ACTION RESEARCH AND DESIGN SCIENCE RESEARCH –MORE SIMILAR THAN DISSIMILAR

Petter Øglænd, Department of Informatics, University of Oslo, P.O. Box 1080 Blindern, 0316 Oslo, Norway, petterog@ifi.uio.no

Abstract

During the past few years there has been a discussion within the information systems community related to the similarities and differences between action research (AR) and design science research (DSR). Some believe the two research paradigms are similar and that they can be matched in ways that can improve the quality of research. Others argue that AR and DSR are based on fundamentally different research paradigms, making a mixed approach difficult. From the viewpoint of doing intervention based software process improvement (SPI) research, the combination of AR and DSR can typically be a critical success factor. Using van Fraassen’s philosophy of science as a starting point, this paper argues how paradigmatic differences between AR and DSR can be overcome by designing AR/DSR based SPI research along the philosophy of constructive empiricism. The idea is illustrated by an empirical example of the evolution of an AR/DSR based SPI research method. The significance of the paper is a support to the group of IS researchers arguing the case of similarity between AR and DSR by pointing out a way of overcoming the paradigmatic criticism made by researchers arguing the case of dissimilarity.

Keywords: Action research, design science research, software process improvement, philosophy of science

1 INTRODUCTION

There is an ongoing debate within the information systems (IS) community regarding the relationship between Action Research (AR) and Design Science Research (DSR) (e.g. Burstein & Gregor, 1999; Cole et al., 2005; Järvinen, 2007; Ilvari & Venable, 2009). Much of the challenge in the debate relates to the fact that AR and DSR are classes of research methods rather than clearly defined methods, although they represent different mindsets in terms of AR originally being a research method in applied psychology (Masters, 1995) and DSR being engineering science (Simon, 1996).

Viewing AR as a problematic research methodology for validating scientific claims, Järvinen (2007) compared it with DSR, finding the two methodologies similar along a string of important parameters and suggested DSR could be used for improving AR designs. Ivari and Venable (2009) responded to this by comparing AR and DSR along other parameters, claiming that they are significantly different. The purpose of this paper is to continue the debate in support of Järvinen by viewing AR and DSR from the model-oriented philosophy of science suggested by Suppes (1967), van Fraassen (1980), Cartwright (1983) and others.

The basic idea in van Fraassen’s “constructive empiricism” is that science should be understood as the activity of making models that fit with empirical data and are useful for making predictions, and not an activity that has anything to do with making truth claims about reality. Although his main area of study is natural science, the reasoning works equally well for social sciences like economics that follow a similar mathematical path (Cartwright, 1983). The idea in this paper is to point towards the similarities in AR and DSR when social constructivism is used as a unifying framework. In order to investigate this, the “V-model of empirical research” in figure 1 is used for showing how AR/DSR can be seen as methods for theory development through model development, similar to how the V-model is used in information systems development and software engineering (Zahran, 1998).
The literature review is structured as an argument in four parts. First there will be a presentation of the Burrell-Morgan matrix to illustrate how Järvinen (2007) and Ivari & Venable (2009) have different perceptions of how to place AR and DSR within the matrix. Then there will be a presentation of an alternative definition of DSR, following the reasoning of Simon (1996). Applying the perspective of van Fraassen (1980), it is then argued that natural science can also be seen as the DSR of building and researching scientific models. The next section argues by the way of Creager (2007) that the model concept is also central to hermeneutical sciences, although the model in such cases may typically consist of exemplary case studies or narratives, viewing history, anthropology and such as a kind of model engineering (e.g. structuralism). Finally, identifying AR as a humanistic research style for doing intervention research, this is seen as DSR in the style of designing intervention strategies for exemplary cases.

The argument consists of using a two-by-two framework of scientific paradigms along the axes of subjective-objective and regulation-change, similar to framework developed by Burrell and Morgan (1979), and then showing how model development can be interpreted as a defining part of scientific practice within each of the four scientific frameworks. The argument is formulated as a part of the literature review in section two. In order to give an example of the practical consequences of this perspective, in section three a framework for empirical application of the model will be explained, using the author’s biography of doing AR/DSR-inspired Software Process Improvement (SPI) research. The results, in terms of how this SPI research methodology has evolved over time, are presented in section four. In section five, the results will be used for further discussion of the model in figure 1, focusing on the models ability to take practical issues into account. The conclusion, in section six, states that the model in figure 1 fits for doing mixed approach research, selecting methodologies from different corners of the subjective-objective-regulation-change matrix, including the subjective-change corner of AR and the objective-change corner of DSR.

