Eating our own Cooking: Toward a More Rigorous Design Science of Research Methods

John Venable[1] and Richard Baskerville[2]
[1]Curtin University of Technology, Perth, Western Australia, Australia
[2]Georgia State University, Atlanta, Georgia, USA
j.venable@curtin.edu.au
baskerville@acm.org

Abstract: This paper argues that Design Science is an appropriate paradigm for research into Research Methods. Research Methods (along with their tools and techniques) are purposeful artefacts, designed and created by people to achieve a specific purpose – i.e. to create new, truthful knowledge. Like other artefacts, research methods vary in their fitness to purpose, i.e. in their utility, depending on their fit and appropriate application to the particular purpose, contexts, and contingencies for which they were developed. Design Science Research aims at developing new purposeful artefacts with evidence of their utility. Applying a DSR perspective to research methods should yield increased utility in the application of research methods, better guidance in applying them and greater confidence in achieving the desired outcomes of applying them. Based on these premises, this paper reviews the basic concerns and issues in Design Science Research (using the balanced scorecard as an example purposeful artefact), then analyses the logical consequences of taking a Design Science perspective on research methods (using the Partial Least Square approach as an example research method purposeful artefact). First, it analyses the various purposes of research methods to clarify the alternative and competing design goals of research methods. Second, it analyses and characterises the types of purposeful (design) artefacts that comprise research methods. Third, it considers issues of the evaluation of research methods. Fourth and finally, it considered the development of design theories of research methods.

Keywords: research method, research design, design science research, evaluation, design theory, research rigour

1. Introduction

There is a long history of discussion, debate, and publication regarding research methods (e.g. 1961; Campbell and Stanley 1963; Bunge 1967; Glaser and Strauss 1967; Miles and Huberman 1994). The ideas and methods used in research have developed and evolved gradually, especially as the research domains and purposes have evolved. The discipline of Information Systems is a recent participant in this debate owing to its comparatively recent origin as a field. However, the domain of information systems has presented methodological challenges that have led to further progress in research methods to better and more fully address the IS domain (e.g. Mumford, Hirschheim et al. 1985; Galliers and Land 1987; Benbasat 1991; Mason 1991).

The development and evolution of research methods has often proceeded in an ad hoc and non-rigorous process, because scientific communities are social phenomena with their own social institutions and mores (Berger and Luckmann 1966; Kuhn 1970; Latour 1987). New research methods and guidelines are posited and developed, written up and communicated to the research community, generally as research essays. But key aspects of rigour in their development may be lacking, such as weak attention to their intended goals and purposes, measures of their effectiveness, efficiency, or efficacy, evidence that the research method meets such goals or qualities, and comparisons to existing methods that provide evidence that the method better achieves its goals.

Because research methods are designed artefacts, they could be better developed using rigorous design science research (DSR) approaches. A DSR approach to the design and development of new or improved research methods would provide a more rational and consistent basis for assuring both their utility and rigour. The rigour from using a DSR approach would come primarily from (a) more precisely stated design theories about research methods and (b) more rigorous evaluation of research methods. Explicating such elements may even help make the peer reviewing of the resulting research more predictable and less capricious.

A seminal paper on DSR by March and Smith (1995) provides support for the use of Design Science Research for the development of research methodology.

Reference this paper as: Venable, J and Baskerville, R. "Eating our own Cooking: Toward a More Rigorous Design Science of Research Methods" *The Electronic Journal of Business Research Methods* Volume 10 Issue 2 2012 (pp 141-153), available online at www.ejbrm.com

> *"Natural science uses but does not produce methods. Design science creates the methodological tools that natural scientists use. Research methodologies prescribe appropriate ways to gather and analyze evidence to support (or refute) a posited theory. They are human-created artifacts that have value insofar as they address this task." (p. 258)*

While March and Smith do not assert that design science creates the methodological tools that behavioural or design scientists use, we propose that this should be the case.

While we are here focussed on DSR as explicated and used within the field of Information Systems, Venable (2010) asserted that DSR applies in all business disciplines. Among others he notes that Van Aken (2004; 2005; 2007) advocates applying DSR within the field of Management.

This paper is organised as follows. The next section reviews the relevant literature on DSR. Section three applies a DSR lens to the goals and objectives of research methods, research methods as purposeful (designed) artefacts, requirements for evaluation of research methods, and design theories in research methods. Section four summarises the paper and identifies areas for further research.

2. Key aspects of design science research for research methodology

In this section we consider relevant literature from the field of Information Systems (IS) on Design Science Research (DSR).

We will define DSR as "Research that invents a new purposeful artefact to address a generalised type of problem and evaluates its utility for solving problems of that type". There are four important concepts in this definition. First, by purposeful artefact, we mean any kind of artefact designed to achieve some human purpose. , A purposeful artefact can be a product or a process; it can be a technology, a tool, a methodology, a technique, a procedure, a combination of any of these, or any other means for achieving some purpose. Second, by invention we mean creation, design, improvement, or adaptation, not just 'pure' invention from scratch (which may not exist anyway). Third, by addressing a generalised type of problem (rather than only a single specific, situated problem), the resulting artefact (and knowledge produced about it and its utility) can be applied over and over for various occurrences of that problem type. Without this generalised outcome, the scope of the resulting knowledge and the purposeful artefact would be too narrow to be useful in other contexts. Fourth, without evaluation, one has no evidence that the produced artefact has utility for solving problems of the type identified. Rigorous evaluation provides evidence that the artefact works and the knowledge created is true and useful. Without evaluation, there is no science in Design Science Research.

