Building Theory in the Sciences of the Artificial

Shirley Gregor
The Australian National University
Canberra
Australia ACT 0200
+61 (0)2 61253749

Shirley.Gregor@anu.edu.au

ABSTRACT
This essay extends Simon's arguments in the *Sciences of the Artificial* to a critical examination of how theorizing in Information Technology disciplines should occur. The essay is framed around a number of fundamental questions that relate theorizing in the artificial sciences to the traditions of the philosophy of science. Theorizing in the artificial sciences is contrasted with theorizing in other branches of science and the applicability of the scientific method is questioned. The paper argues that theorizing should be considered in a holistic manner that links two modes of theorizing: an **interior mode** with the **how** of artifact construction studied and an **exterior mode** with the **what** of existing artifacts studied. Unlike some representations in the design science movement the paper argues that the study of artifacts once constructed can not be passed back uncritically to the methods of traditional science. Seven principles for creating knowledge in IT disciplines are derived: (i) artifact system centrality; (ii) artifact purposefulness; (iii) need for design theory; (iv) induction and abduction in theory building; (v) artifact construction as theory building; (vi) interior and exterior modes for theorizing; and (viii) issues with generality. The implicit claim is that consideration of these principles will improve knowledge creation and theorizing in design disciplines, for both design science researchers and also for researchers using more traditional methods. Further, attention to these principles should lead to the creation of more useful and relevant knowledge.

Categories and Subject Descriptors
K.m [MISCELLANEOUS]

General Terms
Design, Experimentation, Theory

Keywords
Philosophy of Science, Philosophy of Technology, Design Science, Sciences of the Artificial, Practical Sciences, Theory, Theory Building

1. INTRODUCTION
From the last half of the twentieth century onwards there has been an exponential increase in knowledge surrounding technology, including developments not only in information technology (IT) but also in areas such as bio-engineering. These new and complex technologies have not only altered the way many people live and work, but also pose challenges to how we systematically accumulate and disseminate knowledge. It is not clear that the methods that have been used in the traditional sciences, either natural or human science, translate directly to the "artificial sciences". Thus, this essay continues Simon's quest, described in *The Sciences of the Artificial* [37]:

We need a science of design – intellectually tough, analytic, partly formalizable, partly empirical and teachable.

In Simon's terms, the sciences of the artificial "are concerned not with the necessary but with the contingent – not with how things are but with how they might be – in short, with design" ([37] p. xii). In these practical sciences, such as agriculture, medicine and information technology, useful knowledge is required for people to do things in the world; when they act and construct artifacts to achieve some purpose. The artifacts range from improved crop cultivation methods, to surgical procedures, to drugs and to computer software, hardware and systemized processes (methods) for organizing and managing IT.

The author has great regard for Simon's seminal work but aims at taking some of his arguments further. Simon and subsequent work in design science has focussed on knowledge concerning the building of artifacts and has tended to treat the study of artifacts once constructed as something that should be passed back to traditional science methods [21, 29]. In contrast, the premise of this paper is that we need to consider theorizing in a holistic manner across the whole of a science of the artificial, both in the building and the observation of artifacts.

To this end the paper addresses some fundamental questions:

1. How can theorising in the sciences of the artificial be understood in comparison with other branches of science?

2. Does the Scientific Method apply to the Sciences of the Artificial?

3. Is a special framework needed for thinking about theorizing in the Sciences of the Artificial?

4. Do the sciences of the artificial present special challenges for theorizing?

The paper proposes theorizing in both an interior mode, in the design and development of the inner environment of artifacts, and also in an exterior mode, where the artifacts are theorized about in their outer environment. These two modes are seen as "two sides to a coin"; they are intertwined and both contribute to the development of knowledge concerning artifacts in a practical science. The paper differs from prior work in that it applies not only to researchers who personally develop artifacts and theorize about the results in the interior mode (as in some conceptions of design science), but also to researchers who carry out more traditional theorizing in the exterior mode in a manner common in mainstream Information Systems (IS) journals.

Although research and theorizing in relation to design science is increasingly dealt with [15, 21, 29, 37, 41, 43] questions of knowledge creation and theorizing are still relatively unexplored. Processes of theory building and theorizing continue as a subject of debate in many fields. Management journals have had a number of issues devoted to the topic of theorizing, including the *Academy of Management Review* (1989, Vol. 4, No. 4), and *Administrative Science Quarterly* (1995, Vol. 40, No. 3).

The contribution of this paper is that it brings together thinking about knowledge creation and theorizing in the artificial sciences paradigm that has, till now, been relatively piecemeal and suggests a framework and a number of important principles for artificial science theorizing. In this respect the paper contributes to the rather sparse extant work in the philosophy of technology [35]. The essay also has practical significance in that it facilitates the systemizing and capturing of knowledge that informs action and practice in the real world. There is a further ethical dimension in that in the practical sciences we produce knowledge that informs actions that can directly affect people's wellbeing, as in the development of safety critical systems. If system failures occur then they can have very real and immediate negative consequences. The overarching aim guiding the paper is a better understanding of how useful knowledge can be developed in academic research and from practice.

Some clarification of terms used and the underlying position of the author are required as a further introduction to the paper. Theory is seen in essence as a generalized body of knowledge, with a set of connected statements expressing general relationships among constructs that refer to entities of different types, both real-world and theoretical. Consistent with Gregor [14] it is recognized that views on theory vary and that different types of theory can be distinguished, dependent on the primary aims of the theory; whether analysis, explanation, prediction, or prescription for design and action.

The essay rests on the adoption of some form of realist ontology, in the sense that the practical knowledge and theory developed are taken to include constructs that refer to entities existing in the real world independently of human perception. A naïve form of realism is, however, rejected, in that our perceptions of the external world are not taken to be unambiguously "true", but should be treated as provisional and liable to error. Epistemologically it is recognized that there can be many approaches to knowledge development.