2 LITERATURE REVIEW

2.1 Comparing action research and design science research

According to Järvinen (2007), AR is an important but low-level research approach. By analysing and aligning matching activities and characteristics of both AR and DSR, he finds that AR and DSR have five main activities in a research cycle that match, and he also identify additional seven characteristics of AR and six characteristics of DSR that match. Overall, he finds AR and DSR to be fundamentally similar, and suggests AR being more closely associated with DSR rather than necessarily seeing it as qualitative research per se.
If Järvinen is correct in his reasoning, this could prove helpful in the field of software process improvement (SPI) research. DSR has been used in SPI research for investigating and improving specific SPI frameworks like, say, the SW-CMM model for small organizations (Paulk, 1998), but with little focus on how such models fit with cultural issues and organizational politics. AR has been done for researching relationships between technology and change (e.g. Börjesson, 2006), but with little focus on the details of the SPI standards. However, as the SPI practitioner has to deal with both political and technical issues at the same time, an AR/DSR approach, properly balancing social research with engineering research, could provide very useful research.

However, according to Livari and Venable (2009), comparing AR and DSR along the paradigmatic dimensions of ontology, epistemology, methodology and ethics, there are examples of instances of AR and DSR being similar, but on the whole they represent significantly different research approaches. As they use a method for comparing paradigms adopted from Burrell and Morgan (1979), the diagram in figure 2 describes how the author’s interpretation of how the opinions of Järvinen (2007) and Livari & Venable (2009) differ.

As a slight modification of the matrix suggested by Burrell and Morgan (1979), the vertical axis in figure 2 follows Simon’s (1996) convention of action (engineering science; “science of the artificial”) and perception (natural science) rather than the distinction between “regulation” and “radical change” from the original matrix, although the meaning should be more or less the same. In addition to giving a visual presentation of the views on whether AR and DSR are similar or not, the purpose of the diagram is to illustrate the argument of the paper in terms of how a constructivist philosophy of science, by following the steps one to four in the diagram, can eliminate the paradigmatic differences identified by Livari and Venable (2009) and end up with a AR/DSR approach similar to what is suggested by Järvinen (2007).

Rather than applying the various definitions of DSR suggested by IS researchers, the starting point will be a return to the DSR defined by Simon (1996). Applying the perspective of van Fraassen (1980), it is argued that natural science can be seen as the DSR of building and researching scientific models. The next step is to argue by the way of Creager (2007) that the model concept is also central to interpretative research, the model in such cases typically consisting of exemplary case studies or narratives, viewing models used in history, anthropology, psychoanalysis etc. from a viewpoint similar to that of structuralism. The third step is a return back to action and conflict, suggesting models based on (“soft”) game theory, using Checkland’s (1981) soft systems methodology (SSM) as a prime example. Having consistently kept a focus on the models rather than the target system, this results in a unified perspective on science with no fundamental difference in how science is conducted in neither of the four quadrants.
2.2 Design Science Research and Software Process Improvement

Although, as mentioned by Iivari and Venable (2009), DSR may be seen as a research method that has been practiced within engineering for centuries, the source reference on DSR used by most of the current IS debate is Herbert Simon’s 1969 discussion of how engineering science (“the science of the artificial”) compares with natural science (Simon, 1996). In his book, Simon makes no distinction between quantitative research and qualitative research, as is often done in social science (e.g. Rudestam & Newton, 1992), but rather draws the demarcation line between whether the system being researched is a system with a purpose or not. As an illustration, how things interact according to the laws of Newton is an example of natural science (Simon, 1996, p. 1), and how a clock functions is an example of the science of the artificial (ibid., pp. 5-6).