As an example of DSR, consider the development of the now commonplace balanced scorecard (Kaplan and Norton 1992; Kaplan and Norton 1996). While the creation of the balanced scorecard is in the DSR paradigm, Kaplan and Norton undoubtedly did not consider it to be.

The balanced scorecard approach was intended to provide a more holistic, comprehensive, yet still not overly complex picture of firm performance that incorporated not only financial, but also operational performance, to enable better firm management in the long run. The resulting artefact, the balanced scorecard, gives the manager "a fast, but comprehensive view of the business" (Kaplan and Norton 1992, p. 71).

The remainder of this section considers the nature of purposeful (design) artefacts, design theories as formal statements of the knowledge created by DSR, and evaluation in DSR.

2.1 Purposeful (design) artefacts

DSR produces one or more purposeful artefacts. March and Smith (1995) were the first in the IS field to establish the terminology of Design Science. They identified four kinds of what they call 'design artefacts' as the outputs of DSR: constructs, models, methods, and instantiations. Constructs are concepts that are used to describe a problem and specify its solution. Models are defined by March and Smith (1995) as "a set of propositions or statements expressing relationships among constructs" (p. 256) that can be used to "represent situations as problem and solution statements" (p. 256). They define methods as "a set of steps (an algorithm or guideline) used to perform a task" (p. 257). Finally, they define instantiations as “the realization of an artifact in its environment” (p. 258). Interestingly, the above four artefacts do not include “design” per se. However, the model artefact could be used for this, in that it includes “solution statements” (p. 256). A design can then be conceived of as different constructs (the design components) and statements about relationships between them.

For example, consider the balanced scorecard (Kaplan and Norton 1992; Kaplan and Norton 1996) further. Considering March and Smith’s (1995) four kinds of artefacts, there are a number of key constructs or concepts, including the four proposed areas/perspectives on performance measures (financial, customer, internal business, and innovation and learning). Other constructs include the goals and measures to be identified within each perspective. The main model in the balanced scorecard is the view that integrates the four perspectives and shows the links between each of the four perspectives, as well as the goals and their measures within each perspective. Of course, another key concept is that it is up to the firm management to come up with the goals and measures that are appropriate for their organisation. In their original paper, Kaplan and Norton (1992) did not provide a method for how to apply the balanced scorecard, but it is implicit. They did describe an example part of an instantiation of their approach, as carried out at ECI.

Gregor and Jones (2007) distinguish two different kinds of purposeful artefacts that can be designed: product artefacts and process artefacts. A product artefact is a thing, such as a tool, an object, or a system. It can be either physical (e.g. a computer) or abstract (e.g. a diagram notation). A process artefact is similar to a method (March and Smith 1995). Of course, a purposeful artefact may be designed to include both process and product.

The balanced scorecard artefact contains both product and process aspects. The product is a scorecard itself and the constructs/concepts behind it. The process is the method by which the relevant measures of the four perspectives are decided and implemented.

Venable et al (2012) also distinguish between purely technical artefacts (which are admittedly fairly rare, such as a computer network protocol, which has no human users, and socio-technical artefacts, which have human users, possibly with complex social implications from their adoption and use. Importantly, socio-technical artefacts present a much richer milieu concerning the values and interests of the stakeholders as well as complications for the evaluation of the artefact (Venable, Pries-Heje et al. 2012).

The balanced scorecard is clearly a socio-technical artefact, since it has both individual usability aspects as well as group and social aspects in its use and the process of adapting it for a particular organisation.

2.2 Design theory

Design Theories are formalisations of the knowledge that results from Design Science Research. It is generally considered that all high quality research should generate, evaluate, or refine theory.

Walls et al (1992) argued that a proper design theory would have seven components as shown in figure 1: meta-requirements (a general set of requirements that a generalised solution would address), meta-design (a generalised design that could be adapted to a particular problematic situation), a design method (for adapting the meta-design to the particular problematic situation), kernel theories informing the meta-design (according to the meta requirements), kernel theories informing the design method (according to the meta-requirements and meta-design), testable hypotheses to test the meta-design and testable hypotheses to test the design method. They also state that design theories are prescriptive in guiding designers as to what they should do.

Gregor and Jones (2007) developed an alternative, but similar form of design theory with eight parts: Purpose and scope, Constructs, Principles of form and function, Artefact mutability, Testable propositions, Justificatory knowledge, Principles of implementation, and an expository instantiation (optional).

Other authors have proposed simpler structures for design theories, focussing on meta-requirements and meta-design. Venable (2006) asserted that kernel theories and testable hypotheses are not part of a design theory per se and that a design method is not required to form a design theory. He advocated a simpler, non-prescriptive form of theory (which he termed a “utility theory”) containing only (1) the problem space to which a design theory applies (similar to meta-requirements), (2) a solution space describing a design (similar to meta-design), and (3) a relationship between them asserting that the design has utility of some kind (efficiency, effectiveness, etc.) with respect to the problem space. Relationships between components of the problem space and between components of the solution space should also be part of the design/utility theory. Similarly, Baskerville and Pries-Heje (2010) proposed explanatory design theory, which includes three similar parts: (1) General Requirements (similar to meta-requirements), (2) General Components (similar to meta-design), and (3) a relationship between the two which means that the general components together satisfy the general requirements. In essence, the theory explains how to meet the general requirements, i.e. by applying the general components.

