDs		Network Analysis, Descriptive statistics, correlation	
	2016	Global Differences in Online Shopping Behavior:	Experiment, * Experimental design * Training *assigning task based on experimental condition * questionnaire *	recruited through their university's behavioral laboratories, all students had prior experience with online shopping	principled approach for diagnosing which entries qualified as an error	Data cleaning, Incomplete surveys, nonnumerical entries, and missing values		Regression	
	2016	From Warning to Wallpaper: Why the Brain Habituates to Security Warnings and What Can Be	Experiment, * Experimental design * Training *assigning task based on experimental condition * questionnaire *		manipulation check			ANOVA	

	2016	Firm Boundaries, Information Processing Capacity,	Survey *Online survey					PLS, Descriptive statistics, Post hoc analysis	
	2016	Examining Firms' Green Information Technology	Survey *Online survey	Via e-mail and phone conversation				CFA	
	2016	Empirical Assessment of Alternative Designs for	Experiment, * dev of website* Experimental design * Training *assigning task based on experimental condition * questionnaire *					MANOVA, PLS	
	2016	Emotion in IT Investment Decision Making with	case study * semi-structured interview* observation			Coding	Transcription, categorization, pattern identification, coding	panel vector autoregress ion (PVAR)	
	2016	Effects of Social Interaction Dynamics on	Record * database						
	2016	Early Predictions of Movie Success: The Who,	Record * database			Exclusion based on criteria, time frame		Regression Analysis	

	2016	Do Professional Reviews Affect Online User	Record * database			Based on categories			
	2016	Detecting Fraudulent Behavior on Crowdfunding	Record * database				Coding, extraction	Descriptive Statistics	
	2016	Could Deal Promotion Improve Merchants'	Record * database					Econometric, Regression	
	2016	Concurrent IT Sourcing: Mechanisms and	Survey *Online survey	sent prospective respondents a prenotification letter to solicit participation				Descriptive Statistics, econometric, OLS	
	2016	Computer-Mediated Deception: Strategies	Experiment, * dev of website* Experimental design * Training *assigning task based on experimental condition * questionnaire *			cleaned and categorized		regression analysis, decision tree, and support vector machine (SVM) analysis	
	2016	Augmented Virtual Doctor Office: Theory-based	Experiment, * simulation* Experimental design * Training *assigning task based on experimental condition * questionnaire *		Common-Method Bias check			EFA	

	2016	An Examination of Effective IT Governance in	Record * database, Archival data			coding		descriptive statistics and correlations	
	2016	An Empirical Validation of Malicious Insider	Record * database, Archival data			Keyword	Extraction, coding, categorization thematic analysis		
JMIS 2016 ends	2016	A Friend Like Me: Modelling Network Formation in	Record * database, Archival data					descriptive statistics and correlations	
	2017	Winning Back Technology Disadopters: Testing a	Survey *Online survey			Incomplete response		PLS, LISREL	
	2017	Who Do You Think You Are? Common and	Survey *Online survey			Based on certain criteria		PLS, Post Hoc Analysis	
	2017	What Leads to Prosocial Behaviors on Social	Survey *Online survey	e-mailed invitations to 1,500 randomly selected people		incomplete responses		EFA, CFA, Post Hoc Analysis	
	2017	To Cyberloaf or Not to Cyberloaf: The Impact of	Survey *Online survey					CFA, LISREL	

	2017	The Risk Implications of Mergers and Acquisitions	Record * database, Archival data					descriptive statistics, Time-Series Analyses	
	2017	The Consensus Effect in Online	Record * database, Archival data					Descriptive Statistics, ANOVA	
	2017	Gamification in the Workplace: The Central Role of	Survey *Online survey	by contacting several company executives and soliciting their participation in this study		Excluding responses with missing values		PLS	
	2017	Whose and What Social Media Complaints Have	Record * online forum or page *database, Archival data			Time frame		Econometric Analysis	
	2017	Training to Mitigate Phishing Attacks Using	Experiment* Experimental design * Training *assigning task based on experimental condition * questionnaire *			incomplete, no-training condition were excluded		Descriptive statistics, logistic regression analysis	
	2017	The Effects of Process Orientations on Collaboration Technology Use and Outcomes in	Experiment* Experimental design * Training *assigning task based on experimental			incomplete responses	missing data analysis	PLS	

			condition * questionnaire *						
	2017	The Effect of Belongingness on Obsessive- (CS)	Survey *Online survey	Platform selection,	Use of filter question	incomplete responses, attention trap question		PLS	CS used for data collection
	2017	Predicting and Deterring Default with Social Media	Record * database, Archival data			Time frame		Descriptive statistics, logistic regression analysis	
	2017	Integrated Health Information Technology and the	Experiment* Experimental design * Training *assigning task based on experimental condition * questionnaire *					Descriptive statistics, logistic regression analysis	
	2017	How Doctors Gain Social and Economic Returns in	Record * database, Archival data			reduced the data noise (criteria based)		Descriptive Statistics, PLS	

	2017	Gamification in the Workplace: The Central Role of	Survey *Online survey			Excluding responses with missing values		PLS	
	2017	Beyond the Personalization–Privacy Paradox: Privacy Valuation, Transparency Features, and	Experiment* Experimental design * Training *assigning task based on experimental condition * questionnaire (perception reported) *	recruited via e-mail, social networks, forums, and local online classified advertisements	To assure the high quality of our data set, we applied a data-cleaning process to detect satisficing participants	very low response times, who failed to answer a control question, Manipulation Check		CFA, ANCOVA	
	2017	Using IT Design to Prevent Cyberbullying (CS)	Survey *Factorial Survey Design		using established scales, randomizing the appearance of questions, and using different scaling for some measures			PLS, Post Hoc Analysis	

	2017	The Impact of Mental Representations on ICT Related	Survey *Online survey	recruited through multiple channels (e.g., e-mail, face-to-face, invitation letter, social media).				PLS-SEM, Post Hoc Analysis	
	2017	Reflective Technology Assimilation: Facilitating	case study * semi-structured interview* field notes			Coding	Interpretive analysis		
	2017	Platform Structures, Homing Preferences, and	Record * database, Archival data					descriptive statistics	
	2017	IT-Enabled Revenue Cycle Transformation in	case study * semi-structured interview* Observation				Transcription, Diagnostic Process mapping		
	2017	Impediments to Information Systems	Survey *Online survey		pre-screen question			CFA, EFA, PLS	
	2017	Enabling Effective Operational Risk Management	case study * semi-structured interview* Observation				Interpretive analysis		
	2017	An Agile Methodology for the Disaster Recovery	case study * semi-structured interview* Observation				Interpretive analysis		

	2017	A Two-Stage Model of Generating Product Advice:	Experiment* Online experiment *Experimental design *dev of website * Training *assigning task based on experimental condition * questionnaire (perception reported) *					ANOVA, ANCOVA	
	2017	User Motivations in Protecting Information (CS)	Survey *Online survey			Exclusion based on criteria, unreasonably fast completion times, failed attention filter questions		PLS	CS used for data collection
	2017	Risks and Controls in Internet-Enabled Reverse	Delphi *brainstorming * reviews * interview				Interpretive analysis		
	2017	Promoting the System Integration of Renewable	Record * database, Archival data						