In order to elaborate the differences between a natural system and an artificial system, Simon (ibid., pp. 6-7) describes artificial systems (purposeful systems) to consist of an inner environment, an outer environment and an interface. This allows him to formulate the general research problem as that of searching for the optimal interface, like researching the design of a clock in terms of finding the optimal design to fit with the inner environment (requirements regarding the logic of the clock) and the outer environment (requirements regarding use of the clock). As pointed out by Krick (1969), all engineering problems, or design problems in general, consist of finding an optimal design and design process for an artefact.

Even though Simon stresses the need for heuristics in a world of bounded rationality, in order to understand his concept of DSR, the ideal DSR approach is the total rationality of OR. As the name implies, Operations Research (OR) was originally conceptualised as scientific research in terms of understanding military operations and other operations (systems) by developing mathematical models for representing problems, solving the problems, and interpreting the solutions in the context of the system (Beer, 1968).

![Diagram](image)

*Figure 3. OR method and scientific method as two different processes (Phillips et al., 1976, pp. 5-6)*

On the other hand, what is being described in the textbook literature on OR are mathematical methods of industrial engineering (Turner et al., 1987), OR presented as a methodology on how to apply management science in industry, and not as a method for doing research. Phillips et al. (1976, p. 6) makes this point clear in the following way:

*The process of acquiring the conviction that a model actually “works” is commonly called validation. […] “Validation” is a considerably weaker term than “proof” or “verification”. To further clarify the modelling approach to problem solving, contrast it to the experimentally based “scientific method”. […] Here, the first step is the development of a hypothesis which is arrived at, generally by induction, following a period of informal observation. At that point an experiment is devised to test the hypothesis. […] The result of the process is something that purports to be “truth”, “knowledge” or “a law of nature.” In contrast to model conclusions, theories are independently verifiable statements about factual matters. Models are invented; theories are discovered.*
Phillips et al (1976, pp. 4-6) also provide two diagrams in order to illustrate how models are used in OR and how this can be contrasted with the process of theory building (figure 3). However, as is clearly indicated by the quote, the kind of science Phillips et al. have in mind is natural science. The kind of knowledge DSR produces, is knowledge about artefacts, such as the models in OR. On the other hand, mathematical models of reality, such as Newton’s law of universal gravitation, are human constructs, often designed in the language of mathematics, sometimes capable of explaining and predicting empirical observations, but constructs nevertheless, and researchable objects as optimal solutions for the kind of problems envisaged by DSR. In the case of Newton’s law of universal gravitation, the inner environment would be Newton and other scientists researching alternative or better ways to model the relationship between attraction (force), mass and distance between two (large) objects, while the outer environment would be the system of objects under consideration.

From this perspective, the purpose of science is not to produce true facts about the universe, but rather to produce useful models of it, that can be used for prediction and/or explanation (van Fraassen, 1980; Cartwright, 1983). This does not mean that science is disconnected from reality, but it means that the purpose of theories are to explore models, and it is the models that provide the link with reality, not the theories. The idea of splitting theory into a syntactic part (theory) and a semantic part (model) comes from Tarski’s Model Theory and has been suggested as a general framework for science in general (Suppes, 1967).

To illustrate this point further, the interchange between normal science and revolutionary science, as suggested by Kuhn (1996), would correspond with improving models and changing models, without having to commit to whether a model (theory) is true or not. Rather than saying that the geocentric model of world was proved wrong in favour of the heliocentric model, the way to communicate this idea would rather be to say that heliocentric model is a better model, although still only a model.

What has been argued so far, starting with Simon (1996) and ending with a model-oriented philosophy of science (Suppes, 1967; van Fraassen, 1980; Cartwright, 1983), is the following: If we consider the world to be socially constructed, in the sense that the mathematical language and mathematical models are socially constructed, then DSR is the only science there is. Researching SPI would in principle be no different from doing research in physics or chemistry. Rather than trying to figure out how a given organization works, what the processes are, what the culture is like, what the technology is like etc., and what kind of insights we can draw from this, the aim of the research should be to identify what is the best SPI model for understanding and making predictions about this organization, and how can we use the empirical observations for validating or improving the model.