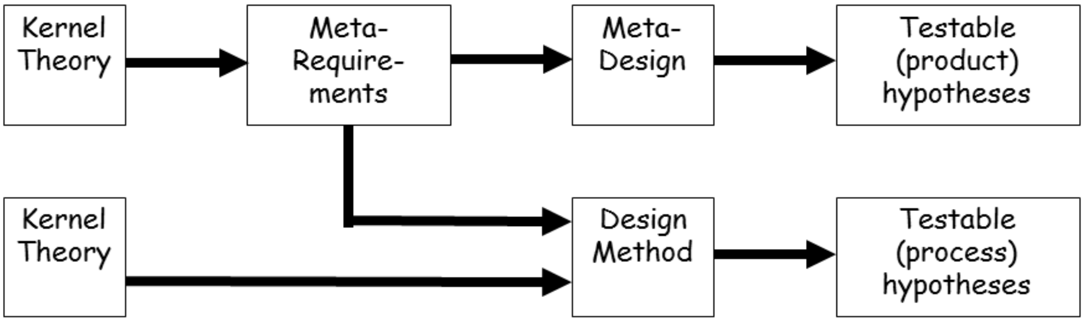

Figure 1: Structure of a design theory (Walls et al, 1992)

Since Kaplan and Norton (Kaplan and Norton 1992; Kaplan and Norton 1996) did not explicitly follow a DSR approach in their development of the balanced scorecard and the ideas of design theories had not yet been invented, it is not surprising that they did not explicitly propose a design theory. However, we can reconstruct at least part of a design theory post hoc, using the structure of Walls et al (1992).

Table 1 shows the different components of design theories according to the different works reviewed above in the left column and in the right column shows what each component of a design theory for the balanced scorecard might look like if Kaplan and Norton had developed one. Similar design theory components from the descriptions above are shown in the same row, The notations “WWES”, “G&J”, “V”, and “B&PH” identify the paper authors who developed the identified components.

Considering Table 1, the requirements shown in the first row are ones that need to be filled in many organisations; hence they are termed meta-requirements or generalised requirements. The particular goals and measures are left non-specific and have to be decided my managers for each particular organisation and their current situation, hence they are part of a meta-design or generalised design, rather than a specific design of a balanced scorecard for a particular organisation. The design method is left implicit in Kaplan and Norton (1992), but would likely require the steps shown. Kaplan and Norton (1992) somewhat notoriously did not provide references to any literature on prior art or to the requirements that they were trying to achieve. Therefore they did not provide any kernel theory. However, we can infer relevant literature which might (better) inform the meta-design and design method. Finally, we surmise some example testable hypotheses or propositions, but many others are possible.

Table 1: Hypothetical design theory for the balanced scorecard

Design Theory Component (Source)	Hypothesised Design Theory for Balanced Scorecard (Kaplan and Norton 1992)
Meta-requirements (WWES) Purpose and scope (G&J) Problem Space (V) General requirements (B&PH)	Enhance the ability of managers to manage a firm in the long run Enhance the ability of managers to manage both financial and operational aspects of a firm Reduce the complexity of managing a firm Develop an integrated and holistic view of the firm Support adapting the approach to specific organisations
Meta-design (WWES) Constructs, Principles of form and function,	Four perspectives (financial, customer, internal business, and innovation and learning) The links between the four perspectives (only the financial and innovation and

Design Theory Component (Source)	Hypothesised Design Theory for Balanced Scorecard (Kaplan and Norton 1992)
Artefact mutability, and Principles of Implementation (G&J) Solution space (V) General components (B&PH)	learning perspectives are not linked to each other) Goals should be included within each perspective One or more measures for each goal Guideline that there should be no more than 20 measures overall to ensure sufficient simplicity and ability to gain oversight of the entire firm
Design Method (WWES)	Decide the goals for each perspective Decide the measures for each goal Decide how to obtain the measures Decide how to present the measures to managers Build or buy and implement an activity system to provide the information Use the measures to manage the firm
Meta-Design Kernel Theories (WWES) or Justificatory Knowledge (G&J)	Could be drawn from extant literature … on why managing is difficult informing the design of the four different perspectives about extant goals and measures that have been identified and used in the different perspectives
Design Method Kernel Theories (WWES) or Justificatory Knowledge (G&J)	Could be drawn from extant literature on … how to collaborate on decision making about management how to present complex information simply
Meta-Design Testable Product Hypotheses (WWES) and Propositions (G&J)	Possible testable hypotheses or propositions re. the meta-design might be … “Use of the balanced scorecard reduces complexity in managing a firm.” “Use of the balanced scorecard improves long term firm management.”
Design Method Testable Process Hypotheses (WWES) and Propositions (G&J)	Possible testable hypotheses or propositions re. the design method might be … “The design method enables effective implementation of balanced scorecards.”
Instantiation (G&J)	A partial example instantiation was provided based on the ECI case study.

2.3 Evaluation

Evaluation is a key activity in Design Science Research (Nunamaker et al. 1991; March and Smith 1995; Hevner et al. 2004; Vaishnavi and Kuechler 2004; Venable 2006; 2006; Peffers et al. 2008). Evaluation is the activity that makes DSR a science. Without it, asserted design theories, utility theories, or technological rules are just untested theories, conjectures, and hypotheses; there is no evidence to back them up.