Much of the argument in the paper applies to a number of the artificial sciences. The focus of the paper, however, is on the disciplines of Information Technology and much of the discussion is couched in terms that apply to Information Systems (IS)

The paper proceeds by directly addressing the research questions that motivated the paper. In the course of the discussion a number of principles are derived to inform theory building in the IT disciplines. Table 1 gives an overview of these principles with their associated research questions as an early orientation to the direction the paper takes.

Table 1: Principles for theorizing in Information Technology as a Science of the Artificial

No.	Key idea	Principle
Question 1: Nature of Artificial Science?		
1	Artifact system centrality	IT system artifacts are central to theorizing
2	Artifact purposefulness	Purposefulness of IT artifacts is recognized and outcomes studied
3	Need for design theory	Different views of theory are needed, including design theory
Question 2: Use of Scientific Method?		
4	Induction and abduction in theory building	The importance of induction and abduction in theory building should be recognized.
5	Artifact construction as theory building	Theory building through artifact construction has special features.
Question 3: Special Framework?		
6	Linked theorizing in interior and exterior modes	Artifacts are studied in complementary interior and exterior modes.
Question 4: Special challenges?		
7	Issues with generality	Complexity in IT disciplines entails problems of scope, generalization and currency.

2. THE NATURE OF THE SCIENCES OF THE ARTIFICIAL ?

We begin with the first question posed:

How should theorising in the sciences of the artificial be understood in comparison with other branches of science?

Milestones in the history of the philosophy of science shed light on the place and nature of the artificial sciences. Although tool development (technology) has been a distinctive feature of humankind since its earliest beginnings, the rapid explosion of technological developments in the 20th century is of such comparative recency that its impact on science and scientific method has scarcely had time to be considered in perspective. Such a perspective is what this essay addresses, and it is why it begins with a short overview of scientific development.

Table 2 shows some key points in the history of Western science[1] and accompanying developments in the philosophy of science. Scientific epochs can be distinguished in terms of the underlying

---

[1] Given space constraints this very brief outline does not mention developments in Eastern, Byzantine and Arab civilizations during the Dark Ages in Western Europe 16].

phenomena which are studied, with successive epochs providing additional foci of study.

Table 2: Scientific development in Western culture

Epoch	Science	Philosophy of science
Early science	Primitive, with some steps towards experimentation (Aristotle).	Science not distinguished from philosophy. Epistêmê distinguished from technê.
Natural Science (17th - 18th century)	Age of Enlightenment/ Age of Reason. Advances in natural sciences (Gilbert, Galileo, Newton, Hooke, Boyle).	Word "scientist" coined (Coleridge/Whewell), Galileo and Bacon and the scientific method (experiments).
Human Science (19th century)	Sociology (Comte), Psychology (James), Psychiatry (Freud).	"Human science" distinguished by Dilthey (1883) (early interpretivism).
Artificial/ Practical Science (20th century)	Increasingly complex artifacts with invention of computers.	Herbert Simon "Sciences of the Artificial" (1969), Strasser "Practical Sciences" (1985).

Distinguishing these epochs shows how advances in science and corresponding shifts in the philosophy of science have occurred hand-in-hand.

In the early science era, science was not distinguished from philosophy, which dealt with a wide range of phenomena. The Greek philosophers did, however, consider the relationships amongst various fields of knowledge. Loosely understood in modern terms, *epistêmê*, or knowledge, was distinguished from *technê*, or practice [33].

With the "Age of Enlightenment" in the 17th and 18th centuries significant advances in science occurred with work by thinkers including Gilbert, Galileo, Newton, Hooke and Boyle. Subsequently, the philosopher-scientist William Whewell invented the word "scientist" in 1833 in conjunction with the poet Coleridge to replace the terms "natural philosopher" and "man of science", which had been used previously [38].

Another shift occurred in the 19th century with the rise of branches of knowledge concerned with the study of human beings, including psychology and sociology. The term "human sciences" was advanced by the German philosopher Wilhem Dilthey. Dilthey:

is best known for the way he distinguished between the natural and human sciences. Whereas the primary task of the natural sciences is to arrive at law-based explanations, the core task of the human sciences is the understanding of human and historical life. [28]

Advances in technology and the complexity of technological artifacts in the 20th century has been accompanied by increased attention to the philosophy of technology, although this attention is still limited. Simon's *Sciences of the Artificial* [37], with the first edition in 1969, is the landmark work that seriously considers

the nature of the disciplines that deal with phenomena that are artifacts. Simon's work contained such a large number of novel ideas that not all of them have been adequately pursued.

The purpose in depicting these different stages in scientific thought is to point out that ideas about epistemology and how science works changed with the phenomena studied. The first "scientists", Glibert and Galileo, were those who were able to apply the scientific method and compare hypotheses with experiments and observations. For some time science concerned itself mainly with naturally occurring phenomena; such as magnetism, the movement of the planets, electricity and the behaviour of gases. In the 19th century, however, interest grew in sciences concerned with living creatures including humans. Darwin published the *Origin of the Species* and the fields of sociology, psychology and psychiatry emerged. Diltthey's concept of the human sciences was important as he saw that the different objects of study meant that the human sciences were relatively independent of the more established natural sciences and that they had to deal with "the more complex networks of the historical world and the actual givens of human beings" [28]. Dilthey's work has influenced thinkers in interpretivism and hermeneutics, including Heidegger, Gadamer and Ricoeur. Even independent of this stream of thought we find questioning as to whether the methods of the natural sciences apply to the social sciences; for example, in concerns with the probabilistic nature of theory [31]

We can take this line of thinking further and argue that the fields of knowledge that concern artifacts again require a fundamental and critical examination of how science is conducted. Note that this was also Simon's concern ([37], p. xi);

The contingency of artificial phenomena has always created doubts as to whether they fall properly within the compass of science.