	2017	Information Security Control Theory: Achieving a	case study * semi-structured interview* Observation			Coding	Transcription, coding		
	2017	Exploring Emerging Hacker Assets and Key	case study * online forum* Observation				classification		
	2017	Beyond Brainstorming: Exploring Convergence in	Experiment*Experimental design *dev of website * Training *assigning task based on experimental condition * questionnaire (perception reported) *					CFA, ANOVA	
	2017	Assessing the Credibility of Decisional Guidance	Experiment*Experimental design *dev of website * Training *assigning task based on experimental condition * questionnaire (perception reported) *					linear mixed modelling (LMM)	

	2017	A Video-Based Screening System for Automated	Experiment*Experimental design *dev of website * Training *assigning task based on experimental condition * questionnaire (perception reported) *					Multivariate Regression	
	2017	A Temporal Study of the Effects of Online	Record * Online forum*database, Archival data					Descriptive Statistics and Correlation	
JMIS 2017 ends	2017	A Data-Mining Approach to Identification of Risk	Record * database, Archival data						
	2016	User personality and resistance to mandatory information systems in empirical test of dispositional resistance to	Survey *Online survey	e-recruiting system				PLS, Post-hoc analysis	

	2016	Understanding individual user resistance	case study * semi-structured interview* Observation* archival data			coding	Transcription, coding, Interpretive analysis	Descriptive stats, logistic regression analysis	
	2016	To share or not to share: the effects of	Survey *Online survey						
	2016	The map and the territory: an ethnographic	case study * semi-structured interview* Observation* archival data				Transcription, indexing, coding, Interpretive analysis	Descriptive stats, regression analysis	
	2016	The effect of personality on IT personnel's	Survey *Online survey						
	2016	Team boundary spanning: strategic	case study * semi-structured interview* blog posts* archival data *log data						
	2016	Review of the empirical business services	Record * research literature			Time frame			
	2016	Netsourcing strategies for vendors: a	case study * semi-structured interview* blog posts* archival data *log data				within caseanalysis, cross caseanalysis		

	2016	Multiple interests or unified voice? Online	case study * semi-structured interview* blog posts* archival data *log data			Coding	Coding, interpretive analysis, Semantic analysis		
	2016	Linking macro-level goals to micro-level	case study * semi-structured interview* Observation* archival data			coding	Transcription, coding, Interpretive analysis		
	2016	Knowledge entrepreneurship: institutionalising wiki-based knowledge management	case study * semi-structured interview* Observation* archival data			coding	Transcription, coding, Interpretive analysis		
	2016	Internal crowdsourcing: conceptual	Record * research literature				coding, Interpretive analysis		
	2016	Identifying generative mechanisms through affordances: a framework for critical realist	case study * semi-structured interview* Observation* archival data			coding	Transcription, coding, Interpretive analysis		
	2016	e-Leadership through strategic alignment:				size, maturity stage of SME	concepts identification, categorization synthesising categories		

	2016	Determinant factors of cloud-sourcing	Record * research literature			Based on criteria (peer-reviewed research articles)			
JIT 2016 ends	2016	Are business users social? A design experiment exploring information sharing in enterprise social systems	Experiment*Experimental design *dev of website * Training *assigning task based on experimental condition * questionnaire (perception reported) *	recruited through thematic promotional releases in magazines, and mailed letters			descriptive statistics, PLS		
	2017	Towards a value theory for personal data	Experiment*Experimental design *online experiment* Training *assigning task based on experimental condition * questionnaire (perception reported) *			outliers		regression analyses	
	2017	Improving strategic flexibility with information technologies: insights for firm	Survey * paper-based questionnaire			missing values		EFA, CFA, PLS	

	2017	Exploring the effects of liminality on	case study * semi-structured interview* archival data, reports, newsletter				Transcription, coding, theme identification, categorization, Interpretive analysis		
	2017	Enhancing the measurement of information	case study * semi-structured interview* Group discussion, archival data, reports, newsletter				PLS) technique.		
	2017	Engaging scientometrics in information	Record * research literature			Keyword, cycling method	Coding, hermeneutically guided computer-assisted textual analysis (CATA) method		
	2017	Voluntary use of information technology: an	Record * research literature				Coding		
	2017	Information sharing and user behavior in	Record * database, Archival data, download			time frame, Data cleaning (based on requirements)	coding		

	2017	The regulatory, technology and market	case study * semi-structured interview* archival data, reports, commentaries						
	2017	How and why organisations use social	case study * semi-structured interview* archival data, reports, press release				thematic analysis, repertory grids		
	2017	Generative innovation: a comparison of	case study * semi-structured interview* observation * archival data, reports, press release				Comparison analysis		
	2017	Social Machines: how recent	case study * semi-structured interview* observation * archival data, reports, press release			Coding	Coding		
	2017	High-frequency trading and its role	Record * database, Archival data, download			missing data	Coding	Descriptive statistics	

	2017	High-frequency trading and conflict	case study * semi-structured interview* observation * archival data, reports, press release				Coding		
	2017	Crossing the next frontier: the role of ICT	Record * database, Archival data, download				Comparison analysis		
	2017	A taxonomy of financial market	Record * database, Archival data, download			Data cleaning, time frame	Cluster analysis		
	2017	The emergence of openness in opensource projects: the case of openEHR	case study * semi-structured interview* observation * archival data, reports, press release				Systematic coding, theme identification		
	2017	Opening the floodgates: the implications	Experiment*Experimental design *dev of website * Training *assigning task based on experimental condition * questionnaire (perception reported) *					descriptive statistics, OLS estimation	

	2017	Inclusive technologies, selective	case study * semi-structured interview* observation * archival data, reports, press release				Transcription, coding, theme identification, categorization, Interpretive analysis		
JIT 2017 ends	2017	Copy, transform, combine: exploring the remix as a form of innovation	Record * database, Archival data, download			copyright issues	coding, Visualizations	regression	
	2016	Why Individuals Participate in Micro-Task Crowdsourcing Work Environment: Revealing Crowdworkers' Perceptions (CS)	Survey *Online survey (open-ended questions)			coding	coding, Causal Mapping		
	2016	Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead	Record * research literature			based on criteria	Coding, categorization, classification		

	2016	Understanding the Role of IS and Application Domain A Knowledge on Conceptual Schema Problem Solving: A Verbal Protocol Study	Experiment*Experimental design * Training *assigning task based on experimental condition * questionnaire (perception reported) *				idiographic analyses, nomothetic analyses		
	2016	The Effects of Information Request Language Usage on Query Formulation Template	Experiment*Experimental design * Training *assigning task based on experimental condition * questionnaire (perception reported) *		Manipulation Check	Manipulation Check		MANOVA, descriptive statistics, Post Hoc Analyses	
	2016	Providing Utility to Utilities: The Value of Information Systems Enabled Flexibility in Electricity	Record * database, Archival data, download			Time frame		descriptive statistics	
	2016	Overconfidence in Phishing Email Detection	Experiment*Experimental design * Training *assigning task based on experimental condition * questionnaire (perception reported) *					Descriptive Analysis, OLS regression analysis	