Venable et al (2012) identify five different purposes for evaluation in DSR:

- Evaluate an instantiation of a purposeful (designed) artefact to establish its utility (including efficacy, effectiveness, and efficiency – or lack thereof) for achieving its stated purpose
- Evaluate the formalized knowledge about a purposeful artefact’s utility for achieving its purpose
- Evaluate a purposeful artefact or formalized knowledge about it in comparison to other extant purposeful artefacts’ utility to achieve a similar purpose
- Evaluate a purposeful artefact or formalized knowledge about it for side effects or undesirable consequences of its use
- Evaluate a purposeful artefact formatively to identify weaknesses and areas of improvement while still under development

There are four questions we should answer in designing an evaluation:

- What is the evaluand (the artefact being evaluated)?
- What aspect or quality of the evaluand is to be evaluated? Artefacts are evaluated for their fitness to purpose. Purposes are extremely varied, but one can identify some key characteristics, such as Checkland’s five E’s : efficiency, effectiveness, efficacy, (a)esthetics, and ethicality (Checkland and Scholes 1990), which are aspects of utility.

- What kind and how much rigour are required of the evaluation? Venable et al (2012) identify two senses of rigour. The first sense is whether an improvement or achievement of purpose can be attributed to the evaluand (i.e. its efficacy), or alternatively whether it could be caused by some confounding variable. The second sense is whether the evaluation can work within the myriad complicating aspects of real use, i.e. its effectiveness.
- What evaluation method(s) should be used? Hevner et al (2004) suggest analytical, case study, experimental, field study, and simulation. Venable (2006) further suggests action research and distinguishes between artificial evaluation and naturalistic evaluation. But which evaluation method(s) should be used?

Venable et al (2012) provide extensive advice on evaluation method choice and design based on a framework matching the prioritised goals of the evaluation to strategies and then to specific relevant methods. Strategies consider whether evaluation should be ex ante (before instantiation) or ex post (after instantiation) and whether it should be naturalistic (real users using the real purposeful artefact on a real task (Sun and Kantor 2006)) or artificial (lacking one of the three realities of Sun and Kantor (2006)) or a combination of these.

Considering evaluation of the balanced scorecard, the evaluand (purposeful artefact evaluated) could be the balanced scorecard itself, some more focussed component part of it, such as different ways to represent the balanced scorecard, and/or the method for designing and using a balanced scorecard in a particular organisation.

Kaplan and Norton (1992) provided a fairly minimal evaluation of the overall approach by applying the method in several different case study organisations. Because this was uncontrolled, it provided only subjective, anecdotal evidence of the effectiveness of the purposeful artefact, not of its efficacy. Anecdotal evidence was also provided of the efficiency of the method. The method was also not compared to other methods.

Other researchers have also evaluated the balanced scorecard approach, e.g. Mooraj et al (1999) and Malina and Selto (2001). These evaluations were more rigorously performed, but still used case studies and largely considered subjective opinions (e.g. the perceived attributes of the balanced scorecard for strategy formulation and communication (Mooraj et al. 1999)). Both evaluations were generally positive, but could not be said to rigorously evaluate the efficacy of the balanced scorecard.

Unfortunately, constructing a fully-controlled evaluation of the efficacy of the balanced scorecard is probably impossible; the key anticipated benefit (dependent variable) – improved firm performance in the long term – is subject to many confounding variables.

Other desired benefits, such as improved ability to manage, improved ability to deal with both financial and operational performance in an integrated and holistic way, and reduced complexity, are all subjective in nature and could be evaluated by asking people about their subjective opinions on whether, how well, or how much better the purposes are achieved.

3. Analysis of research methods using a design science research lens

This section considers what lessons may be learned from applying Design Science Research to research methods and how DSR might improve the rigour of research methods. We look at four main areas: goals (and constraints) of research methods, design artefacts in research methods, the need for evaluation of research methods, and design theory in research methods.

As an example for the purpose of this paper, we will use the Partial Least Squares (PLS) method (Wold 1974; Wold 1982; Lohmöller 1989; Chin 1998; Chin and Newsted 1999) for analysing quantitative data in the context of Structured Equation Modelling (SEM or PLS-SEM). One reason for this is that the use of PLS for business research is a matter of extensive debate (Marcoulides and Saunders 2006; Marcoulides et al. 2009; Gefen et al. 2011; Hair et al. 2011; Ringle et al. 2012; Rönkkö et al. 2012).

3.1 Goals and purposes of research methods

This section identifies and discusses some of the various goals to be met by research methods. Goals include both the primary goals as well as practical goals that enable their effective use. We discuss those first, then implications from a DSR perspective.

3.1.1 Primary goals

The primary goals of research methods include both general goals and goals specific to the research project. Rigour and relevance are general goals desired of all research and research methods. Research methods must also be appropriate for answering the specific research question or achieving the specific goal of the research.

Rigour is concerned with the reliability of research. Properly following a research method should ensure that the research findings are correct or that the probability of the findings being incorrect is sufficiently low.

Relevance is concerned with whether the findings are useful – either to theory and further research or to practice. While following a research method may not ensure relevance, some research methods may be more amenable to investigating topics that are relevant.

Research questions may be of several kinds or types. Some research questions are concerned with theory testing (and extension). Some are concerned with theory building. Some are concerned with understanding and explanation while others are concerned with prediction (cf. Gregor 2006). Some are concerned with evaluation. Some are concerned with inventing new technologies. Some are concerned with human emancipation. Different research methods are more or less appropriate for answering different kinds of research questions.