Our overview of the history of the sciences suggests that there could be value in distinguishing the sciences of the artificial from other major branches of science, to allow their special features to be carefully examined. The artificial sciences can be seen as constituting a third major form of science, in addition to the natural sciences and the human sciences. This is the view of the philosopher Strasser [39] who looked at the study of artifactual phenomena from the viewpoint of "the phenomenology of scientific life" (p. 55) (although he preferred the term "practical sciences" rather than "sciences of the artificial"). His definition is that a practical (artificial) science is:

a science which is conceived in order to make possible, to improve, and to correct a definite kind of extra-scientific praxis. (p. 59).

Figure 1 follows Strasser in distinguishing among different types of science. This categorization is approximate and some fields of study will lie across categories. Economics and sociology, for example, have both descriptive (human science) and prescriptive (artifactual) aspects. Strasser provides good arguments, however, for avoiding the term "applied science" for those sciences which have, as their primary aim, the development of knowledge that guides the construction of artifacts and interventions in the world, as in exemplars such as medicine, agriculture and information technology. In an applied science, the application of knowledge is incidental: it is not the primary aim of the science. Distinguishing

the artificial sciences separately, with useful knowledge as the reason for their existence, means that their special nature can be teased out.

Notwithstanding what has been argued so far, the view that the artificial sciences should be recognized as "science" is not universal. Hatfield ([18] p. 946) in the *Shorter Routledge Encyclopaedia of Philosophy* states that "the word 'science' now means primarily natural science, examples of which are physics, astronomy, biology, chemistry, geology and psychology, and it applies secondarily to social sciences such as economics and sociology". There is no mention of the word science applying thirdly to sciences of the artificial, which we will attribute to a narrow and somewhat old-fashioned outlook in the philosophy of science.
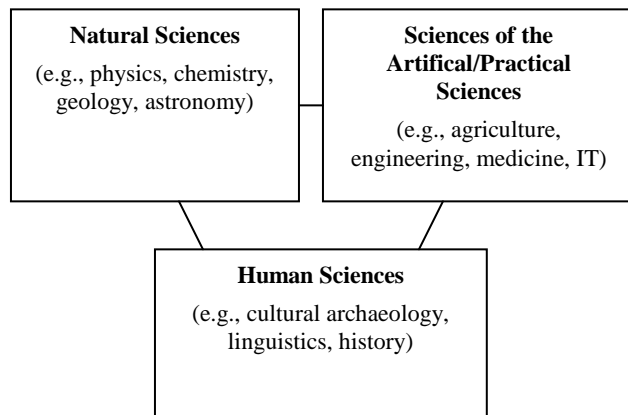

Figure 1: Three interrelated forms of science

In summary, the answer to our first guiding question is that the sciences of the artificial can be seen as a type of science, but one that likely differs from the natural and human sciences, which each in themselves have developed distinctive modes of enquiry. To illustrate the distinctive features of the sciences of the artificial, three principles are advanced to guide theorizing in IT, working from the conceptualization of IT as an artificial science.

Principle 1: IT system artifacts are central to theorizing.

An Artificial Science by definition concerns artifacts and we would expect to find an artifact playing a central role in theorizing. This point has been argued by others including [3, 24, 32, 46]. This principle distinguishes theory that belongs to IT from reference or kernel theories; theories that can be useful in the study of artifacts, as explanations for artifact behaviour or for design ideas, but do not have an IT artifact as a primary focus. An example is a theory of interpersonal trust, which belongs as a reference theory to the social sciences, whereas theory about how trust is engendered in online communications could belong to IS.

It hardly seems necessary to demonstrate that a further distinguishing and fundamental feature of IT is that it is intimately concerned with artifacts that are complex systems that involve both technical and social components, although the degree of emphasis on the technical versus the social aspects varies across the fields constituting IT. Basic definitions of a computer or

information system use words such as input, output, control, feedback and external environment. The first computers appeared after general systems theory had been advanced [1, 44] and relied on many of its concepts. Social science reference theories commonly talk in terms of social systems. Characteristics of systems are: they are open to, and interact with their environment; they acquire new properties through emergence and continually evolve, and; the parts of a system interact to form a whole which is independent of the separate constituents. Systems concepts include the system-environment boundary, input, output, processes, state, hierarchies, goal-directedness, and information [22]. IT-based artifacts that are systems include a computer program, operating systems, the Internet, an ERP system, an online auction and an online social network.

Conceptualizing IT as a science concerning complex system artifacts has a number of implications. That systems need to be treated differently in scientific reasoning and explanation has been argued for by biologists, who are also intimately concerned with systems, albeit those that animated by life. Nagel ([31] p.401) gives a comprehensive treatment of this claim and concludes that there are indeed good reasons for differentiating biology from the physical sciences:

One is the dominant place occupied by teleological explanations. The other is the use of conceptual tools uniquely appropriate to the study of systems whose total behaviour is not the resultant of the activities of independent components.

These considerations apply equally as well to the study of computer-based information systems as they do to biology and Simon has used systems theory concepts extensively in the *Sciences of the Artificial*.

Taking this foundational principle as a premise leads to a number of the other principles that are subsequently developed.

Principle 2: Purposefulness of IT artifacts is recognized and outcomes studied

A distinguishing feature of an artifact is that it serves some purpose, although purposes can be many and varied: for example, a jug has holding liquid as one purpose. This concept dates back to Aristotle in his depiction of the *causa finalis*, the final cause or end of an artifact - "what it is for" - one of the four causes of any thing [23]. The artifact's purpose relates to the context in which it is used. Heidegger [19] gave the example of a silver chalice, where in order to understand its purpose, we need to understand the religious ritual in which the chalice is to be used. The purpose of the artifact may not always be that of the original designer and some of its uses and effects may be unintended. Nevertheless, in studying artifacts it is needful to consider the goal, end or aim of the artifact, as originally intended or as arising in use.