	2016	Organizational Violations of Externally Governed Privacy and Security Rules: Explaining and Predicting Selective Violations under Conditions of Strain and	Record * database, Archival data,				Transcription, coding, Interpretive analysis		
	2016	Neural Correlates of Protection Motivation for Secure IT Behaviors: An fMRI Examination	Experiment**lab experiment *Experimental design * Training *assigning task based on experimental condition * questionnaire (perception reported) *			Coding	Cluster analysis, Interpretive analysis	regression analysis	
	2016	Mapping the Corporate Blogosphere: Linking Audience, Content, and Management to Blog Visibility	Record *, blogs					Descriptive Analysis, regression analysis	

	2016	Job Satisfaction in Agile Development Teams: Agile Development as Work Redesign	Survey *Online survey (open-ended questions)	recruited our sample using Empanel, a data-collection company that specializes in recruiting Internet-based survey panels		Rejected less than stipulated time		PLS	
	2016	Information Technology, Cross-Channel Capabilities, and Managerial Actions: Evidence from the Apparel Industry	Record * database, Archival data, news media			coding	coding, structured content analysis, interpretive analysis		
	2016	Information Systems Control: A Review and Framework for Emerging Information Systems Processes	Record * research literature				Mapping, interpretive analysis		
	2016	Inconsistent and Incongruent Frames During IT-enabled Change: An Action Research Study into Sales Process Innovation	case study * semi-structured interview* observation * archival data, reports, press release				Transcription, coding, theme identification		

	2016	Impression Formation and Durability in Mediated Communication	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * questionnaire (perception reported) *					CFA, MANOVA	
	2016	Ideational Influence, Connectedness, and Venue Representation: Making an Assessment of Scholarly Capita	Record * research literature				social network analysis, Venue Affiliation Analysis	PLS	
	2016	Graph-based Cluster Analysis to Identify Similar Questions: A Design Science Approach	Record * database, Archival data,				Cluster analysis		
	2016	Entangled Stakeholder Roles and Perceptions in Health Information Systems: A Longitudinal Study of the U.K. NHS N3 Network	case study * semi-structured interview* observation * archival data, reports, commentaries				Interpretive Analysis		

	2016	Enhancing Analysts' Mental Models for Improving Requirements Elicitation: A Two-stage Theoretical	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					CFA	
	2016	Culture, Conformity, and Emotional Suppression in Online Review	Record * database, Archival data,				coding, robustness checks		
	2016	Choosing a Fit Technology: Understanding Mindfulness in Technology Adoption and Continuance	Survey *Online survey			Based on criteria		PLS, Descriptive statistics, Post hoc analysis	
	2016	Business Process and Information Technology Alignment: Construct Conceptualization, Empirical	Record * database, Archival data,			Based on criteria	coding, interpretive analysis		

	2016	A Rift in the Ground: Theorizing Evolution of Values in Crowdfunding Communities Anchor	Record * database, Archival data,				open, axial, and selective coding		
JAIS 2016 ends	2016	A Knowledge-centric Examination of Signalling and Screening Activities in the Negotiation for Inform	case study * semi-structured interview* * archival data, reports, logs				Transcription, coding, Interpretive analysis		
	2017	Using Information Systems in Innovation Networks: Uncovering Network Resource	case study * semi-structured interview* * archival data, reports, logs				coding, central theme identification		
	2017	Use of Online Networking Services from a Theoretical Perspective of	Survey *Online survey			based on criteria		PLS, Descriptive Statistic	
	2017	Understanding User Adaptation toward a New IT System in Organizations: A Social Network Perspective	Survey * paper based* archival data					PLS, Descriptive Statistic	

	2017	Understanding Online Health Information Use: The Case of People with Physical Disabilities	Survey *Online survey	posted invitations in two general social networking				PLS	
	2017	The Role of Business Intelligence and Communication Technologies in Organizational Agility	Survey *Online survey					descriptive statistics and correlation	
	2017	The Online Waiting Experience:	Experiment**lab experiment * dev of website *Experimental design * Training	We recruited participants by distributing invitational	Manipulation Check			MANOVA and ANOVA, CFA	
			*assigning task based on experimental condition * interview (perception reported) *	fliers and making class announcements to undergraduate business students at two large U.S. universities					

	2017	The Doing of Datafication (And What this Doing Do	case study * semi-structured interview* * archival data, reports, logs				coding, central theme identification		
	2017	Shared Benefits and Information Privacy	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *			based on criteria (home owners)		PLS Analysis	
	2017	Robbing Peter to Pay Paul: Surrendering Privacy	Focus group				Transcription, coding, central theme identification, categorization	PLS	
	2017	Representing Crowd Knowledge: Guidelines for	case study * semi-structured interview* * archival data, reports, logs				Transcription, coding, central theme identification, categorization		

	2017	Political Manoeuvring During Business Process	case study * semi-structured interview* * archival data, reports, logs			coding	Transcription, coding, central theme identification, categorization		
	2017	Insights from an ICT4D Initiative in Kenya	case study * semi-structured interview* * archival data, reports, logs				Transcription, coding, central theme identification, categorization		
	2017	Heuristic Principles and Differential Judgments	case study * semi-structured interview* * archival data, reports, logs				Transcription, coding, central theme identification, categorization		
	2017	Health Information Systems and Accountability	case study * semi-structured interview* * archival data, reports, logs						
	2017	Got Phished? Internet Security and Human	Survey *Online survey				CFA		
	2017	Exploring the Dialectics Underlying	case study * semi-structured interview* * archival data, reports, logs				Transcription, coding, central theme identification, categorization		

	2017	Examining Real Options Exercise Decisions in	Survey *Online survey					Descriptive Statistics, ANOVA	
	2017	Evaluating Business Process Maturity Models	Record * research literature and Delphi * QUESTIONNAIRE			CODING	coding, assessment and re-assessments		
	2017	Effects of Animation on Attentional Resources o	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *		Manipulation and Control Checks			ANOVA, MANOVA	
	2017	Do Shareholders Value Green Information Technology	Record * database, Archival data,				transcription, coding, theme identification, categorization, interpretive analysis		
	2017	Controlling for Lexical Closeness in Survey Research	Survey *Online survey					CFA	

	2017	Challenging Dominant Frames in Policies for IS	case study * semi-structured interview* * archival data, reports, logs			Time frame	rhetorical strategy analysis		
	2017	Business Intelligence Capability_ The Effect of Top Management an	Survey *Online survey					descriptive statistics, correlation	
	2017	An Integrated Temporal Model of Belief and Attitude Change_ An	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					PLS	
	2017	An Extensive Examination of Regression Models with a Binary Outcome	Record * database, Archival data,					OLS	