A primary goal of PLS-SEM is to support rigorous analysis of quantitative data. The relevance of its results are mostly dependent on the relevance of the questions and the data collected, but, as a positivist, reductionist approach dealing with quantitative data, there are many areas where more interpretive approaches may bear more fruit. In general, it is recommended that PLS-SEM be applied for exploratory research (Gefen et al. 2011) and theory building and prediction rather than theory testing. PLS is most relevant and appropriate for social and behavioural research, including business research (Gefen et al. 2011; Hair et al. 2011).

PLS-SEM (along with other SEM forms, i.e. co-variance-based SEM or CB-SEM) also has a key advantage in its support for latent variables (or concepts), which cannot be directly measured.

3.1.2 Practical goals

There are many practical goals for research methods that affect their ability to be applied and to achieve the primary research goals given above. Among them are ease of learning, ease of use, cost, time required, facilities or other resources required, access to research subjects, ability to deal with more people and research subjects, and depth of engagement with research subjects.
These goals are often in the form of constraints – constraints on available researcher time, expertise, funding, access to research subjects, and others. The list is probably not comprehensive.

A key practical advantage for PLS is its (supposed) ability to work with small(er) sample sizes. However, this is under much debate and widely criticised (Marcoulides and Saunders 2006; Marcoulides et al. 2009; Gefen et al. 2011; Ringle et al. 2012; Rönkkö et al. 2012). Smaller sample size requirements is often a reason for using PLS, with over 1/3 of papers using PLS published in MIS Quarterly citing this as a reason for choosing PLS (Ringle et al. 2012).

PLS, particularly with the software support available today (e.g. SmartPLS), is quite easy to learn and use, so in some sense is suited to novice researchers with limited time resources. However, it can also be argued that the constraints and limitations of the technique, as well as reporting standards (Gefen et al. 2011; Ringle et al. 2012), are quite difficult to learn as we shall see below.

3.1.3 Lessons from DSR

The goals above are analogous to the DSR idea of meta-requirements. The different research methods are analogous to the DSR idea of meta-designs. The different goals/meta-requirements form contingencies that are addressed with differing levels of utility by the different research methods.

A key issue is that the number of goals and constraints is fairly large and to some extent they conflict. Evaluating the different contingencies is complex and difficult as it is a combinatorial problem without hard yes/no answers. However, the provision of a clear system of design theories with research goals as the contingencies could aid in the selection of an appropriate research method for a particular research situation. See for example Venable et al (2012) for a framework and method to aid researchers in the selection and design of evaluation methods for DSR purposeful artefacts.

Moreover, there may be room to develop a system or tool to aid in the selection of research methods based on the specific goals and priorities of a research project. For example, the approach described in Baskerville and Pries-Heje (2008) for helping managers to select an appropriate change management methodology might also be applied to a system to help selection of research methods. Given that research questions sometimes evolve based on the feasibility of researching them, such a tool might also be modified to support choosing/designing research questions.

3.2 Design artefacts in research methods

As described in section 2.1, design artefacts are distinguished as product and process artefacts (Gregor and Jones 2007), as well as classified as constructs, models (including designs), methods, and instantiations (March and Smith 1995; Hevner et al. 2004).

What kind of design artefact is a research method? A research method is clearly a process or method artefact. Research methods often (always?) make use of various tools, techniques, and frameworks for organising and doing the research work. These are clearly product artefacts. Descriptions of how to use a tool, technique, or framework are in turn process artefacts.

Importantly, tools and techniques may sometimes be used in other methods than where they were developed. They have an existence of their own and could be used in other ways. They have their own goals and instructions for how to use and adapt them. Any attempt to conduct DSR on research methods needs to consider the possibility that research tools and techniques have uses both as part of and outside of any particular research method.

Considering PLS, there are many artefacts which could be evaluated. PLS is an algorithm (process), implemented in software (product), with a user interface (product), and also procedures (process) and standards (product) for its application and submission of results for publication. One could evaluate the PLS algorithms, the correctness and usability of PLS software, how easy it is to learn and apply the standards or practice, whether PLS achieves the rigour that is required when applied with its constraints and standards (its efficacy), the efficiency of its application by people (time, resources used, etc.), and the effectiveness of its use (obtaining rigour in actual practice by researchers).

3.3 Evaluation of research methods

In the introduction, we asserted that research methods have evolved in a non-rigorous way, without the benefit of the ideas of Design Science Research. A key aspect of DSR is evaluation of new technologies.

3.3.1 Why evaluate research methods?

Research methods are tremendously important to the progress of human knowledge and the improvement of the human condition. Currently, we lack evidence of the effectiveness, efficiency, and efficacy of research methods. Some evidence clearly is provided by the evident progress of human knowledge, but enabled by which research methods? Other evidence has been developed through rational argument in research essays. However, without more rigorous evidence, how do we really know that research methods are (sufficiently) effective, efficacious, or efficient?

We also lack evidence about whether research methods achieve other desirable, practical goals – e.g. ease of learning, ease of use, and time and resource requirements.

Moreover, we especially lack evidence about the relative efficiency and effectiveness of research methods, i.e. compared to each other.

In order to make better judgments in the choice of research methods, it would be very helpful to have better evidence than our currently ad hoc assertions about them.

For example, Rönkkö et al (2012) recently evaluated the use of PLS for theory testing who found that many models developed using PLS and published in top IS journals may in fact by incorrect. Of course the fault may lie in the application of PLS rather than the PLS algorithms or their software implementations. Importantly, their research shows that, whether the PLS technique would work (determine correct models) if applied correctly, the PLS technique as recently applied in the IS field, is (or has been) *ineffective* in practice, as used and published by IS researchers, reviewers, and journal editors.