Further, given this distinguishing feature, theorizing is more satisfactory if some assessment is made of the outcomes of the artifact's use: whether it achieves a goal or purpose in some way. The constructor of artifacts is usually expected to give some demonstration that the artifact at least works. The observer of artifacts will often evaluate the efficacy or consequences of its use.

This idea finds varied expression across a number of fields. In medicine the evidence-based approach uses the PICO model,

where P stands for patient or population, I for intervention (drug or procedure), C for comparison (against what alternatives) and O stands for outcomes (what you can hope to accomplish, measure, improve or affect) [40]. We find something of this thinking also in Van Aken ([42] p. 29) where he says in management design propositions "the independent variable must describe something of value to the organization, like financial performance". In IS, Jarvinen [26] proposed that we should use a goal function for measuring the goodness of a new artifact. The goal function could cover both intended and unintended consequences of the developed artefact.

Principle 3: Different views of theory are needed, including design theory.

The conceptualization of IT as a science concerned with artifacts leads us to the question of the place of theory in the Sciences of Artificial. Simon was not much concerned with theory and other researchers in the design science tradition have tended to follow along his path.

The case for different types of theory, however, depending on the purpose of the theorising, is made by [14] who distinguished five interrelated types of theory: Type 1 - theory for analysing; Type II - theory for explaining; Type 3 – Theory for theory for predicting; Type IV- theory for explaining and predicting; Type V – Theory for design and action. Gregor and Jones [15] provide further arguments for recognizing design knowledge as theory and show in detail the structure and components of a design theory (Type V theory). These authors conclude (p. 312) that:

The unambiguous establishment of design knowledge as theory gives a sounder basis for arguments for the rigor and legitimacy of IS as an applied discipline. A craft can proceed with copying of one example of a design artifact by one artisan after another. A discipline cannot.

Design theory will include prescriptive statements such as the technological rules of Bunge, which are "an instruction to perform a finite number of acts in a given order and with a given aim" ([5], p. 132). Van Aken ([4], p. 227) expressed these rules as taking the form 'If you want to achieve Y in situation Z, then something like action X will help".

The need for design theory as a special form of theory is something that distinguishes the artificial sciences from other sciences. In non-design disciplines such as the natural sciences, Type V theory is not necessary and there is no need for prescriptive statements that will guide design and action.

In IS, the lingering influence of conventional scientific views of theory lead to charges that papers are "atheoretic" if they do not include reference theories of the non-design type. Novice researchers are nervous about doing work that does not have "theory". An illustration is the plight of a PhD candidate who was investigating project management success. In response to a question as to whether she had considered knowledge of project management methodologies, she replied "but that's not theory". Support for a counter argument is provided by published papers that do not contain any theory other than design theory (an example is Iversen et al. [25], where design theory is developed from other design theories). In other branches of IT, such as Computer Science and Software Engineering, the type of

knowledge that is here referred to as design theory tends to be unquestioningly adopted as the norm.

Some further discussion of problems concerning design theory are discussed under the question of special challenges for the artificial sciences against our fourth question.

3. APPLICABILITY OF THE SCIENTIFIC METHOD?

The second question addressed is:

Does the Scientific Method apply to the Sciences of the Artificial?

This question is asked because theorizing at heart relies upon the methods judged suitable for knowledge development. Being scientific is often equated with the use of scientific methods, although the lack of consensus of what is meant by scientific methods is probably as great as the lack of agreement on precisely what the word science means. In very general terms, science is supposed to employ methods that compare "hypotheses with experiments and observation to weed out the wheat from the chaff" ([16] p. 63).

Possibly the most generally accepted view of the scientific process today is the "hypothetico-deductive method" (H-D method), named by William Whewell in his *History of the Inductive Sciences* [49], which built upon the work of Francis Bacon in the 17th century. Current forms of the hypothetico-deductive method owe much to the logician Charles Pierce in the 19th century, who combined the H-D method with symbolic logic and recognized three primary modes of reasoning that are at play in scientific enquiry: abductive, deductive and inductive reasoning. The H-D method can be described as ([7] p. 409):

Step 1: Explanatory hypotheses or conjectures are generated as a result of earlier inductions, a guess or creative imagination. The hypotheses are accompanied by statements of initial conditions.

Step 2: Predictive hypotheses are deduced from the conjectures.

Step 3: The hypotheses are subjected to experimental or observational tests. If the tests do not support the hypotheses then they are rejected. If repeated tests over time are in agreement with the hypotheses then they can be said to be rendered more probable.

These steps present an idealized form of the theorizing process. Step 1 represents a theory building phase, while Step 2 represents theory testing. This model is presented here in some detail because again there are differences of opinion that impinge on the artificial sciences.

Some divergences in opinion accompany historical shifts in thinking, particularly in regard to the competing claims of induction versus deduction as the dominant mode of reasoning and differences in opinion as to what are the important means of arriving at conjectures in Step 1, the theory building phase. There are different degrees of emphasis in Step 1 upon the importance of induction from evidence versus guesses and creative imagination.

Some areas of IT, notably IS, are influenced by what management scholars have to say about theorizing and a number of these scholars have de-emphasized the importance of building

inductively upon evidence. For example, Weick [48] saw theory construction as "disciplined imagination". Possibly the management field has been overly influenced by philosophers of science following Karl Popper, who notably had little interest in where conjectures came from in the first place and was also opposed to the idea that theories became "stronger" as they survived more and more tests, owing to problems with inductive reasoning [34]. Although inductive scepticism has a place, the philosophy of science now has more balanced views in tune with earlier views such as those of Bacon, Whewell and Pierce and there are again more varied views on how theory development can occur (see [13]).

There are also differing views of theory building (Step 1) in other areas of the human sciences. Examples include the grounded theory work of Glaser and Strauss [12]. Merton ([30] p. 47), writing about sociological theory, advocated the development of theory on an adequate base of "antecedent empirical enquiry". Bourgeois [4] provides a description whereby theory of the middle range is generated in a non-linear process with seven steps including theory generation by induction from an empirical base.