	2017	Affect Infusion and Detection through Faces in Computer-mediated	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *		Manipulation Checks			Descriptive Statistics, PLS	
	2017	A Paradox of Progressive Saturation_ The Changing Nature of Impro	case study * semi-structured interview* observation				transcription, coding, pattern identification, categorization, interpretive analysis		
	2017	Information Communicating Personal Health an Integration of Privacy	Survey *Online survey			missing values		principal component analysis (PCA), regressions analysis	
EJS 2016		Factors influencing the intention to comply with data protection regulations in hospitals: based on gender differences in behaviour and deterrence	Survey *Online survey					PLS-SEM	

		Representing small business web presence content: the web presence pyramid model	case study * semi-structured interview* * archival data, reports, logs			Time frame	coding, interpretive analysis		
		Effectiveness of top management support in enterprise systems success: a contingency perspective of fit between leadership style and system life-cycle	case study * semi-structured interview* * archival data, reports, logs				transcription, coding, within-case analysis, cross-case analysis, theme identification, categorization		
		Mindful revolution or mindless trend? Examining agile development as a management fashion	case study * semi-structured interview*				transcription, coding, theme identification, categorization, comparative analysis		
		Handle mergers and acquisitions with care: the fragility of trust between the IT-service provider and end-users	Survey *Online survey					PLS-SEM, Descriptive statistics	
		Measuring eGovernment success: a public value approach	Survey *Online survey					CFA, Regression analysis	

		Achieving dynamic capabilities with cloud computing: an empirical investigation	case study * semi-structured interview*				transcription, coding, theme identification, categorization, interpretive analysis		
		Dispositional and situational factors: influences on information security policy violations	Survey *Online survey			based on condition		Descriptive statistics	
		Web assurance seal services, trust and consumers' concerns: an investigation of e-commerce transaction intentions across two nations	Survey *Online survey		added a series of questions to check the validity of response	invalid and incomplete responses		EFAs, PLS	
		Why different trust relationships matter for information systems users	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					PLS	

		Utilizing big data analytics for information systems research: challenges, promises and guidelines	Record * database, Archival data,			duplicate reviews	exploratory data analysis, Latent Dirichlet Allocation (LDA) algorithm, interpretive analysis		-3
		The inner and the outer model in explanatory design theory: the case of designing electronic feedback systems	Survey *Online survey				PLS, Post hoc analysis		
		Work routines as an object of resistance during information systems implementations: theoretical foundation and empirical evidence	Survey *Online survey				PLS, Post hoc analysis		
		Identity metamorphoses in digital disruption: a relational theory of identity	case study * semi-structured interview* * archival data, reports,			coding	transcription, coding, theme identification, categorization, interpretive analysis thematic analysis		

		How users perceive and respond to security messages: A Neurosis research agenda and empirical study	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *		manipulation-check				
		The impact of media on how positive, negative, and neutral communicated affect influence unilateral concessions during negotiations	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					Descriptive statistics, PLS	
		An empirical analysis of the factors and measures of Enterprise Architecture Management success	Survey *Online survey					PLS	

		Untangling the complex role of guilt in rational decisions to discontinue the use of a hedonic Information System	Survey *Online survey			based on criteria		Descriptive statistics, CFA	
		Triggered essential reviewing: the effect of technology affordances on service experience evaluations	Record * database, Archival data,			time frame		Descriptive statistics	-1
		Collaborative partner or opponent: How the messenger influences the deaf effect in IT projects	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					MANOVA, ANOVA, PLS	
		Fragmentation or cohesion? Visualizing the process and consequences of information system diversity, 1993–2012	Record * research literature			keyword frequency	Co-word analysis		

		Alignment in an inter-organisational network: the case of ARC transmittance	case study * semi-structured interview* * archival data, reports,				hermeneutic analysis		
EJIS 2017		Information systems security policy implementation in practice: from best practices to situated practices	case study * semi-structured interview* observation				description mode of analysis, s visualization		
		Mixed results in strategic IT alignment research: a synthesis and empirical study	Record * database, Archival data,				PLS, MANOVA		
		Examining the intended and unintended consequences of organisational privacy safeguards	case study * semi-structured interview* * archival data, reports,				Transcription, coding, categorization, interpretive analysis		
		Useful business cases: value creation in IS projects	case study * semi-structured interview* * archival data, reports,				Transcription, coding, categorization, interpretive analysis		-1

		The concept of mindfulness in information systems research: a multi-dimensional analysis	Record * research literature				Thematic analysis		
		The public procurement of information systems: dialectics in requirements specification	case study * semi-structured interview* * archival data, reports,				Transcription, coding, categorization, narrative analysis		
		Variable use of standards-based IOS enabling technologies in Australian SMEs: an examination of deliberate and emergent decision-making processes	case study * semi-structured interview* * archival data, reports, and Survey				Transcription, coding, categorization, narrative analysis	CFA	
		Interorganizational dependence, information transparency in interorganizational information systems, and supply chain performance	Record * database, Archival data,				(ANOVA, PLS		

		Conflicts and complements between eastern cultures and agile methods: an empirical investigation	case study * semi-structured interview* * archival data, reports, and Survey				Transcription and translation, coding, categorization, interpretive analysis		-3
		The funeral industry and the Internet: on the historical emergence and destabilization of strategic paths	case study * semi-structured interview* * archival data, reports,				Transcription, coding, categorization, narrative analysis		
		“What else is there...?”: reporting meditations in experiential computing	case study * semi-structured interview* * archival data, reports, and Survey				Transcription, coding, categorization, narrative analysis		
		Once upon a time: Crafting allegories to analyse and share the cultural complexity of strategic alignment	case study * semi-structured interview* * archival data, reports,				Transcription, coding, categorization, comparative analysis		

		Information quality, user satisfaction, and the manifestation of workarounds: a qualitative and quantitative study of enterprise content management system users	Survey *Online survey					PLS	
		It's complicated: explaining the relationship between trust, distrust, and ambivalence in online transaction relationships using polynomial regression analysis and response surface analysis	Experiment*simulation *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *		Manipulation checks	d unusable or missing data		PLS	
		Online product review as an indicator of users' degree of innovativeness and product adoption time: a longitudinal analysis of text reviews	Record * database, Archival data, Reviews					Regression analysis	

		Constraint-based and dedication-based mechanisms for encouraging online self-disclosure: Is personalization the only thing that matters?	Record * database, Archival data, websites					EFAs, PLS	
		Service robots in hospitals: new perspectives on niche evolution and technology affordances	case study * semi-structured interview* Q-methodology					CFA	
		Balancing exploration and exploitation of IT resources: the influence of Digital Business Intensity on perceived organizational performance	Survey * mail-based survey			Missing data		PLS	
		Impact of executive compensation on the execution of IT-based environmental strategies under competition	Survey *Online survey and *Archival data					CFA, PLS	

		The role of context in IT assimilation: A multi-method study of a SaaS platform in the US non-profit sector	Survey *Online survey and *Archival data			incomplete survey responses	permutation analysis	CFA	
		Taking stock of organisations' protection of privacy: categorising and assessing threats to personally identifiable information in the USA	case study * semi-structured interview* * archival data, reports,			Transcription, coding, categorization, Exploratory analysis			
		Which phish get caught? An exploratory study of individuals' susceptibility to phishing	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *				CFA, Post hoc analysis		
		Organizational information security policies: a review and research framework	Record * research literature			Time frame	Coding, theme identification		

