

# **Patterns of Testing Theory with Human Subjects: A Design Science Perspective**

*Completed Research Full Papers*

**Ijeoma Enwereuzo**

Victoria University of Wellington  
Ijeoma.enwereuo@vuw.ac.nz

**Pedro Antunes**

Victoria University of Wellington  
Pedro.antunes@vuw.ac.nz

**David Johnstone**

Victoria University of Wellington  
David.johnstone@vuw.ac.nz

## **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.

## **Keywords**

Theory, theory testing, design science, pattern theory.

## **Introduction**

The key foundations to advance scientific knowledge include the testing and re-testing of theory (Niederman and 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 and Zapata-Phelan 2007). Theory testing thus requires collecting empirical data either confirming or disconfirming the set of propositions articulated by a theory (Lokke and Sorensen 2014). Depending on the research discipline, empirical data can be collected from various types of sources and using various types of instruments.

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 characterize the essential aspects of the phenomena of interest covered by the IS field: the relationships between humans, organizations and technology. Identifying and characterizing 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 (Bhattacharjee 2012). The confluence between humans, organizations, and technology generates varied and rich types of data, which the researcher has to analyze 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 conceptualize 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 emphasizes the development of innovative and useful IS artefacts (Hevner and 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 synthesizes 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.

## **Research Approach**

We adopt the DSR paradigm for this study (Hevner et al. 2004). DSR supports problem-solving research centered on the creation of innovative artefacts. By artefact we mean anything that can be developed as an artificially made object (Gregor and Hevner 2013). In that category, we include abstractions such as the framework and model proposed in this study (Gregor and Hevner 2013). The DSR paradigm articulates three research cycles: relevance, rigor, 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 and Chatterjee 2010). In our case, the application environment concerns researchers wanting to test theory, and relevance concerns improving research processes involving theory testing. The rigor 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 and 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”).

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 synthesized 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 generalized solutions to recurring design problems (Zigurs and Khazanchi 2008). In the DSR context, the existing theory used in artefact development has been designated kernel theory (Gregor and Jones 2007). The adoption of kernel theory in artefact development is considered important to strengthen the relationship between artefact and knowledge base (Gregor and 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 realization of the pattern model.

In DSR, artefact evaluation seeks to deliver two complementary types of contributions (Sonnenberg and vom Brocke 2011). One is a 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 the general knowledge about artefact development (Gregor and 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. Next, we discuss the development of the first artefact in more detail.

## **Conceptual Framework**

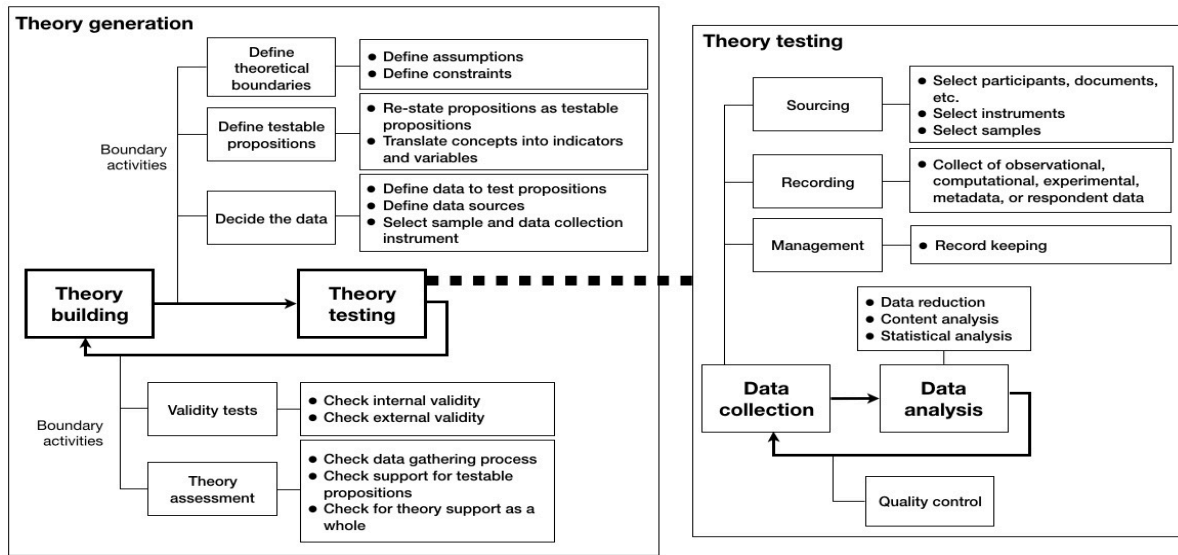
Theory helps scientists logically describe and explain a phenomenon, process or sequence of events. Gregor (2002; Gregor 2006) suggests a classification into five theory types. A Type I theory is descriptive,