3.3.2 How could we evaluate research methods?

Like other designed artefacts, research methods should be evaluated for their utility in achieving their goals, including primary goals of rigour and relevance, suitability to type of research questions or research domain and topic, as well as secondary, practical goals.

Rigour as it is assessed in the reviewing process is a rather subjective and often assessed by simply confirming that an established research method has been followed, assuming that the outcome will therefore be correct. Sometimes, sets of criteria are set out for what makes research rigorous, trying to make the criteria as specific and clear as possible, even objective if possible. But what if the method itself or the criteria are wrong?

The most convincing evidence of the rigour of a research method would come from the detection of errors in findings and determining error rates. Where errors are found in findings, the research method itself is in some way responsible (after all, it didn’t prevent the error), but additional causes are possible (e.g. misapplication of the method). Finding errors would require triangulation of findings through other research, which may not be possible or thought to be worth the research expenditure.

For some research methods, especially quantitative ones, the efficacy of the method to provide rigour can be demonstrated through mathematical analysis or proof.

Alternatively, research conducted could be subjected to critical analysis and ratings of rigour. The rigour ratings could then be compared and correlated across projects using different research methods. Unfortunately, mixed- and multi-method approaches would complicate these analyses. It may also be helpful analyse failed research projects, where outcomes are not accepted for publication. One might also consider whether a research method correlates with high or low quality journals.

To assess the performance/utility of research methods in achieving goals in other, practical areas one could survey users of various research methods about their perceptions of ease of learning and ease of use of research methods, tools, and techniques. One could evaluate resource consumption by gathering data on time required, resources consumed, costs, etc. and correlating them with research methods.

The above suggestions relate to extant research methods, but what about the development of new methods where there isn’t a history of application? Where a method is new, evaluation is largely limited to analytical kinds of evaluation, e.g. rational argument, surveys or focus groups with experts in research methodology, or possibly some form of simulation. Empirical, naturalistic evaluation will likely be very limited at first, with limited trials, probably using qualitative evaluation methods.

Considering evaluation of PLS, Rönkkö et al (2012) did not evaluate the process by which PLS was applied; instead it examined the results of the application of PLS and publication of PLS results. It did so be re-running the analyses, but with modified data, including (1) switching the data between two variables and (2) substituting completely random data for two variables. They found that in a significant portion of the models re-analysed with incorrect data, the model was still supported by the analysis. Since models with incorrect data were supported, presumably the models themselves may be incorrect since the evidence through the use of PLS is extremely weak.

3.4 Design theory for research methods

Design theories would be important in a Design Science of Research Methods.

3.4.1 Why design theories of research methods?

Venable (2006) observed that theory is the way that academics communicate with each other about research. Clear, formalised statements of theory are needed to reduce misinterpretation and to allow other researchers to accurately research the same topics. They are also important when developing new research methods as one needs to be able to clearly articulate how a new research method is different from existing ones.

Clear statements of design theories may also be useful for practitioners – in this case the practitioners are researchers.

For these reasons, we propose that formal design theories could and should be developed for every research method. Doing so would aid the on-going development and progress in research methods and in the long run should increase our ability to generate new, valid knowledge.

3.4.2 What should be in a design theory for a research method?

At a minimum, a design theory for a research method needs three things: (1) meta-requirements (Walls et al. 1992), a model of the problem space (Venable 2006) to be addressed by the research method, a purpose and scope (Gregor and Jones 2007), or generalised requirements (Baskerville and Pries-Heje 2010), (2) a meta-design (Walls et al. 1992), a model of the solution space (Venable 2006), principles of form and function (Gregor and Jones 2007), or a generalised design (Baskerville and Pries-Heje 2010), and (3) a prescriptive statement (Walls et al. 1992; Gregor and Jones 2007), an asserted utility relationship (Venable 2006), or an explanatory relationship (Baskerville and Pries-Heje 2010) between the first two. Other aspects of design theories are optional.

All aspects of design theories need to be stated as precise, well-defined constructs. The meta-requirements/model of the problem space/purpose and scope should include and clearly state the goals (and possibly the constraints) for which the research method is suited (or possibly the level to which they are suited, such as low cost). The primary goals that should be addressed include research rigour and relevance and the level of each that can be achieved or expected, the type(s) of research question for which the method is suited, and the domain(s) for which the research method is suited. Optional goals may also be addressed (we recommend this) by statements about the anticipated levels of ease of learning, ease of use, cost, time required, facilities or other resources required, access to research subjects, ability to deal with more people and research subjects, and depth of engagement with research subjects.

The meta-design/model of the solution space/principles of form and function should include a detailed description of the research method's process, including steps, actions to take, tasks, decisions, and iterations, as well as tools and techniques to use, and when and how to use them. Descriptions of tools and techniques may refer to other design theories for detail about them.

Table 2 gives an example of a design theory for the PLS approach to quantitative data analysis. In this case, we take on the results of the Rönkkö et al (2012) evaluation and propose some improvements to PLS as it is currently implemented.