		Antecedents and outcomes of information privacy concerns in a peer context: An exploratory study	Survey *Online survey					PLS	
		An empirical study on the susceptibility to social engineering in social networking sites: the case of Facebook	Experiment**lab experiment * role play *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					CFA, ANOVA, PLS	
		Adverse consequences of access to individuals' information: an analysis of perceptions and the scope of organisational influence	case study *Focus group				iterative analysis, Transcription and translation, coding, categorization, interpretive analysis		

ISJ 2016		Affordances of social media in collective action the case of Free Lunch for Children in China	case study * semi-structured interview* * archival data, reports,				Transcription and translation, coding, categorization, interpretive analysis		-1
		Business model development, founders 'social capital and the success of early stage internet start-ups a mixed-method study	case study * semi-structured interview* * archival data, reports,				Transcription and translation, coding, categorization, interpretive analysis	SNA	
		Business models and opportunity creation How IT entrepreneurs create and develop business models under uncertainty	case study * semi-structured interview* * archival data, reports,			Categorizati on	Transcription and translation, coding, categorization, interpretive analysis		
		Can the outside-view approach improve planning decisions in software development projects	Experiment**lab experiment *Experimental design * Training *assigning task based on experimental condition * interview				MANOVA, ANOVA, PLS		

			(perception reported) *					
		Cultures of participation—for students, by students	case study * semi-structured interview* observation* archival data, reports,			iterative analysis, Transcription and translation, coding, categorization, interpretive analysis		-2
		Developing ecological sustainability a green IS response model	case study * semi-structured interview* archival data, reports,		coding	iterative analysis, Transcription and translation, coding, categorization, interpretive analysis		

		Hitting a moving target, a process model of	case study * semi-structured interview* archival data, reports,			coding	Cross-case analysis, Transcription and translation, coding, categorization, comparative analysis		
		Impact sourcing ventures and local	case study * semi-structured interview* observation* archival data, reports,			coding	iterative analysis, Transcription and translation, coding, categorization, interpretive analysis		
		Internet aggression in online communities a contemporary deterrence perspective	Survey *Online survey				PLS		

	Leveraging virtual business model innovation, a framework for designing business model development tools	case study *Focus group				Transcription and translation, coding, categorization, interpretive analysis		
	Perceived barriers to effective knowledge	case study * semi-structured interview* observation* archival data, reports,				Transcription, coding, categorization, mapping, comparative analysis		
	Reconciling global and local needs a canonical action research project to deal with workarounds	case study * *Action research *semi-structured interview* observation* archival data, reports,			coding	Transcription and translation, coding, categorization, interpretive analysis		
	Reflections on Information Systems	case study * semi-structured interview* archival data, reports,				Transcription, coding, categorization, interpretive analysis		

		The influence of ethnicity on organizational commitment and merit pay of IT workers the	Survey *Online survey					PLS	-2
		Towards an understanding of the role of business intelligence systems in organisational knowing	case study * semi-structured interview,				iterative analysis, Transcription, coding, categorization, theme identification, interpretive analysis		
		Understanding the influence of absorptive capacity and ambidexterity on the process of business model change – the case of on-premise and cloud-computing software	case study * semi-structured interview* archival data, reports,				iterative analysis, Transcription, coding, categorization, theme identification, interpretive analysis, Within-case analysis, Cross-case analysis		

		Winner's regret in online C2C Auctions an automatic thinking perspective					Descriptive stats, CFA	
ISI 2017		Applying configurational analysis to IS behavioural research a methodological alternative for modelling combinatorial complexities	Survey *Online survey				PLS	
		Career transition antecedents in the information technology area	Survey *Online survey				Descriptive stats, PLS	
		Understanding buyers' loyalty to a C2Cplatform the roles of social capital, satisfaction and perceived effectiveness of-commerce institutional mechanisms	Survey *Online survey				Descriptive stats, PLS	

		User behaviours after critical mobile application incidents the relationship with situational context	Survey *Online survey					Descriptive stats, ANOVA	
		A typology of user liability to IT addiction	Survey *Online survey, interview				coding	ANOVA	
		Digitally enabled disaster response the emergence of social media as boundary objects in a flooding disaster	case study * semi-structured interview* archival data, reports,				iterative analysis, Transcription, coding, categorization, theme identification, interpretive analysis		
		Service encounter thinklets how to empower service agents to put value co-creation into practice	Experiment**lab experiment *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					MANOVA, ANOVA, PLS	

	An activity theoretic analysis of the mediating role of information systems in tackling climate change adaptation	case study * semi-structured interview* archival data, reports,				Transcription, coding, categorization, theme identification, interpretive analysis		
	Characteristics of IT artefacts a systems thinking-based framework for delineating and theorizing IT artefacts	Record * research literature				Coding, interpretive analysis		
	How does business analytics contribute to business value	Record * database, Archival data, websites				Coding, interpretive analysis		
	Driving business transformation toward sustainability exploring the impact of supporting IS on the performance contribution of eco-innovations	case study * semi-structured interview* observation* archival data, reports,				iterative analysis, Transcription, coding, categorization, theme identification, interpretive analysis		

		How IT executives create organizational benefits by translating environmental strategies into Green IS initiatives	Survey *Online survey			incomplete survey responses, outliers		PLS	
		Information systems absorptive capacity for environmentally driven IS-enabled transformation	Survey *Online survey					PLS	
		Winning the SDG battle in cities how an integrated information ecosystem can contribute to the achievement of the 2030sustainable development goals	case study * semi-structured interview* archival data, reports,			iterative analysis, Transcription, coding, categorization, theme identification, interpretive analysis			

		Applying a critical approach to investigate barriers to digital inclusion and online isocyanato king among young people with disabilities	case study * semi-structured interview				Transcription, coding, categorization, theme identification, Thematic analysis		
		Open gentrification	case study * semi-structured interview* observation* archival data, reports,				iterative analysis, Transcription, coding, categorization, theme identification, mapping, interpretive analysis		
		Starting open source collaborative innovation, the antecedents of network formation in community source	case study * semi-structured interview* observation* archival data, reports,				iterative analysis, Transcription, coding, categorization, theme identification, interpretive analysis		

	The roles of mood and conscientiousness in reporting of self-committed errors on IT projects	Experiment**lab experiment *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					Descriptive stats, PLS	
	Entering the field in qualitative field research a rite of passage into a complex practice world	case study * semi-structured interview* observation				Transcription, coding, categorization, theme identification, interpretive analysis		
	Issues that support the creation of ICT workarounds towards a theoretical understanding of feral information systems	case study * semi-structured interview* observation* archival data, reports,				Transcription, coding, categorization, theme identification, interpretive analysis		-1