answering the “what is” question by seeking to characterize 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 occur (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). In this paper, we focus on Type III and Type IV theories because they both involve predictive elements focused on testable propositions. Goode and 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.

Researchers spend time building theory; and then spend more time testing theory (Viswesvaran and Ones 1995). De Vaus (2013) and Bitektine (2007) developed conceptual frameworks that describe the major activities involved in theory building. They include important concepts such as the definition and operationalization 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 focused on theory testing, we need to disentangle the two concepts. To accomplish that goal, we first conceptualize the theory generation cycle (Figure 1), 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 what happens in the world (Lynham 2002). This simple, mutually re-enforcing, the 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 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. 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 and 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).

Testing can be conceptualized 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



**Figure 1. Conceptual framework of theory testing**

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.

## 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 systematic 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 (Paré et al. 2016).

## 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 analyzing data that reflects the frequency of topics found in the literature. The review procedure includes the following stages (Okoli and 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 recognized 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 building. We use a 2-year period for the review because this frame seems sufficient to develop a comprehensive view of the topic. Furthermore, during the review, we found that we got close to a saturation point, where not much new data is collected (Fusch and 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 and 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 the 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 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). After this stage, we had a collection of 404 articles to review.

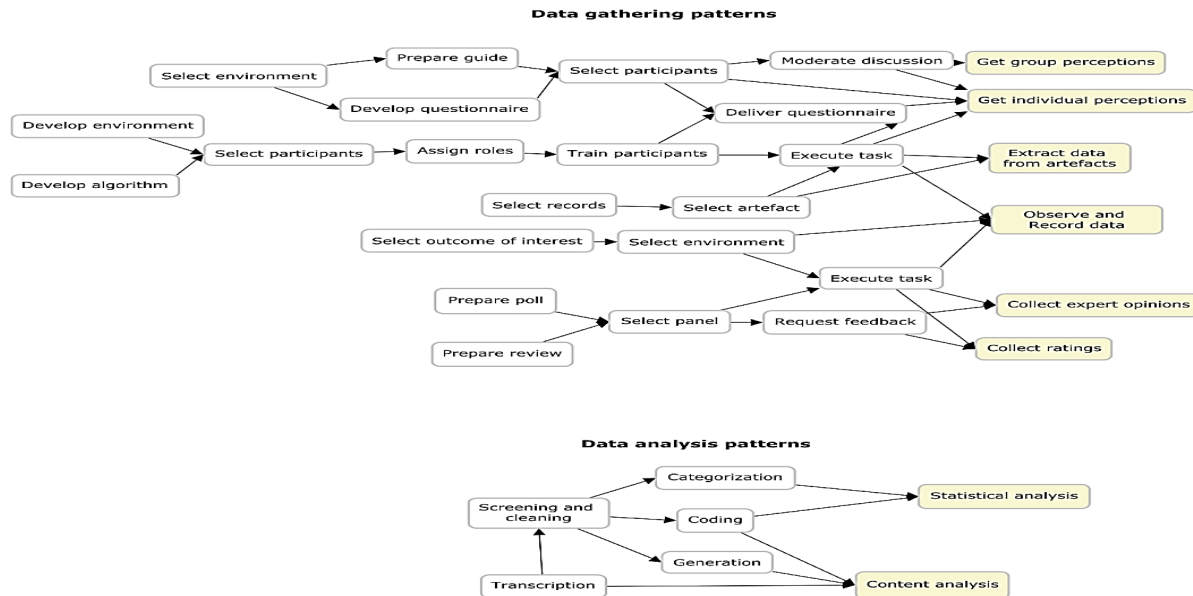
**Data Extraction.** This stage involves a detailed reading of specific sections of each article and then coding relevant data for posterior analysis. According to Paré et al. (2016), there are three main operational aspects of data extraction. The first one is deciding what parts of an article should 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. 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 synthesized into a usable artefact. To synthesize the review data, we adopt pattern theory, which is discussed below.