Views on whether and how the scientific method applies to the sciences of the artificial and technology are mixed. Bunge ([6], p. 174) says that methodologically, technological research "is no different from scientific research". However he presents a rather simplistic view of how solutions to technical problems are invented, ignoring much of the trial-and-error process by which new inventions are developed. Simon ([37], pp. 114-124) limited his discussion of the discovery of new knowledge to optimization problems and means-end analysis, where ordinary mathematical deductive logic can be used to identify the best option from a range of identifiable alternatives.

Designers, however, are not confronted only by optimization problems, but also by problems where the range of potential solutions is large and not identifiable at the design point. Such problems can not normally be solved by deductive logic. An engineer faced with the problem of building a bridge over a ravine has no clear guidance from deductive logic as to which design to implement. Often the best that can be done is to think that such and such a design worked in similar situations in the past, and reason inductively to assume that the design is likely to work again in this similar situation (engineering analysis indicating that there is no reason to believe it will not work). Thus, as has been pointed out by a number of authors, inductive logic is necessary in theorizing about artifacts.

The above reasoning leads to two more theorizing principles for the Sciences of the Artificial.

Principle 4: The importance of induction and abduction in theory building should be recognized.

This principle recognizes that deductive logic continues to be important in theory testing, but that in design disciplines especially the theory building phase is important and must employ induction and abduction in addition to creativity and imagination.

The proposal of abduction as an alternate reasoning mechanism is an attempt to understand and describe what practicing scientists do in theory discovery (theory building). Abduction is "inference to the best explanation" and refers to a process in which a

hypotheses or theory is developed as the best explanation of data that is available [45].

Principle 4 stresses the importance of grounding theory development empirically in available evidence, whether it is through the construction of artifacts or in observing what is being done in practice.

This principle has important implications from what is currently seen as "theory development" in IS. It implies that the generation of research models in IS deductively from reference theories for quantitative testing should be done with extreme care and not without some degree of grounding in IT use in practice. Often it appears that researchers have added variables to their models with only slight attention to their importance in something like a "shopping-basket" approach. Reference theories can indicate that a large number of explanatory variables could have some relationship with outcome variables and the inclusion of many variables will lead to more variance being explained in statistical models. This result does not mean that key factors that designers of interventions can manipulate to bring about desired results have been identified. Recent disquiet about the rather ad-hoc extensions to the Technology Acceptance Model (TAM) is an illustration of this problem (see the special issue in Journal of the AIS, 2007, Volume 8, titled Quo Vadis TAM?).

The special case of theory building when constructing artifacts deserves further elaboration, as discussed below.

Principle 5: Theory building through artifact construction has special features

Theory building through artifact construction, in what is referred to as design science, presents some special challenges for the H-D method. In no branch of science outside the artificial sciences does theory building involve the researcher or others actually constructing the objects of study.

There is an expanding literature on how knowledge development occurs in design fields [e.g., [17]. Some literature deals with processes of design and innovation and how designers work in practice [18]. Design itself can be thought of as more an art than a science. Atwood, McCain and Williams [2] give a useful overview of how the design community thinks about design in general, across many fields. Simon (1996) argued against design as a rational decision process and proposed that human designers, when confronted with a myriad of design choices, are likely to settle for good or satisfactory solutions rather than optimal ones.

Kuechler and Vaishnavi [27] provide a good overview of design science in IS. Design science research activities are often described in terms of design-build-evaluate cycles and as a problem solving process. Hevner et al. [21] utilize a means-end analysis conception of design activity. Other design research activities that can lead to design knowledge include action research and collaborative clinical research [42].

The design science literature tends to deal mostly with the first-hand situations where the individual or group constructing the artifact are the same people who reflect on the design and produce design theory. In Van Aken's terms this is the "**developing case study**" approach [41].

However, design theory can also be produced by researchers who reflect at second-hand on what others have done in constructing

artifacts. For example, IT professionals can produce artifacts that are new and interesting from a research point of view. For design theory to be produced, reflecting upon what has been done is required [36] and design principles need to be abstracted. This systemization of knowledge gained through practice is a legitimate academic activity and one that has led to a number of influential design theories. For example, Davenport and Short [9] abstracted ideas from case studies in 19 companies to first depict the general method of business process redesign. Van Aken ([41] p. 232) refers to this type of activity as an "**extracting case study**" or "best-practice" approach and notes that it has produced a number of very powerful technological rules (design theory), such as the Kanban-system and Just-in-Time approaches.

The reasoning that produces design theory from the process of artifact construction is theory building using induction – reasoning from the particular to the general – but does not rely on simple forms of induction (like counting number of instances). The fact that the theorizer can point to the steps in the process that led to the final form of the artifact, possibly showing also why these steps were taken, provides credible evidence of cause and effect.

Further, experimentation in design research can have a different role from that usual in the theory testing cycle of the H-D method. In a building phase, designers think of an idea then try it out informally to see if it works, make a decision then proceed on to other design decisions. This experimentation is part of the process of designing rather than being the experimental method as proposed in science to test cause-effect relationships. In anything more than a trivial design problem the designer will make very many design decisions and it would be infeasible to test every design decision point by conducting a formal experiment.

4. IS A SPECIAL FRAMEWORK NEEDED?

The third question addressed is:

Is a special framework needed for thinking about theorizing in the Sciences of the Artificial?

It is argued that a special framework is needed for thinking about theorizing in the Sciences of the Artificial, to clearly distinguish it from other types of science and also to illustrate two general modes of research activity and theorizing that are obvious when one examines the research activities that occur. These two modes are termed the interior and exterior modes. The interior mode is where theorizing is done to produce theory for design and action, with prescriptive statements about how artifacts can be designed, developed and brought into being. The exterior mode, includes other types of theory, which aim primarily at analysing, describing and predicting what happens as artifacts exist and are used in their external environment. Depicting these two modes as separate but linked allows us to study them individually but also in relation to one another.