		Risks to Effective Knowledge Sharing in Agile Software Teams A Model for Assessing and Mitigating Risks	case study * semi-structured interview* observation* archival data, reports,				Transcription, coding, categorization, theme identification, interpretive analysis Within-case analyses, Cross-case analysis		
ISR 2016		Why Do Adults Engage in Cyberbullying on Social Media? An Integration of Online Disinhibition and Deindividuation	Survey *Online survey (CS)	Platform selection				PLS	CS used for data collection
		Turbulent Stability of Emergent Roles: The Dualistic	Record * database, Archival data, websites, articles				coding, categorization, theme identification, interpretive analysis		
		Toward a Theory of Remixing in Online Innovation	Record * database, Archival data, posts				incomprehensible, inconsistent with our focus, Uncategorized posts	Descriptive Statistics, Regression Analyses	

	The Impact of Fake Reviews on Online Visibility: A	Record * database, Archival data, reviews			Time frame			
	The Impact of Competing Ads on Click Performance in	Record * database, Archival data, reviews			Time frame, keywords			
	The Double-Edged Sword of Backward Compatibility: The	Record * database, Archival data, reviews					Model fit	
	Secret Admirers: An Empirical Examination of Information	Record * database, Archival data, reviews					Descriptive Statistics, Correlation Matrix	
	Research Note—When Do Consumers Value Positive	Record * database, Archival data, reviews					Descriptive Statistics and Correlations, Cross-Sectional Analysis of Reviews	
	Research Note—Using Expectation Disconfirmation	Record * database, Archival data, reviews			Time frame		PLS	
	Research Note—IT Outsourcing and the Impact of Advisors	Record * database, Archival data, contracts			Time frame		Descriptive Statistics, Correlation Matrix, Sensitivity Analysis	

		Research Note— Information Technology, Customer	Record * database, Archival data, contracts			Time frame		Descriptive Statistics	
		Research Note—In CARSs We Trust: How Context-Aware	Experiment* *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					Descriptive stats, PLS	
		Research Note— Gamification of Technology-Mediated	Experiment* *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					Descriptive stats, PLS, ANCOVA, ANOVA	
		Research Note— Designing Promotion Ladders to Mitigate	Record * database, Archival data, records			Time frame			
		Research Note— Content and Collaboration: An Affiliation	Record * database, Archival data, projects, articles					Descriptive Statistics and Variable Correlation s	

		Research Note—Cloud Computing Spot Pricing Dynamics:	Record * database, Archival data, projects, articles			Time frame		Descriptive Statistics	
		Research Note—Are Online Labor Markets Spot Markets for (CS)	Experiment* *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *	Through an interface provided by Mturk			Interpretive analysis		
		Research Note—An Internet-Enabled Move to the Market	Record * database, Archival data, projects, articles					Descriptive Statistics	
		Rate or Trade? Identifying Winning Ideas in Open Idea	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception			Based on conditions		Descriptive Statistics and Correlations, Regression Analysis	

			reported) *					
		Product Line Extension in Consumer Software Markets in (CS)	Survey *Online survey				Descriptive Statistics and Correlations,	
		Open Content, Linus' Law, and Neutral Point of View	Record * database, Archival data, projects, articles				Descriptive Statistics, Regression Analysis	
		On the Longitudinal Effects of IT Use on Firm-Level	Record * database, Archival data, projects, articles				Descriptive Statistics	
		More Harm Than Good? How Messages That Interrupt Can	Experiment**lab experiment * dev of website		We excluded those with		fMRI analysis, ANOVA	

			*Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *			colour blindness			
		Managing Citizens' Uncertainty in E-Government Services	Survey *Online survey				regression analyses, Descriptive Statistics and Correlations		
		Intellectual Property Norms in Online Communities: How	case study * semi-structured interview* survey				coding, interpretive analysis		
		Implementation of an Information and Communication	case study * semi-structured interview* survey				coding, interpretive analysis		

		Framing Innovation Opportunities While Staying	case study * semi-structured interview* observation* archival data, reports,				coding, interpretive analysis		
		Folding and Unfolding: Balancing Openness and	case study * semi-structured interview* observation* archival data, reports,			coding	Transcription, coding, categorization, theme identification, interpretive analysis		
		Facilitating the Transformational: An Exploration of	case study * semi-structured interview* observation* archival data, reports,			coding	Transcription, coding, categorization, theme identification, interpretive analysis		
		Excessive Dependence on Mobile Social Apps: A Rational	Record * database, Archival data, projects, articles					descriptive statistics	

		Examining the Continuance of Secure Behavior: A	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					Descriptive Statistics, ANOVA, PLS	
		Effect of Knowledge-Sharing Trajectories on Innovative	Record * database, Archival data, projects, articles			coding, categorization,		OLS regression, Descriptive Statistics	
		Does Product Market Competition Drive CVC Investment?	Record * database, Archival data, projects, articles			Time frame		descriptive statistics	
		Creating Value in Online Communities: The Sociomaterial	case study * semi-structured interview* observation* archival data, reports,				Transcription, coding, categorization, theme identification, interpretive analysis	and the correlation	

		Coordinating Interdependencies in Online Communities: A	Record * database, Archival data, projects, articles and interviews				Exploratory data analysis, Transcription, coding, categorization, theme identification, interpretive analysis	ANCOVA, regression	
		Comparing Open and Sealed Bid Auctions: Evidence from	Record * database, Archival data, projects, articles and observation					regression analysis	
		Characteristics and Economic Consequences of Jump Bids	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *	solicitation emails sent out to all registered undergraduate students at the school				Descriptive Statistics, regression analysis	
		Assessing the Impact of Granular Privacy Controls on	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental					Descriptive Statistics, regression analysis	

			condition * interview						
			(perception reported) *						
ISR 2017		Does Information Technology Improve Open Innovation	Record * database, Archival data, projects, articles						
		Governance Practices in Platform Ecosystems: Navigating	case study * semi-structured interview* observation* archival data, reports,				Transcription, coding, categorization, theme identification, interpretive analysis		
		Disconfirmation Effect on Online Rating Behavior	Record * database, Archival data, projects, articles					Descriptive Statistics	

		Continued Voluntary Participation Intention in Firm-	Record * database, Archival data, projects, articles					Descriptive Statistics and Correlation	
								S	
		Unravelling the Alignment Paradox: How Does Business—IT	Survey *paper-based survey					regression	
		Understanding Consumers' Attitudes Toward Controversial	Survey *Online survey					PLS, CFA	
		Temporal Motivations of Volunteers to Participate in	case study * semi-structured interview* observation* archival data, reports,				thematic data analysis		
		Technical Systems Development Risk Factors: The Role of	Survey *paper-based survey					Descriptive Statistics, Correlations, OLS	
		Seeking Value Through Deviation? Economic Impacts of IT	Record * database, Archival data, projects, articles					Descriptive Statistics, OLS	