In the SPI literature there are numerous SPI models, such as ISO 9000, ISO 15504, CMM, Bootstrap (e.g. Zahran, 1998; Hoyle, 2006), CMMI (Chriissis et al., 2003), various agile methods (Schwaber, 2004; Poppendieck & Poppendieck, 2003; Cockburn, 2002), attempts at balancing discipline and agile (Boehm & Turner, 2004), general overviews based on best practices and/or research findings (e.g. Dybå et al., 2002; Sommerville, 2001) or models focusing on social issues and critical theory rather than technical guidelines (e.g. Nielsen & Kautz, 2008). The type of DSR suggested for turning the SPI process into a research process could ask all sorts of research questions, but common to all the questions, whether they are designed to be answered through statistical hypothesis testing or to be explored by some kind of qualitative approach, is that researching the question should produce new knowledge explicitly linked to validating or improving a given model.

The model being investigated does not necessarily have to be the espoused model of the organization. As illustrated by Seddon (1997), Brunsson et al. (2000) and others, the fact that people in an organization say they are following a given quality standard, for example ISO 9000, does not necessarily mean that the ISO 9000 model is the best model for understanding SPI in the organization. Perhaps the best model for understanding how SPI is conducted in some organization claiming to be following a given SPI model is that they have no model at all. As the criterion for the type of DSR suggested here is that it is necessary to focus on a particular SPI model, then the SPI model in such a case would be the model called "having no SPI model at all", and then try to see whether the empirical observations can be used for adding new knowledge to the model of “having no SPI model at all”, validating this model in terms of how good it is at explaining and predicting.
The literature review so far should explain step one in figure 2. In the next section follows an explanation of the remaining steps two, three and four.

2.3 The debate concerning Action Research and Design Science Research

Although Simon (1996) talks about design in a wide sense, including architects, painters, musical composers, managers, and decision makers of all kinds, big and small, his book is a book about decision making in complex environments, replacing the model of economic man (selfish and rational) with an alternative model where the part about rational has been replaced by heuristic, e.g. using “rules of the thumb” for making rational decisions in an environment that is not fully predictably or cannot be fully understood. Due to this focus, there seems to be no loss of information if one should interpret the book to be about robots rather than people. In fact, the book is considered one of the foundational books in the field of artificial intelligence (Russell & Norvik, 2003, p. 29).

Action research, on the other hand, is often associated with qualitative research as the problems that concern people working with action research often has to do with trying to figure out how people think, feel and behave. Rather than assuming the model of economic man, metaphors or models for understanding action research can be that of jazz musicians trying to improve their way of playing or detectives trying to solve a murder mystery (Person et al., 2007).

The fact that action research is primarily used for making social change or gaining psychological insight does not necessarily mean that it is incompatible to the engineering logic of DSR. In the foundational paper on action research for improving the social conditions among social minorities (Lewin, 1946), one of the major issues brought up is the importance of establishing irrefutable knowledge through the use of measurements, statistical analysis and hypothesis testing. As pointed out by Gold (1999), both Lewin and his contemporaries saw action research as a research approach within the tradition of “scientific management” (Taylor, 1911), although Lewin’s hypotheses were often different from Taylor’s, e.g. wanting to prove that collaboration among workers would have a better long-term effect on productivity than Taylor’s focus on individual financial incentives.

Although Järvinen (2007) bases his comparison of AR and DSR upon contemporary writings within the IS tradition, returning as we do now to the roots of AR and DSR, it is not surprising that he finds the two methods to be similar, especially as he concludes his paper by indicating how elements from the engineering science of DSR can be used for improving the validity of the claims made through qualitative arguments in AR. Livari and Venable (2009), on the other hand, choose to compare AR and DSR in terms of ontology, epistemology, methodology and ethics, finding differences on all accounts, but nevertheless suggest a framework for combining AR and DSR, looking particularly at the issue of applying AR as a part of a DSR project.

In the case of SPI, a mix of DSR and AR can be useful as much research on the failure of SPI project have been attributed to social and psychological issues (Sommerville, 2001). On the other hand, as statistical process control and cyclic PDCA research designs are inherent research paradigms that follow automatically as a part of many SPI models, such as the CMM models (Humphrey, 1989), it may initially not be quite obvious how the “story telling” results of AR can be used for validating or improving the SPI under consideration. Nevertheless, as shown in the literature review of natural science as DSR above, the fact that sense-making by constructing stories are examples of social constructions, just like the mathematical models, the way people believe they are following a given SPI standard when they are not, or believe they are not following a SPI standard when they actually are doing so, can provide insights on the communicative powers of the SPI standards, their ability to be understood, their ability to be misunderstood, etc. Although AR as a research paradigm may not produce evidence that helps validate a given SPI model, the paradigm may be helpful in giving examples of social contexts for indicating how and why SPI models are understood or misunderstood.