### ***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 and 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 generalized solutions to classes of problems (Bider et al. 2013). 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 organization of activities within a pattern, which define a coherent system of knowing what to do (Alexander 1999). And *implementation* considers how the pattern realizes 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 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 support. They are activities that need to be carried out in order 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. 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 2 and discussed below.



**Figure 2. 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.**

The analyzed body of research exhibits a clear separation of concerns between data collection and data analysis, which supports the theory testing cycle defined in Section 3 and illustrated in Figure 1. We start the literature synthesis by discussing the 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. In order to build the patterns, we started from the primary activities and then analyzed their relationships with secondary activities by moving backward. 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 analyze these patterns is to start from the end and move backward. 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, socio-material 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 and Jarvenpaa 2016) and (Li et al. 2017) are two examples), and can use various techniques, such as open and axial coding (Miles and 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)).

## 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 and 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 and Warfel 2004). 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 conceptualize 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 Ph.D. students conducting research projects in the IS field and IS Academics. Ph.D. students were selected because they represent the main target audience for this research, while the IS Academics were also used in order to get their opinions and feedback based on their years of experience. 15 participants (10 Ph.D. 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. 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 also to cater for the diversity of research problems and methods that are typical in the IS field. The exercises were then operationalized 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, focusing 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.

### ***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 80% were able to sort the cards into the right categories. 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 in order 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.

### ***Understanding of Pattern Model (E2)***

This step evaluated how easy it was for participants to interact and relate to the pattern model and apply the model to their current project. 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, 4 of the participants said that it was very easy to understand the model and relate that to their research project, 9 of the participants said it was somewhat easy, and 1 participant responded that it was not easy at all to understand. 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 'backward' 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.

### ***Understanding of the Elements defined by the Pattern Model (E3)***

This step evaluated how easy it was for participants to understand the elements described in 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 labeled "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.

### ***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 conceptualize their research project. From the result, 40% of the participants said that the model was very useful, while 60% said it was somewhat useful. 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 seen as 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.

## 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 realized that very often theory testing per se has not been considered as an independent phenomenon of interest. It is present in most research endeavors but is often entangled with theory building and taking a subsidiary role.

Our first undertaking consisted in dissociating 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 and 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 organized 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 distills 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 understanding 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 utilize principles and technological rules in their endeavors (in this case, theory testing).

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 emphasizes 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.

## 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), pp. 71-82.
- Bacharach, S. 1989. "Organizational Theories: Some Criteria for Evaluation," *Academy of management review* (14:4), pp. 496-515.
- Barrett, A., and Edwards, J. 1995. "Knowledge Elicitation and Knowledge Representation in a Large Domain with Multiple Experts," *Expert Systems with Applications* (8:1), pp. 169-176.
- Bhattacharjee, A. 2012. "Social Science Research: Principles, Methods, and Practices".