	Mitigating Diminishing Returns to R&D: The Role of	Record * database, Archival data, projects, articles					Correlations and Descriptive Statistics	
	Learning Effects of Domain, Technology, and Customer	Record * database, Archival data, projects, articles					Correlations and Descriptive Statistics	
	Direct and Indirect Information System Use: A	Survey *paper-based survey					Correlations and Descriptive Statistics, Post Hoc Analyses	
	A Theory of Responsive Design: A Field Study of Corporate	case study * semi-structured interview* observation* archival data, reports,				Thematic Analysis.		
	Electronic Commerce, Spatial Arbitrage, and Market	Record * database, Archival data, projects, articles					Descriptive Statistics	
	Managerial Incentives and IT Strategic Posture	Record * database, Archival data, projects, articles					Descriptive Statistics, OLS	
	Extending the Concept of Control Beliefs: Integrating the	Survey *Online survey					Descriptive Statistics and Correlation	

							s, Regression	
		Politics and Information Technology Investments in the	Record * database, Archival data, projects, articles				Descriptive Statistics	
		Popularity or Proximity: Characterizing the Nature of	Record * database, Archival data, projects, articles				Correlations and Descriptive Statistics	
		Providing a Window of Opportunity for Converting eStore	Record * database, Archival data, projects, articles				Descriptive Statistics	
		Research Commentary—Diversity of the Information	Record * research literature				coding, categorization, interpretive analysis	
		Anonymizing and Sharing Medical Text Records	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview				Cluster analysis	

			(perception reported) *					
		Coping Responses in Phishing Detection: An Investigation	Survey *Online survey				PLS, Post Hoc Analysis	
		Cosearch Attention and Stock Return Predictability in	Record * database, Archival data, projects, articles				Correlations and Descriptive Statistics, Co-viewing Analysis	
		Designing for Diagnosticity and Serendipity: An	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *				MANOVA, ANOVA, PLS	

		Information Feedback, Targeting, and Coordination: An	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					Regression	
		Information Technology, Revenues, and Profits: Exploring	Record * database, Archival data, projects, articles					Descriptive Statistics	
		Logic Pluralism in Mobile Platform Ecosystems: A Study of	case study * semi-structured interview* observation* archival data, reports,				Transcription, coding, categorization, theme identification, interpretive analysis		
		The Impact of Institutional Distance on the Joint	Survey *Online survey					PLS, Correlations and Descriptive Statistics	

		How Can We Develop Contextualized Theories of	case study * semi-structured interview* focus groups* archival data, reports,			coding	Transcription, coding, categorization, theme identification, interpretive analysis		
		Is Tom Cruise Threatened? An Empirical Study of the	Record * database, Archival data, projects, articles					Descriptive Statistics, PLS, Regression	
		On Buyer Selection of Service Providers in Online	Record * database, Archival data, projects, articles					Descriptive Statistics, PLS,	
		Understanding Voluntary Knowledge Provision and	Experiment**lab experiment * dev of website *Experimental design * Training *assigning task based on experimental condition * interview (perception reported) *					Correlation , OLS	

Form D

Crowdsourcing Theory Testing Activities

Article Title	Reference	Intent	Activities	How to Crowdsource	Reason
How Can We Develop Contextualized Theories of Effective Use? A Demonstration in the Context of Community-Care Electronic Health Records	Burton-Jones, A., & Volkoff, O. (2017). How can we develop contextualized theories of effective use? A demonstration in the context of community-care electronic health records. <i>Information Systems Research, 28(3)</i> , 468-489.	Get group perceptions (Focus Group)	Select Outcome of interest	Cannot be Crowdsourced	
			Prepare guide	Cannot be Crowdsourced	
			Select participants	* Select the desired CS platform that suits the study * Screen/profile crowd or specify desired crowd * Specify number of participants needed * Select participants based on criteria	

			Moderate discussion	<p>Assign Task to selected participant(s)</p> <ul style="list-style-type: none"> • Give participants details regarding <ul style="list-style-type: none"> - Selected site - Desired participants in the discussion - Questions to be asked and discussion - Give desired output information to participants (i.e. Recordings, video, etc.) - Give timeline for the said task 	
			Get group perceptions	<ul style="list-style-type: none"> * Contact participants after deadline for task completion has expired * Collect desired output information from participants (i.e. Recordings, video, etc.) 	
		Get individual perception (Semi-structured interview)	Select outcome of interest	Cannot be Crowdsourced	
			Prepare questionnaire/guide	Cannot be Crowdsourced	

			Select participants	<ul style="list-style-type: none"> * Select the desired CS platform that suits the study * Screen/profile crowd or specify desired crowd * Specify number of participants needed * Select participants based on criteria 	
			Moderate discussion	<p>Assign Task to selected participant(s)</p> <ul style="list-style-type: none"> • Give participants details regarding <ul style="list-style-type: none"> - Selected site - Desired participants in the discussion - Questions to be asked and discussion - Give desired output information to participants (i.e. Recordings, video, etc.) - Give timeline for the said task 	
			Get individual perceptions	<ul style="list-style-type: none"> * Contact participants after deadline for task completion has expired * Collect desired output information from participants (i.e. Recordings, video, etc.) 	
		Data extraction from artefacts (Data from Documents, projects, emails, etc)	Select records	Cannot be Crowdsourced	

			Select artefact	Cannot be Crowdsourced	
			Select participants	<ul style="list-style-type: none"> * Select the desired CS platform that suits the study * Screen/profile crowd or specify desired crowd * Specify number of participants needed * Select participants based on criteria Assign Task to selected participant(s) 	
			Data extracted from artefacts	<ul style="list-style-type: none"> • Give participants details regarding <ul style="list-style-type: none"> - Artefacts and records selected - Desired data needed from the artefacts - Give desired output information to participants (i.e. Recordings, video, write-ups etc.) - Give timeline for the said task - Contact participants after deadline for task completion has expired - Collect desired output information from 	

Web assurance seal services, trust and consumers' concerns: an investigation of e-commerce transaction intentions across two nations	Kim, D. J., Yim, M. S., Sugumaran, V., & Rao, H. R. (2016). Web assurance seal services, trust and consumers' concerns: an investigation of e-commerce transaction intentions across two nations. <i>European Journal of Information Systems</i> , 25(3), 252-273.	Get individual perception (Online survey)	Select outcome of interest	Cannot be Crowdsourced	
			Develop questionnaire	Cannot be Crowdsourced	
			Select participants	<ul style="list-style-type: none"> * Select the desired CS platform that suits the study * Screen/profile crowd or specify desired crowd * Specify number of participants needed * Select participants based on criteria 	
			Deliver questionnaire	<ul style="list-style-type: none"> * Send survey link to participants via desired or selected method, e.g. CS platform, Email, social media etc. 	
			Get individual perceptions	<ul style="list-style-type: none"> * Contact participants after deadline for task completion has expired * Collect desired output information from participants 	

				(i.e. completed survey or response etc.)	
Affect Infusion and Detection through Faces in Computer-mediated Knowledge-sharing Decisions	Fehrenbacher, D. D. (2017). Affect Infusion and Detection through Faces in Computer-mediated Knowledge-sharing Decisions. <i>Journal of the Association for Information Systems</i> , 18(10), 703-726.	Get individual perceptions (Lab Experiment)	Develop environment (Website development)	<p>Can be difficult to CS if its development cannot be done online</p> <ul style="list-style-type: none"> * Select the CS platform that suits the study * Select participants based on skill <ul style="list-style-type: none"> * Give the desired design, information and outlook of the proposed website * Give a time frame * Collect output from participant * Check quality and select the best 	<p>It is possible that some of the activities cannot be CS, especially if the lab has to be physical and cannot be simulated or done online</p>
			Select participants	<ul style="list-style-type: none"> * Select the desired CS platform that suits the study * Screen/profile crowd or specify desired crowd * Specify number of participants needed * Select participants based on criteria 	<p>Can only be done if the lab experiment can be done online or the task is CS to an individual or group to physically recruit participants and run the experiment</p>
			Assign roles	* Based on experimental design, assign roles to participants with conditions	