The kind of insights produced by this approach may be somewhat similar to Berger and Luckman’s (1966) sociology of knowledge, not primarily focused on whether people have knowledge in a scientific sense (i.e. valid knowledge), but more focused on trying to understand mechanisms of persuasion or other
social mechanisms, such as the relationship between power and knowledge in the case of a particular SPI model, perhaps adding design features to the model that can make it more easy to adopt.

Knowledge Management (KM) should be considered an important part of SPI, as can be illustrated by the fact that total quality management (TQM) is often described as a fad that lasted about five-to ten years (1985-1995), while the more or less identical approach called Six Sigma has lasted twenty years and is still popular. Could it be that familiar and easily interpretable terms like “total”, “quality” and “management” made TQM both an instant success and also quickly forgotten, while the more cryptic yet easily remembered phrase Six Sigma (6s) sounded sufficiently “scientific” and “mystical” at the same time in order to make it last? Legge (2002), among others, has written about TQM in this context, stressing the importance of how SPI frameworks like TQM are communicated, regardless of what their scientific merits may be.

Of particular interest in the case of knowledge produced by the AR approach is the history of the SPI models. In presenting the literature review of the AR and DSR research models above, the focus has partially been on explaining the similarities and differences between AR and DSR based on the history of the models. Even though Simon (1996) is clearly focused on the timeless aspects of decision making, the search and evolution for such ideas is embedded in the social world through history. Even though mathematical concepts and ideas are eternal ideas, the mathematics of Pythagoras and Euclid being independent of time and social circumstances, the history of mathematics is constantly being rewritten, changing with fashions, politics and paradigms of the age (Boyer & Merzbach, 1989).

AR can play an important part in providing insights on how psychological and social factors related to the production, spread, implementation and success of various SPI models. To summarize the core idea of this literature review, in the AR/DSR model suggested in this section, DSR aims at constructing, analysing and validating scientific models (SPI models) while AR aims at writing the history of science. Both science and the history of science play important roles in the development of science.

Returning to step two in figure 2, going from natural science to humanities in the case of SPI means a change of focus from construction of models about the world to constructions of models about the mind. Piaget’s (1970) structuralism is an example of trying to gain insights from mathematical models without necessarily doing calculations, an approach that appears to ease physicists (Sokal and Bricmont, 1998) and cause enthusiasm from mathematicians (Devlin, 1997; Tasic, 2001). From the viewpoint of van Fraassen (1980) and others, the model is more important than theory for producing scientific knowledge, and as argued by Creager (2007), models play a particularly important role in qualitative research; case studies, exemplary narratives, classification systems etc. can all be viewed as models in the sense van Fraassen (1980) focuses on.

The step from humanities to action research, step three in figure 2, could be illustrated by Checkland’s (1981) arguments concerning the difficulty in performing operations research (DSR) when there is no proper understanding of what the problem is and what should be optimised. Along with Checkland’s approach, there are many “problem structuring methods” (Rosenhead & Mingers, 2001) that can be used as models in the context of doing action research. Similar to how Simon (1996) differentiates between engineering science and natural science, the difference between interpretative research and action research, from the model perspective, is that AR models are models of action, i.e. strategies in what a mathematician might conceptualize as the game theoretical description of the organization (e.g. Rosenhead & Mingers, 2001, chapter 10-11).

3 METHODOLOGY

As the aim of this paper is to validate a model for doing AR/DSR as an integrated part of SPI, a model that was presented as a V-model in figure 1 and will now be referred to as the “V-model for SPI Research” (VSPIR). In the language of DSR, this means that the artefact in consideration is the VSPIR as represented in figure 1. In the context of AR, the focus is on the social system surrounding the artefact, i.e. the social surrounding of the VSPIR method. This means that the object of the study will be the socio-technical system consisting of the VSPIR and its shareholders. More specifically, we are interested in building a socio-technical model of this target system by combining the technical drawing in figure 1
with an exemplary narrative generated by investigating the VSPIR in use and performing interventions with the VSPIR in use.