- Bitektine, A. 2007. "Prospective Case Study Design: Qualitative Method for Deductive Theory Testing," *Organizational Research Methods* (11:1), pp. 160-180.
- Colquitt, J., and 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), pp. 1281-1303.
- De Vaus, D. 2013. *Surveys in Social Research*. Routledge.
- Estelles-Arolas, E., and Gonzalez-Ladron-de-Guevara, F. 2012. "Towards an Integrated Crowdsourcing Definition," *Journal of Information Science* (38:2), pp. 189-200.
- Fusch, P., and Ness, L. 2015. "Are We There Yet? Data Saturation in Qualitative Research," *The qualitative report* (20:9), p. 1408.
- Goode, W., and Hatt, P. 1952. "Methods in Social Research,").
- Gregor, S. 2006. "The Nature of Theory in Information Systems," *MIS quarterly* (30:3), pp. 611-642.
- Gregor, S., and Hevner, A. 2013. "Positioning and Presenting Design Science Research for Maximum Impact," *MIS quarterly* (37:2), pp. 337-355.
- Hevner, A., and Chatterjee, S. 2010. *Design Research in Information Systems: Theory and Practice*. Springer Science & Business Media.
- Hevner, A., March, S., Park, J., and Ram, S. 2004. "Design Science in Information Systems Research," *Management Information Systems Quarterly* (28:1), pp. 75-105.
- Kampffmeyer, H., and Zschaler, S. 2007. "Finding the Pattern You Need: The Design Pattern Intent Ontology," in *International Conference on Model Driven Engineering Languages and Systems*. Berlin, Heidelberg: Springer, pp. 211-225.
- Kumar, R. 2005. *Research Methodology*. Australia: Pearson.
- Lokke, A., and Sorensen, P. 2014. "Theory Testing Using Case Studies," *Electronic Journal of Business Research Methods* (12:1), pp. 66-74.
- Lu, Y., Gupta, A., Ketter, W., and Van Heck, E. 2016. "Exploring Bidder Heterogeneity in Multichannel Sequential B2b Auctions," *Management Information Systems Quarterly* (40:3), pp. 645-662.
- Lynham, S. 2002. "The General Method of Theory-Building Research in Applied Disciplines," *Advances in developing human resources* (4:3), pp. 221-241.
- Maaninen-Olsson, E., Wismén, M., and Carlsson, S. 2008. "Permanent and Temporary Work Practices: Knowledge Integration and the Meaning of Boundary Activities," *Knowledge Management Research & Practice* (6:4), pp. 260-273.
- Miles, M., and Huberman, A. 1994. *Qualitative Data Analysis*. Thousand Oaks, California: SAGE.
- Niederman, F., and March, S. 2015. "Reflections on Replications," *AIS Transactions on Replication* (1).
- Okoli, C., and Schabram, K. 2010. "A Guide to Conducting a Systematic Literature Review of Information Systems Research," *Sprouts Work. Pap. Inf. Syst* (10), p. 26.
- Paré, G., Tate, M., Johnstone, D., and Kitsiou, S. 2016. "Contextualizing the Twin Concepts of Systematicity and Transparency in Information Systems Literature Reviews," *European Journal of Information Systems* (25:6), pp. 1-16.
- 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.
- Rittel, H., and Webber, M. 1973. "Dilemmas in a General Theory of Planning," *Policy sciences* (4:2), pp. 155-169.
- Sein, M., Henfridsson, O., Puroo, S., Rossi, M., and Lindgren, R. 2011. "Action Design Research," *MIS quarterly* (35:1), pp. 37-56.
- Selander, L., and Jarvenpaa, S. 2016. "Digital Action Repertoires and Transforming a Social Movement Organization," *MIS Quarterly* (40:2), pp. 331-352.
- Simon, H. 1996. *The Sciences of the Artificial*. Cambridge, USA: The MIT Press.
- Sonnenberg, C., and vom Brocke, J. 2011. "Evaluation Patterns for Design Science Research Artefacts," *European Design Science Symposium*: Springer, pp. 71-83.
- Spencer, D., and Warfel, T. 2004. "Card Sorting: A Definitive Guide," *Boxes and Arrows*, p. 2.
- Viswesvaran, C., and Ones, D. S. 1995. "Theory Testing: Combining Psychometric Meta-Analysis and Structural Equations Modeling," *Personnel Psychology* (48:4), pp. 865-885.
- Zigurs, I., and Khazanchi, D. 2008. "From Profiles to Patterns: A New View of Task-Technology Fit," *Information systems management* (25:1), pp. 8-13.