			Train participants	* Give a brief demonstration of what is expected, can be done by giving a list of instructions, link that leads to a recorded demonstration etc.	For CS to be used, training has to be done online, or the researcher goes a step forward by video recording the training session with demonstrations of what participants should know and giving the recordings to participants, if training cannot be done online, then CS cannot be used
			Run task	* Give a time frame for the completion of the activity	
			Deliver questionnaire	* Get individuals perception by sending survey link to participants via desired or selected method, e.g. CS platform, Email, social media etc. or by asking participants to describe their perceptions	
			Get individual perceptions	* Contact participants after deadline for task completion has expired * Collect desired output information from participants (i.e. completed survey or response etc.)	

Form E

Theory Testing Activities and Crowdsourcing Outcomes

Activity	Cannot Crowdsource (X)	Difficult to Crowdsource (D)	Easy to Crowdsource (E)	Reason
Prepare guide	✓			CT, NATC
Develop Questionnaire	✓	✓		CT
Select participants			✓	PC, SC,
Moderate discussion			✓	WT, RT
Get group perceptions			✓	WT, RT
Develop environment		✓		CT, CDT
Develop Algorithm		✓		CT, CDT
Assign roles			✓	PC, SC, WT
Train participants		✓		CDT, CT
Run task			✓	PC, SC, WT, PT
Deliver questionnaire			✓	PC, SC, WT, PT, RT
Get individual perception			✓	PC, SC, WT, PT, RT
Select records	✓			CT
Select artefact	✓			CT
Data extraction from artefact			✓	PC, SC, WT, PT, RT
Select outcome of interest	✓			CT
Select environment	✓			CT
Observation			✓	PC, SC, WT, PT, RT
Prepare poll	✓			CT

Prepare review	✓			CT
Select panel			✓	PC, SC, WT, PT, RT
Request feedback	✓			CT
Collect ratings			✓	PC, SC, WT, PT, RT
Collect expert opinions		✓		
Transcription			✓	PC, SC, WT, PT, RT
Screening and cleaning			✓	PC, SC, WT, PT, RT
Categorization			✓	PC, SC, WT, PT, RT
Coding		✓	✓	PC, SC, WT, PT, RT
Generation		✓	✓	PC, SC, WT, PT, RT
Content analysis			✓	PC, SC, WT, PT, RT
Statistical analysis	✓			CT

Where:

NATC represents No access to context

CDT = Coordinated Task

WT = Well-defined Task

RT = Remote ask

SC = Screened Crowd

PC = Profiled Crowds

CT = Contextual Task

PT = Pooled Task

Appendix B. Ethics Approval

This research involves human participants in its activities: Card sorting, Interviews and observation. The human ethics application was approved by the Human and Ethics Committee of the School of Information Management at Victoria University of Wellington, New Zealand.

Reference: 0000025042

Information sheet and consent form are presented below.

Form A

Information Sheet



Decision Support Tool for Theory Testing

INFORMATION SHEET FOR PARTICIPANTS

Thank you for your interest in this project. Please read this information before deciding whether or not to take part. If you decide to participate, thank you. If you decide not to take part, thank you for considering my request.

Who am I?

My name is Enwereuzo Ijeoma and I am a Doctoral student in Information Management at Victoria University of Wellington. This research project is work towards my dissertation.

What is the aim of the project?

This project aims to develop a decision support tool for theory testing. This tool will assist researchers make decision as to what the best method to select to test a particular theory, and to give recommendations as to what activity(ies) within the theory testing process can be crowdsourced or not.

This research has been approved by the School of Information Management's Human Ethics Committee.

How can you help?

If you agree to take part, I will engage you in a card sorting task. You will be asked to sort some cards according to their respective categories, and a brief interview will be conducted. The cards are made from sheets of paper, square in shape and contains words or sentences which will be sorted based on different categories. The task will take about an hour and will be conducted in one of the meeting rooms

on level 4 of the Rutherford House, at the Pipitea campus. The task will be conducted at your convenient and free time. I will audio record the interview with your permission and write it up later. You can choose to not answer any question or stop the discussion at any time, without giving a reason. You can withdraw from the study by contacting me at three weeks after the interview. If you withdraw, the information you provided will be destroyed or returned to you.

What will happen to the information you give?

This research is confidential. This means that the researchers named below will be aware of your identity, but the research data will be combined, and your identity will not be revealed in any reports, presentations, or public documentation.

Only my supervisors and *myself*, will read the notes or transcript of the task and discussion. The discussion and task transcripts, summaries and any recordings will be kept securely and destroyed 5 years after the research ends.

What will the project produce?

The information from my research will be used in my PhD dissertation and in academic publications and conferences.

If you accept this invitation, what are your rights as a research participant?

You do not have to accept this invitation if you don't want to. If you do decide to participate, you have the right to:

- choose not to answer any question;
- ask for the recorder to be turned off at any time during the discussions;
- withdraw from the study three weeks after the interview;
- ask any questions about the study at any time;
- receive a copy of your interview recording;
- receive a copy of your interview transcript;
- read over and comment on a written summary of your interview;
- be able to read any reports of this research by emailing the researcher to request a copy.

If you have any questions or problems, who can you contact?

If you have any questions, either now or in the future, please feel free to contact:

Student:

Name:

University email address:

Supervisors:

Name:

Role:

School:

Phone:

Human Ethics Committee information

If you have any concerns about the ethical conduct of the research, you may contact the Victoria University HEC Convenor: Associate Professor Susan Corbett. Email susan.corbett@vuw.ac.nz or telephone +64-4-463 5480.

Form B

Consent form



Decision Support Tool for Theory Testing

CONSENT TO PARTICIPATE

This consent form will be held for 5 years.

Researcher: Ijeoma Enwereuzo, School of Information Management, Victoria University of Wellington.

- I have read the Information Sheet and the project has been explained to me. My questions have been answered to my satisfaction. I understand that I can ask further questions at any time.
- I agree to take part in a card sorting exercise and interview.

I understand that:

- I may withdraw from this study three weeks after the interview, and any information that I have provided will be returned to me or destroyed.
- The information I have provided will be destroyed 5 years after the research is finished.
- Any information I provide will be kept confidential to the researcher and the supervisors.
- I understand that the results will be used for a PhD dissertation, academic publications and presented at conferences.
- My name will not be used in reports, nor will any information that would identify me.
- I would like a copy of the recording of my interview: Yes No
- I would like a summary of my task and interview: Yes No
- I would like to receive a copy of the final report and have added my email address below. Yes No

Signature of participant: _____

Name of participant: _____

Date: _____

Contact details: _____