As the VSPIR has been developed and used by the author, struggling with SPI for the past 18 years, the empirical study will be conducted as retrospective research by testing against this autobiographical material. As the development of the research method has been a story of conflict and politics, the story should fit with an integrated AR/DSR approach, as is assumed by the VSPIR model.

Both DSR and AR are often described as cyclic research processes, going through AR steps like diagnosis, action planning, action taking, evaluation and specified learning (Sussman and Evered, 1978), or following the steps of the design cycle in the case of DSR. In the case of VSPIR we follow the six steps defined in figure 1 as the scientific model is designed, implemented and validated. Depending on the outcome of the validation, a new cycle will start.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Period</th>
<th>Organizations involved</th>
<th>Scientific publications</th>
<th>Other publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1991-98</td>
<td>DNMI, UiO</td>
<td>17</td>
<td>302</td>
</tr>
<tr>
<td>2</td>
<td>1999-present</td>
<td>DNMI, SKD, UiO</td>
<td>34</td>
<td>210</td>
</tr>
</tbody>
</table>

Table 4. Research cycles and count of theoretical output

As illustrated in table 4, the case to be investigated here consisted of two cycles. Within each cycle the VSPIR method was executed repeatedly, producing the publications counted on the far right columns of the table, but the VSPIR model was only properly evaluated twice. In the first cycle the method was validated by having the publications assessed by independent scholars as to whether they would comply with the requirements of a PhD. The final part of the second cycle consisted of the author being part of a formal PhD program, being able to more actively be a part of the testing of the VSPIR method himself. For conducting hypothesis testing according to DSR principles, the Shewart (1938) idea of interpreting statistical process control (SPC) as testing the null hypothesis of the process being in stable is used.

4 CASE STUDY OF AR/DSR RESEARCH MODEL

4.1 First cycle

Design of model. In 1992, I started working for the Climate Department of the Norwegian Meteorological Institute. Although my work description consisted of working with meteorologists in information systems development in terms of developing a climate database KLJENAS and surrounding systems, I was given the formal title of research scientist. This had more to do with organizational politics than what we were asked to do. Some time prior to me entering the organization, there had been an issue about people with the same education working for the R&D department and getting better pay because they were following the salary tables for research scientists while the meteorologist in the Climate Departments were following a different and less fortunate table. As most of the climate meteorologists were doing scientific consultancy in terms statistical analysis of weather stations, some of them also doing scientific research, it was decided to establish a scientific climate research journal with ISSN number, and change the formal work descriptions of all meteorologists to research scientists, regardless of what they were doing.

Although there was no peer pressure in writing and publishing research, when socializing with others in private or outside the Climate Department, I felt a psychological need for justifying the title of research scientist, so after getting on with the practical work of designing and developing the KLJENAS system, I targeted the topic of meteorological quality control to be relevant and interesting based on my background in industrial mathematics. The initial model thus consisted of the want to do research.

Implementation of model. Deciding to specialize in quality control proved to be a good choice. During the first few years, I managed to analyze and design algorithms for performing quality control. The first publication was a mathematical analysis of a method called the dip test (Ogland, 1993). For this contribution, I got good feedback from colleagues at DNMI, and I also got feedback from the US and (much later) from Puerto Rico, where the method was referred to as a part of a masters degree on
geophysical quality control. Based on this instant success, I continued sharing my time between developing KLIBAS and doing design research on quality control, contributing to conferences, a Nordic research project that later developed into a European Union research project. The implementation of the VSPIR model thus consisted of the method and the community of practice.

**Validation of model.** Although I felt happy with the organization, my financial situation had not improved at the same rate as society in general, so I asked whether it might be possible to use my extensive list of publications as an argument for moving up to the next research scientist level. I was asked to select some representative work to have it evaluated by the university, and if my competence was considered to be on a PhD level, I would get promoted. The result of the evaluation, however, was that despite hundreds of publications, the fact that the body of work had been published without peer review, resulted in rejection, so I remained within the same position. Both the technical and the social aspects of the VSPIR model were seen to represent a challenge.