Table 2: Hypothetical (partial) design theory for the partial least squares method

Design Theory Component (Source)	Hypothesised Design Theory (partial) for an improved Partial Least Squares method
Meta-requirements (WWES) Purpose and scope (G&J) Problem Space (V) General requirements (B&PH)	Support use of latent variables and both formative and reflective measures. Support use of small data sets (within minimum constraints). Support ease of learning and application by novice researchers. Support easy building and testing of models (efficiency and ease of use). Support easy calculation and application of constraints to ensure adequate rigour (i.e. correctness) of results.

Design Theory Component (Source)	Hypothesised Design Theory (partial) for an improved Partial Least Squares method
Meta-design (WWES) Constructs, Principles of form and function, Artefact mutability, and Principles of Implementation (G&J) Solution space (V) General components (B&PH)	PLS algorithms Software implementation of the algorithms Guidelines and constraints for PLS results testing and result publication Easy software management of data sets and results of analyses Automatic support for determining sufficient sample size for the model to be tested Automatic testing of constraints and alerting to constraint violations Automatic testing of a saturated model and comparison to the proposed model Automatic re-testing of models with partially switched or partially random data to ensure model correctness and strength of evidence
Design Method (WWES)	Determine the model and measures to be used. Input the model into the software and calculate the required sample data size – revise model and recalculate as needed. Collect and input the data. Run the model Read alerts as to whether all the constraints and supporting calculations (e.g. f^2, Q^2, and q^2 (Ringle, Sarstedt et al. 2012)) meet acceptable conditions. Read automated explanations as to why any unmet constraints are a problem and suggestions for how to correct the problem.
Meta-Design Kernel Theories (WWES) or Justificatory Knowledge (G&J)	Could be drawn from extant literature, e.g. … on standards and constraints for PLS (e.g. as discussed in this paper) on usability of software implementation of complex models
Design Method Kernel Theories (WWES) or Justificatory Knowledge (G&J)	Could be drawn from extant literature, e.g. … on teaching research design and conducting research on research risks to be managed
Meta-Design Testable Product Hypotheses (WWES) and Propositions (G&J)	Possible testable hypotheses or propositions re. the meta-design might be … "The improved PLS approach and software are easy to learn and use." "Use of the improved PLS approach ensures rigour of resulting models"
Design Method Testable Process Hypotheses (WWES) and Propositions (G&J)	Possible testable hypotheses or propositions re. the design method might be … "The design method enables effective enactment of the improved PLS approach.
Instantiation (G&J)	None yet, this is only a proposal!

4. Conclusion

Research Methods are designed artefacts. Currently, there is no adequate basis for assuring the utility and rigour of research methods. The evolution and development of methods needs a more rational basis. This paper proposes that DSR could (and should) provide such a basis. If we are to be certain that research methods are reliable and achieve their purpose, they should be developed with a rigorous DSR approach. The rigour would come from (a) more precise statements of design theories of research methods and (b) more rigorous evaluation of research methods.

In order to progress toward a rigorous DSR approach to research methods, we have applied DSR concepts to research methods, tools, and techniques. We have identified the primary and secondary (practical) goals of research methods, which would be part of design theories about research methods and serve as criteria for evaluation of the utility of research methods. We have also used DSR concepts to identify types of design artefacts in research methods, which would define the research method in a design theory. The configuration of constructs and artefacts that make up a research method are what are evaluated for how well they fulfil the research method's goals. This understanding should also lead to more consistent evaluation of research methods.

Applying a DSR approach to research about research methods should enhance the quality of research methods and therefore the quality of the research conducted using them.

What we have proposed in this paper is only a first step. Work is now needed to analyse existing research methods to develop clear design theories about them. The basis for this already exists in the literature; it simply needs critical reading and formalising in design theories. Consideration also is needed re designing and carrying out appropriate evaluations of research methods, to provide evidence of their effectiveness, efficiency, efficacy, etc. in achieving their disparate purposes. Once a body of literature on design theories and evaluations of research methods is developed, there would be a clearer basis for further research on research methods.

References

Baskerville, R. and J. Pries-Heje (2010). "Explanatory Design Theory." Business & Information Systems Engineering **2010**(5): 271-282.

Benbasat, I., Ed. (1991). The Information Systems Research Challenge: Experimental Research Methods. Boston, Harvard University Press.

Berger, P. and T. Luckmann (1966). The Social Construction of Reality: A Treatise in the Sociology of Knowledge. New York, Anchor Press.

Bunge, M. (1967). Scientific Research I: The Search for System. New York, Springer-Verlag.

Campbell, D. and J. Stanley (1963). Experimental and Quasi-Experimental Designs for Research. Boston, Mass., Houghton Mifflin.

Checkland, P. and J. Scholes (1990). Soft Systems Methodology in Action. Chichester, J. Wiley.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. Modern Methods for Business Research. G. A. Marcoulides. Mahwah, NJ, USA, Lawrence Erlbaum Associates**:** 295-336.

Chin, W. W. and P. R. Newsted (1999). Stuctural equation modelling analysis with small samples using partial least squares. Statistical Strategies for Small Sample Research. R. H. Hoyle. Thousand Oaks, CA, USA, Sage Publications**:** 307-342.

Galliers, R. and F. Land (1987). "Choosing appropriate information systems research methodologies." Communications of The ACM **30**(11): 900-902.

Gefen, D., E. E. Rigdon, et al. (2011). "An Update and Extension to SEM Guidelines for Administrative and Social Science Research." MIS Quarterly **35**(2): iii-xiv.

Glaser, B. G. and A. L. Strauss (1967). The Discovery of Grounded Theory, Strategies for Qualitative Research. Chicago, Aldine Publishers.