Figure 2 gives a picture of this framework and its chief idea is expressed in a further principle.

Principle 5: Theorizing is done in both interior and exterior modes.

This division into two modes follows Simon's ([37] p. 7) insights into systems complexity where the division between the inner and outer environment "is highly convenient". The separation of the inner from the outer allows the simplification of tasks. The detail of the inner environment can be hidden when we talk about an artifact attaining its goals and only minimal assumptions may need to be made about the inner environment, as the same end may be achieved by different mechanisms. It is this principle which underlies the mastering of complexity by decomposing hierarchically ordered systems into sub-systems where details of the sub-systems operation are "hidden". However Simon appears not to think that theorizing methods in the exterior mode are different in the artificial sciences. His view is that:

Given an airplane, or given a bird, we can analyse them by the methods of natural science without any particular attention to purpose or adaptation, without reference to the interface between what I have called the inner and outer environments. (p. 7).

It is here that respectfully we disagree with Simon. As argued under Principles 1 and 2, in IT if we are studying artifacts in the exterior mode, our theorizing is improved if we consider a special feature of artifacts, namely their goal-directedness, and we focus on design features that are linked to the achievement of goals.

These two modes of theorizing are seen as "two sides of a coin". Both are needed in the design disciplines of IT and they are complementary to each other. Theory in the exterior mode can include propositions such as "A system with feature X will perform better on measure M than a system without feature X". If empirical testing shows that this proposition is supported then the proposition can be "turned around" or **inverted** to give a design proposition: "If you want to achieve M then include feature X".

It helps to distinguish the two modes in terms of theorizing as they involve different activities and different ways of thinking about theory. Moreover, a single piece of research in a journal article or thesis is likely to include research conducted in either one mode or the other and it is well to consider what is regarded as acceptable theorizing in each mode. A researcher who has devoted a great deal of effort to showing how a new artifact can work, and in showing that it is a novel artifact, might not have the time and effort to do the comprehensive evaluation that would be expected if he/she was working in the exterior mode assessing an artifact constructed by someone else. The IT design communities would do well to reflect and define their expectations in this respect.

5. SPECIAL CHALLENGES?

The fourth question addressed is:

Do the sciences of the artificial present special challenges for theorizing?

The answer here is "yes" and the challenges are many and varied. In part these challenges stem from the relative newness of the disciplines concerned with complex artifacts, but also to their complexity itself and their rapid evolution.

To take just one of these artifacts, the Web, Hendler et al., ([20] p. 68) note that:

The Web is different from most previously studied systems in that it is changing at a rate that may be of the same order as, or perhaps greater than, even the most knowledgeable researcher's ability to observe it...Web scientists need new methodologies for

gathering evidence and finding ways to anticipate how human behaviour will affect development of a system that is evolving at such as amazing rate.
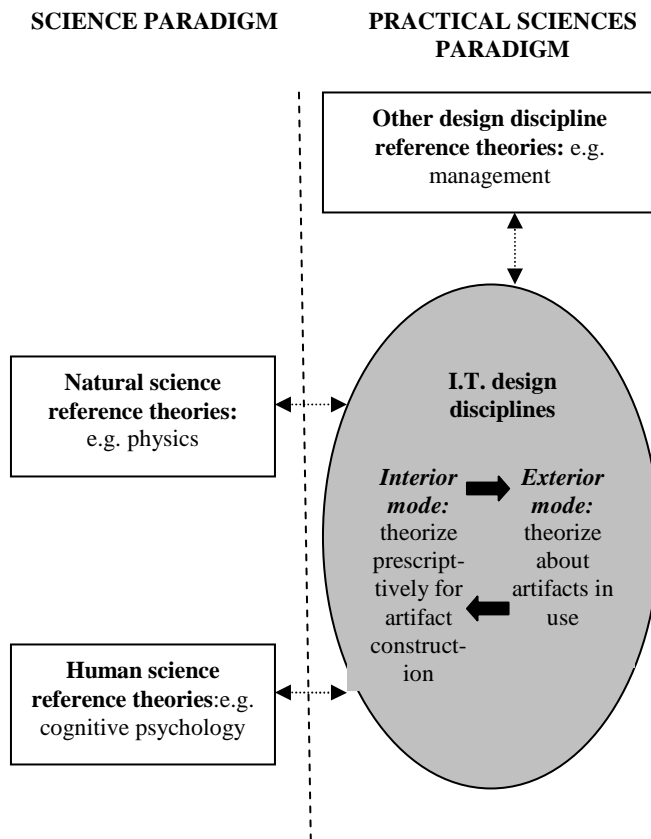
SCIENCE PARADIGM   **PRACTICAL SCIENCES PARADIGM**

Figure 2: A framework for theory building in IT as a science of the artificial

The challenges for theorizing are so extensive that they can only be touched upon here. A number of challenges relate to the degree to which generalizations can be made about artifacts. The goal of high-level "grand" theories with wide applicability in IT disciplines is illusive. Perhaps the most likely prospects remain general systems theory and its derivatives, which provide such general statements as the Law of Requisite Variety: only variety in a system's responses can keep down variety in outcomes when the system is subjected to a set of disturbances [1]. Weber gives a generalized theory of representation, which aims to model the desirable properties of information systems at a deep level and be a theory native to IS [47]. Apart from these few instances, the imprecise nature of knowledge relating to the behaviour of humans interacting with technology and the changing nature of technologies themselves suggests that highly generalized, unchanging laws as can be found in science will be rare in IT design disciplines.

Some authors have advocated an intermediate position of mid-range theory: theory that has moderate coverage and can easily lead to testable hypotheses. Mid-range theory is seen as particularly important for practice disciplines [30], such as IS. Similarly, Van Aken sees that technological rules are not general knowledge, but rather mid-range theories of practice. "They are only valid for a certain application domain, a range of settings that have key attributes in common with the settings in which the rules were developed and tested" ([42] p. 238).