The learning from this incident was to make contact with the university in order to find out what was required for getting enrolled in a PhD program and focus more strongly on writing scientific reports, simplifying the software engineering part of the work and put more effort into the scientific part.

4.2 Second cycle

**Design of model.** Although doing DSR-style research on software process improvement and quality control was highly interesting from a personal point of view, my diagnosis was that it might be interesting also to look for other organizations where I could continue contributing and developing my research interests. After the result of the previous evaluation, I consequently started looking at other places I could do research driven SPI. This was a mental redesign of the VSPIR model, especially the social part, but also the technical part.

**Implementation of model.** Late 1999, I left DNMI and started a career first as quality management consultant at the Directorate of Taxes (SKD) and quickly got promoted to director of quality management for the IT function. As the head of the IT department was a civil engineer with strong beliefs in science and technology, I gradually found out how to adjust the methods I had used for implementing SPI at DNMI for implementing SPI at the much larger organization of DNMI. Despite working against the odds most of the time, the approach and ideas I developed were extremely successful, up to a certain point (Ogland, 2008).

I continued publishing research reports and going to conferences, but this time there were no available research outlets with ISSN number, and the documents were made part of the internal SKD archive, not generally accessible to the public. However, at a point I managed to get two employees that were useful for discussing how to make the approach more scientific.

**Validation of model.** Despite spending the same amount of time and energy on improving the SPI system as I had been doing at DNMI, in the end I made some political mistakes and was asked to terminate my position as director of quality management. In exchange I was given the opportunity of getting enrolled in a PhD program and do SPI research.

The period 1999-2005 had been a period of exploring many SPI frameworks, such as CMM, ISO 9000, ISO 15504, ISO 12207, CobiT and EFQM, the main challenge was the question of how to design quality management systems in complex organizations, in this case meaning the complexity of a politically turbulent organization. I have later written much about specified learning in looking back upon this period while also continuing collecting data and doing AR/DSR at SKD on a small scale. The scientific output so far has consisted of one paper contributing insights as a result of an interpretive analysis of the organization, two papers contributing theory on the engineering problems of SPI design, and six papers contributing to the AR strategy of how to deal with political challenges.

At the end of the second cycle, the VSPIR model appears to be functioning much more consistently with the intensions of the model.
4.3 Analysis

As pointed out in section 2.1, although a model validation is weaker than hypothesis testing, it should be possible to make parameter estimations and test hypotheses about the VSPiR model in order to build knowledge. What we would like to confirm in this subsection is that there is a relationship between following the suggested research model and productivity and/or quality of research.

![Graphs](image)

**Figure 5.** SPC chart for estimating productivity (left) and run chart for indicating quality (right)

The statistical process control (SPC) chart on the left side of figure 5 indicates research productivity in terms of scientific publications. Process average and control limits have been calculated for each of the two cycles of research, ignoring the extreme value of 1999 for reasons explained in the end of section 4.1 above. In order to predict whether entering a PhD program in 2006 had any impact on productivity, the SPC parameters for the second AR cycle are based on data from 2000 to 2005.

Interpreting the SPC as an hypothesis test, as described in the methodology section, what we see is that the null hypothesis is rejected at the end of the first cycle, which has already been mentioned, but the second cycle remains stable despite the fact that the final few points of the curve represent publications done during the PhD program. In other words, there is no change in productivity.

The right hand side of figure 5 is supposed to indicate research quality by calculating the h-index (Hirsch, 2005) from publications identified by Google Scholar. The h-index points towards a quality improvement as I joined the PhD program, but the only references identified by Google Scholar are instances where I refer to my own work.

5 DISCUSSION

The purpose of the case study was to validate the implied AR/DSR model presented in the literature review section. Although the model was more or less a straight forward AR/DSR cycle research design, using the similarities pointed out by Järvinen (2007) by the way of linking the research model with the philosophy of science of people like Suppes (1967), van Fraassen (1980) and Cartwright (1983), the object of research was the model rather than reality, a somewhat dual approach to the description of the scientific method presented in a textbook on operations research (Phillips et al., 1976). By using this dual approach both on DSR and AR, it was suggested that the result should be theoretical contributions in terms of both validating the model and contributing to the history of science for this part of the model being tested.