Gregor, S. (2006). "The Nature of Theory in Information Systems." MIS Quarterly **30**(3): 611-642.

Gregor, S. and D. Jones (2007). "The Anatomy of a Design Theory." Journal of the Association for Information Systems **8**(5): 312-335.

Hair, J. F., C. M. Ringle, et al. (2011). "PLS-SEM: Indeed a Silver Bullet." Journal of Marketing Theory and Practice **19**(2): 139–151.

Hevner, A. R., S. T. March, et al. (2004). "Design Science in Information Systems Research." MIS Quarterly **28**(1): 75-105.

Kaplan, R. S. and D. P. Norton (1992). "The Balanced Scorecard: measures that drive performance." Harvard Business Review(January-February): 71-80.

Kaplan, R. S. and D. P. Norton (1996). Balanced Scorecard: Translating Strategy into Action. Cambridge, MA, USA, Harvard Business School Press.

Kuhn, T. S. (1970). The Structure of Scientific Revolutions. Chicago, University of Chicago Press.

Latour, B. (1987). Science in Action. Boston, Harvard.

Lohmöller, J.-B. (1989). Latent Variable Path Modeling with Partial Least Squares. Heidelberg, Germany, Physica.

Malina, M. A. and F. H. Selto (2001). "Communicating and controlling strategy: An empirical study of the effectiveness of the balanced scorecard " Journal of Management Accounting Research (2001): **13**: 47-90.

March, S. T. and G. F. Smith (1995). "Design and natural science research on information technology." Decision Support Systems **15**(4): 251-266.

Marcoulides, G. A., W. W. Chin, et al. (2009). "A Critical Look at Partial Least Squares Modeling." MIS Quarterly **33**(1): 171-175.

Marcoulides, G. A. and C. Saunders (2006). "PLS: A Silver Bullet?" MIS Quarterly **30**(2): iii-ix.

Mason, R. (1991). MIS experiments: A pragmatic perspective. The Information Systems Research Challenge: Experimental Research Methods. I. Benbasat. Boston, Harvard University Press**:** 3-20.

Miles, M. B. and A. M. Huberman (1994). Qualitative Data Analysis: A Sourcebook of New Methods. Newbury Park, Calif, Sage.

Mooraj, S., D. Oyon, et al. (1999). "The Balanced Scorecard: A Necessary Good or an Unnecessary Evil?" European Management Journal **17**(5): 481-491.

Mumford, E., R. Hirschheim, et al., Eds. (1985). Research Methods in Information Systems. Amsterdam, Elsevier Science Publishers.

Nagel, E. (1961). The Structure of Science: Problems in Scientific Explanation. London, Routledge and Kegan Paul.

Nunamaker, J. F. J., M. Chen, et al. (1991). "Systems Development in Information Systems Research." Journal of Management Information Systems **7**(3): 89-106.

Peffers, K., T. Tuunanen, et al. (2008). "A Design Science Research Methodology for Information Systems Research." Journal of Management Information Systems **24**(3): 45-77.

Ringle, C. M., M. Sarstedt, et al. (2012). "A Critical Look at the Use of PLS-SEM in MIS Quarterly." MIS Quarterly **36**(1): iii-xiv.

Rönkkö, M., K. Parkkila, et al. (2012). "Use of Partial Least Squares as a Theory Testing Tool - An Analysis of Information Systems Papers." Proceedings of the 2012 European Conference on Information Systems (ECIS 2012): n.p.

Sun, Y. and P. B. Kantor (2006). "Cross-Evaluation: A new model for information system evaluation." Journal of the American Society for Information Science and Technology **57**(5): 614-628.

Vaishnavi, V. and W. Kuechler. (2004). "Design Research in Information Systems." Retrieved October 2, 2004, from http://www.isworld.org/Researchdesign/drisISworld.htm.

van 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.

van Aken, J. E. (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, J. E. (2007). "Design Science and Organization Development Interventions: Aligning Business and Humanistic Values." THE JOURNAL OF APPLIED BEHAVIORAL SCIENCE **43**(1): 1-17.

Venable, J. (2006). A Framework for Design Science Research Activities. Proceedings of the 2006 Information Resource Management Association Conference Washington, DC, USA.

Venable, J., J. Pries-Heje, et al. (2012). "A Comprehensive Framework for Evaluation in Design Science Research." Proceedings of the 7th International Conference on Design Science Research in Information Systems and Technology (DESRIST 2012).

Venable, J. R. (2006). "The Role of Theory and Theorising in Design Science Research." Proceedings of the First International Conference on Design Science Research in Information Systems and Technology, Claremont, CA, USA, 24-25 February, 2006: 1-18.

Venable, J. R. (2010). Information Systems Design Science Research as a Reference Discipline for Other Business Disciplines. International Academy of Business and Public Administration Disciplines Conference New Orleans, International Academy of Business and Public Administration Disciplines**:** 1049-1061.

Walls, J. G., G. R. Widmeyer, et al. (1992). "Building an information system design theory for vigilant EIS." Information Systems Research **3**(1): 36-59.

Wold, H. (1974). "Causal Flows with Latent Variables: Partings of Ways in the Light of NIPALS Modelling." European Economic Review **5**(1): 67-86.

Wold, H. (1982). Soft Modeling: The Basic Design and Some Extensions. Systems Under Indirect Observations: Part II. K. G. Jöreskog and H. Wold. Amsterdam, Netherlands, North-Holland.