The foregoing leads to the final principle:

Principle 7: Complexity in IT disciplines entails problems of generality in theorizing

There are other challenges. For example, problems of moving from descriptive to prescriptive statements are explored in some depth by Edgley, who distinguishes between "What is the case?" (questions of fact and science) and "What is to be done?" (questions of action) [11]. Edgley shows that prescriptive statements about action do not necessarily follow from descriptive knowledge. To extrapolate, imperative statements about design and action cannot be deduced from scientific knowledge. For example, as a matter of science one might know that an insecticide kills an insect pest. One cannot deduce from this knowledge that "in order to kill the insect pest one must use insecticide": there may be other ways of killing the pest. This interesting point has implications for the degree and manner in which theorizing in the exterior mode informs artifact construction and theorizing in the interior mode. However, it is not pursued further at this point.

6. CONCLUDING REMARKS

This essay addresses complex issues, those of knowledge creation and theory building, and it does so from the point of view of the sciences of the artificial (alternatively referred to as practical sciences or design disciplines), a perspective which has been little considered. Thus, the arguments made are advanced for discussion and to further debate, rather than being regarded as fixed and certain ideas. There are many new ideas appearing in the literature of the artificial sciences, which as a study of the design of complex artifacts dates back only for about 50-60 years. It is expected that it will be some time before a common understanding and means of describing our problems emerge. This essay is just one step along the path.

The essay was framed around a number of fundamental questions in order to place theorizing in the artificial sciences against the broad tradition of the philosophy of science. Addressing these questions has led to a general framework (Figure 2) within which a number of approaches to theorizing can be seen as complementary and to a number of high-level principles for theorizing (see Table 1).

In answer to our first driving question, it is concluded that the artificial sciences are a third broad form of science in addition to the natural and human sciences. The artificial sciences can be regarded as a science because they use scientific-type methods of observation and experiment to test theory. It is sensible to distinguish this third category, however, because it has features that make theorizing different from other categories of science. Artificial science theory concerns artifacts that are constructed with some purpose and in theorizing it is needful to study the relationship between the means of achieving some outcome and the outcome which is achieved. A special type of theory with prescriptive statements, design theory, is needed to capture the

knowledge developed. Principles 1, 2 and 3 resulted from these ideas.

The second question asked whether the artificial sciences followed the scientific method. The conclusion is "yes, but with some special features". First, it was argued that the artificial sciences require the acceptance of inductive methods in theory building, which is in accordance with some streams of thought amongst philosophers of science, but not all (Principle 4). Further, the processes of theory building can involve steps which do not normally belong in any of the other categories of science: namely the construction of the object of study by the researcher/theorizer or a practitioner (Principle 5). Theory building around artifact construction provides for more sophisticated forms of inductive reasoning.

The third question asked whether a special framework is needed for theorizing in an artificial science. The suggestion is "yes", and following Simon an interior and exterior mode of theorizing are distinguished (Principle 6). Unlike Simon, however, theorizing in the exterior mode is not seen as equivalent to normal natural science theorizing, because, as argued under Principles 1,2 and 3, the objects of study are artifacts and need to be treated differently in theorizing. Design type theory is seen as the ultimate aim of research in the artificial sciences and work in the exterior mode should inform the design theory generated in the interior mode.

The fourth question asked whether there are special challenges for theorizing in the artificial science. The answer again is "yes". The complexity and rate of change of modern technologies stretch our abilities to formulate general and high-level theory. The theory that is developed is liable to be limited in its scope and in the level of generalization possible, and also in the time it retains currency (Principle 7).

A limitation of the paper is that it does not cover in detail questions of theory building and research methods. Thus the advantages and limitations of different research approaches are not discussed. Further, at heart the essay is a result of analytic reasoning; the principles derived are arrived at by logical extension from the defining characteristics of IT and its nature as a design discipline.

The contribution of the paper is that it provides a high-level framework for thinking about the different types of theorizing in the paradigms of natural science versus that of design. Many of the ideas in the paper stem from Simon's seminal work, yet there are differences, particularly in the identification of the need for inductive logic. Further, Simon did not give much thought to the actual process of theorizing and both he and subsequent writers have not explicated the special features of theorizing in the exterior mode in a design discipline.

The principles are in part conjectures. In future work it would be valuable to explore whether the principles appear, even implicitly, in what could be regarded as "good" theorising or seminal work in IS/IT. Further work could also usefully be done to make the study of the processes of theory building more transparent. Space precludes such examination in this essay.

7. ACKNOWLEDGMENTS

8. REFERENCES

[1] Ashby, W.R. 1956 An Introduction to Cybernetics. Chapman & Hall.

[2] Atwood, M., McCain, K. and Williams, J. 2002. How does the design community think about design? Symposium on Designing Interactive Systems 2002: 125-132

[3] Benbasat, I. and Zmud, R. 2003, The identity crisis within the IS Discipline: Defining and communicating the discipline's core properties", MIS Quarterly, 27, 2, 183-194.

[4] Bourgeois, L. 1979.Toward a method of middle-range theorizing, Academy of Management Review, 4, 443-447.

[5] Bunge, M. 1967 Scientific Research II: The Search for Truth. Springer Verlag.

[6] Bunge, M. 2003. Philosophical inputs and outputs of technology. In Philosophy of Technology The Technological Condition, R. Schaff and V. Dusek, Eds. Blackwell Publishing.

[7] Butts, R. 1999. Hypothetico-deductive method. In The Cambridge Dictionary of Philosophy, R. Audi, Ed. Cambridge University Press.

[8] Cross, N. 2001. Designerly ways of knowing: Design discipline versus design science. Design Issues, 17, 3, 49-55.

[9] Davenport, T. and Short, J. 1990. The new industrial engineering: Information technology and business process redesign. Sloan Management Review, 31, 4, 11-27.