Looking at the case study in order to validate the research model in figure 1, the first chart in section 4.3 said that there had been a stable output of about two scientific publications each year for 18 years, from 1991 to the present. The second chart confirmed the 1998 external assessment of the research methodology, namely that it was a low-quality approach. Although the approach had been continually developing for the whole period of 18 years, it was only during the past two years, as the author had been
enrolled in a formal PhD program, that the h-index showed any results at all, and the only reason for this was the fact that the author had published papers where he made references to his own work.

Nevertheless, despite the fact that the results are not impressive, the study shows that it is possible to do SPI research following the VSPIR method developed during the case period.

Of the first 33 scientific publications prior to entering the PhD program, four could be characterised as natural science, contributing in the field of climatology, while the remaining 29 would be DSR, dealing with models and methods for quality control and process improvement within the SPI context of DNMI, all published as conference or journal papers with ISSN identification, but without proper peer review procedures. The next nine papers were related to preliminary results at SKD, and although they were published in the non-scientific SKD journal, they could perhaps be seen as action research papers.

It is only with the final nine papers, however, written as a part of the PhD program, that the method has matured to a level where it could be considered proper academic research. As pointed out in section 4.2, the papers contributed to three of the four quadrants in figure 2; DSR, AR, and humanities but not natural science. All in all, however, the case study illustrates how the research methodology had no difficulty crossing all paradigms in figure 2 by constantly focusing on the model of the target system and making small adjustments in order to investigate the model as the target system changed between natural and artificial, problems dealing with meaning or problems dealing with control.

The main idea in figure 1, namely that there is a correspondence between the V-model of constructive learning (Novak & Govin, 1984), the “constructive empiricism” of van Fraassen (1980) and the V-model of systems development (Zahran, 1998), has been validated in the sense that the 18 year evolution of the research methodology in the case study has implicitly been working according to this idea, something that has been made explicit due to the publications during the past few years.

6 CONCLUSION

The empirical part of this study was motivated by the fact that doing research in the Software Engineering discipline of Software Process Improvement (SPI) deals with both organizational issues and technological issues, thus making it unclear whether the engineering approach of Design Science Research (DSR) or the humanistic approach of Action Research (AR) is the better, or whether one should mix the two, if that is epistemologically possible. Within the current debate in the Information Systems (IS) community, some argue that AR and DSR are fundamentally similar (e.g. Järvinen, 2007), while others argue that they are fundamentally different (e.g. Ivvari & Venable, 2009).

As a central part of the argument for AR and DSR being fundamentally different had to do with paradigmatic matters identified through the matrix developed by Burrell and Morgan (1979), the idea of the study was to bypass such paradigmatic matters by suggesting an isomorphism between the Burrell-Morgan matrix and a matrix that uses van Fraassen’s (1980) philosophy of science of anti-realism (or ignoring the issue of ontology) and rethinking different epistemologies in a way that turn them all into social constructivism.

The isomorphism is based on the idea that all kinds of science can be seen to deal primarily with model construction and secondary with theory development, especially if theory is supposed to be based on mathematics or linguistic constructions that perform in the same manner as mathematics does in physics, engineering and economics. By a model van Fraassen (1980) means a non-linguistic concept, typically something that can be represented or partially represented by a drawing. Following the observations by Creager (2007), for those parts of academia where mathematical theory plays a minor part or no part at all, models are used in terms of model systems, cases, exemplary narratives etc., thus making Ivvari and Venable’s observed difference between AR and DSR into a difference between what kind of models are being produced. A typical model investigated by DSR might be an algorithm or a procedure, and a typical model in AR may be an exemplary narrative of how a change happened in a given context.

As the empirical case of software process improvement (SPI) research showed, working with mathematical models and exemplary narratives at the same time seems like a more natural way of designing SPI research than representing different types of problem by different paradigmatic points of
view, at least when it is possible to collapse all seemingly different paradigms into the unifying van Fraassen perspective. In other words, the scientific contribution of this paper has been to point out how Järvinen’s (2007) argument about AR and DSR being similar is an important contribution to SPI research, stressing the usefulness of a van Fraassen-like perspective, focusing on the scientific model as a mediator between theory and target system (the circle in figure 1), in order to bypass the paradigmatic differences between AR and DSR pointed out by Livari and Venable (2009).

References