[10] Dilthey, W. 1883 Introduction to the Human Sciences. (See online Stanford Encyclopaedia of Philosophy).

[11] Edgley, R. 1969 Reason in Theory and Practice. Hutchinson University Library.

[12] Glaser, B., and Strauss, A. 1967 The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine Publishing Company.

[13] Godfrey-Smith, P. 2003 Theory and Reality. University of Chicago Press.

[14] Gregor, S. 2006. The nature of theory in Information Systems. MIS Quarterly, 3, 30, 611-642.

[15] Gregor, S. and Jones, D. 2007. The anatomy of a design theory. Journal of the Association of Information Systems, 8, 5, Article 2, 312-335.

[16] Gribbin, J. 2003. Science a History 1543-2001. Allen Lane The Penguin Press.

[17] Goldkuhl, G. 2004. Design Theories in Information Systems – a Need for Multi-grounding, Journal of Information Technology Theory and Application, 6, 2, 59-72.

[18] Hatfield, G. 2005. Scientific method. In The Shorter Routledge Encyclopedia of Philosophy. Routledge.

[19] Heidegger, M. 1993. The question concerning technology. In Basic Writings, San Francisco: Harper, pp. 311 -341 (Translated from Martin Heidegger (1954) Vortrage and Aufsatze, Gunther Neske Verlag, Pfullingen, pp. 13-44.

[20] Hendler, J., Shadbolt, N., Hall, W., Berners-Lee, T. and Weitzner, D. 2008. Web science: Approach to understanding the Web. Comm. Of the ACM, July, 51, 7, 60-69.

[21] Hevner, A., March, S., Park, J. and Ram, S. 2004. Design science in information systems research, MIS Quarterly, 28, 1, 75-105.

[22] Heylighen,F. and Joslyn, C. 1992. Systems theory. In F. Heylighen, C. Joslyn and V. Turchin (editors): Principia Cybernetica Web (Principia Cybernetica, Brussels), URL: http://pespmc1.vub.ac.be/systheor.html (Accessed August, 2008.)

[23] Hooker, R. 1996. Aristotle: The four causes-Physics II.3, http://www.wsu.edu/~dee/GREECE/4CAUSES.HTM (Accessed January, 2009)

[24] Iivari, J. 2003. Towards information systems as a science of meta-artifacts, Communications of the AIS, 12, 37, Nov., 568-581.

[25] Iversen, J., Mathiassen, and Nielsen, P. 2004. Managing process risk in software process Improvement: An action research approach. MIS Quarterly, 28, 3, 395-434.

[26] Jarvinen P. 2007. On reviewing results of design research. In Proceedings of ECIS 2007, St Gallen June 7-9.

[27] Kuechler, W. and Vaishnavi, V. 2008. The emergence of design science research in Information Systems in North America, Journal of Design Research. 7, 1,1–16.

[28] Makkreel, R. 2009. Wilhelm Dilthey, The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/dilthey/>. (Accessed November, 2008)

[29] March, S. and Smith, G. 1995. Design and natural science research on Information Technology. Decision Support Systems, 15, 251-266.

[30] Merton, R. 1968. On Theoretical Sociology. New York: Free Press.

[31] Nagel, E. 1979. The Structure of Science. Hackett Publishing Co.

[32] Orlikowski, W. and Iacono, C. 2001. Research commentary: Desperately seeking the "IT" in IT Research – A Call to Theorizing the IT Artifact, Information Systems Research, 12, 2, 121-134.

[33] Parry, R. 2007. Episteme and Techne. In Stanford Encyclopedia of Philosophy. URL: http://plato.stanford.edu/entries/episteme-techne/ Accessed November, 2007).

[34] Popper, K. 1980. The Logic of Scientific Discovery. Unwin Hyman.

[35] Scharff, R. and Dusek, V. Eds. 2003. Philosophy of Technology: The Technological Condition: An Anthology. Blackwell Publishing.

[36] Schön, D. 1983 The Reflective Practitioner. Basic Books.

[37] Simon, H. 1996 The sciences of the artificial, 3rd ed. MIT Press.

[38] Snyder, L. 2008 William Whewell. The Stanford Encyclopedia of Philosophy (Fall 2008 Edition), Edward N. Zalta (ed.), URL = http://plato.stanford.edu/archives/fall2008/entries/whewell/ (Accessed November, 2008)

[39] Strasser, S. 1985 Understanding and Explanation Basic Ideas Concerning the Humanity of the Human Sciences. Duquesne University Press.

[40] University Libraries (2008). Real-Time EBM Program: The P I C O Model. URL: http://www4.umdnj.edu/camlbweb/ebm/picomodel.htm (Accessed August, 2008)

[41] Van Aken, J. 2004. Management research based on the paradigm of the design sciences: The quest for field-tested and grounded technological rules. Journal of Management Studies, 41, 2, 219-246.

[42] Van Aken, J. 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.

[43] Venable, J. 2006. "The Role of Theory and Theorizing in Design Science Research", First International Conference on Design Science Research in Information Systems and Technology, Claremont, California, pp. 1-18.

[44] von Bertalanffy, L. 1968 General System Theory Foundations, Development, Applications New York: George Braziller

[45] Vogel, J. 2005. Inference to the best explanation. In The Shorter Routledge Encyclopaedia of Philosophy, E. Craig, Ed. Routledge.

[46] Weber, R. 1987. Toward a theory of artifacts: A paradigmatic base for Information Systems research. Journal of Information Systems, 1,1, Spring, 3-19.

[47] Weber, R. 1997 Ontological Foundations of Information Systems. Coopers & Lybrand.

[48] Weick, K. 1989. Theory construction as disciplined imagination, Academy of Management Review, 14, 4, 516-531.

[49] Whewell, W. 1857. A History of the Inductive Sciences. (Parts I, II, III). 3rd edn.Thomas Nelson. (Reproduced in facsimile, 1967)