IS RESEARCH PROGRESS WOULD BENEFIT FROM INCREASED FALSIFICATION OF EXISTING THEORIES

Complete Research

Salovaara, Antti, Aalto University School of Business, Finland, antti.salovaara@aalto.fi
Jani Merikivi, Aalto University School of Business, Finland, and Huazhong Normal University, Wuhan, Republic of China, jani.merikivi@aalto.fi

Abstract
In empirical sciences, one of the best-known measures for a theory’s strength is its falsifiability. This principle, originally introduced by philosopher Karl Popper (1902–1994), holds that good theories make bold and empirically testable claims that survive repeated attempts of falsification, i.e., attempts to prove that a theory is invalid. According to Popper, scientific progress requires provisional falsifiable theories and their refutations that show where the existing theories need to be corrected. This is not, however, how majority of research in Information Systems Science appears to operate. Instead, much research follows an inductivist approach where researchers attempt to extend theories to new domains and obtain positive empirical confirmation. Such research, however, is weaker than falsification in terms of validation. We exemplify this research practice by tracing the history of IS use model development and by presenting examples of studies that suggest how falsification can be applied in IS research to examine existing theories’ boundary conditions. We summarize this essay by suggesting how falsification can be integrated fruitfully into replication and comparison studies.

Keywords: Popper, Falsification, Scientific progress, IS use, Boundary conditions

1 Introduction
Over the years, the Information Systems (IS) research community has gradually come to an agreement that several different research traditions have a legitimate place within IS research. These include not only the positivist, traditionally quantitative research tradition, but also the interpretivist, traditionally qualitative tradition (Walsham, 1995) and action research (Baskerville and Wood-Harper, 1996), among others. While most researchers within the IS community tend to identify with some traditions more than others, the traditions do not operate in silos but interact with each other with a goal of producing higher-quality theories and findings about the relationships between humans, technologies and organizations.

The traditions therefore share the same methodological question: what are the research methods and approaches that will best support the advancement of scientific knowledge within the shared discipline. While the practical applicability of the findings is also crucial, scientifically the question of progress relates to the theories that the field strives to develop. Progress can manifest in many ways: new theories may account for phenomena that have not been previously addressed, theories may replace old ones, and they may be repaired.

In this paper we argue that the progress within IS science rests too much on the induction-based approach in which the goal is to develop models with increasing fit with the reality, and too little on existing models’ correction, refutation, and critical examination that may give impetus to the development of new models. Following the terminology in use in the philosophy of science, we call the two
approaches *inductivist* and *falsificationist* research traditions and approaches. While the inductivist tradition has a longer history, reaching back to the 16th century, falsificationist tradition is less than a hundred years old, but defines much of the research carried out nowadays especially in natural sciences.

In this paper we first look at both research approaches. We identify the known problems in inductivist research and then present the benefits of falsificationism, followed by an encouragement to increase that kind of research within IS science. Even though the writings on the philosophy of science on which we will base our argumentation have paid more attention to quantitative research, the discussion presented in this paper is applicable both to quantitative and qualitative research in IS science.

### 1.1 Traditional inductivist view on the scientific method

Examination and discovery of the laws of nature has been a central part of scientific inquiry since the ancient times, and gained increasing importance when Galilei and Newton made their discoveries on gravity and the laws of classical mechanics. The method by which the nature should be studied to produce reliable laws became an important problem. Francis Bacon (1561–1626) has been credited as the person who first described the *method of induction* by which discoveries can be made. In short, the method consists of careful observations that are generalized as laws that the scientist then tries to confirm with more observations. Bryan Magee (1973, p. 56) summarized this method as consisting of six steps: 1. observation and experiment; 2. inductive generalization; 3. hypothesis; 4. attempted verification of hypothesis; 5. proof or disproof; 6. knowledge. This process is followed in much of the scientific research also today.

The inductive method emphasizes generation of theories of increasingly accurate knowledge about the reality. However, it has a logical flaw whose formulation has been credited to David Hume (1711–1776): no amount of confirmatory evidence will give us a full certainty that a proposition will be true. This criticism, commonly known as Hume’s problem or the fallacy of affirming the consequent, is often exemplified with an example of black swans. No matter how many white swans are observed these observations will not ensure that all swans are white. Hume’s observation was disconcerting to the scientists who wanted to demarcate their knowledge from mere subjective opinions, lay knowledge, superstition and religious beliefs. The question of demarcation between scientific and other kind of knowledge became an ongoing challenge for scientists.

Finally, in the beginning of the 20th century, the logical positivists in the Vienna circle developed a new solution to the demarcation problem. They wanted to apply the newly discovered methods of logic to natural science, and this way provide a general method that removes all ambiguous and empirically non-verifiable statements from the scientific language. They maintained that scientific knowledge could only be based on direct observations and conclusions that can be deduced from them through logical, deductive reasoning (e.g., Ackermann, 1976). If a theory, stated in a manner that fulfilled these criteria, was verified with confirming empirical observations, it was part of scientific knowledge. This requirement initially served as a means to demarcate scientific knowledge from other beliefs. However, its inability to address Hume’s problem caused a failure in demarcation: its principles had an implication that generalized laws would be considered non-scientific. The problem was that general laws apply to infinite number of physical instances but there can always be only a finite number of observations. Because general laws could not be proven empirically, they could not be part of science (ibid, p. 9). Because of problems such as these, logical positivism eventually lost much of its appeal in philosophy of science.

### 1.2 Falsification as a scientific method

Karl Popper was one of the critics of logical positivism and contributed to its eventual rejection. Popper accepted that Hume’s problem caused a failure in demarcation: its principles had an implication that generalized laws would be considered non-scientific. The problem was that general laws apply to infinite number of physical instances but there can always be only a finite number of observations. Because general laws could not be proven empirically, they could not be part of science (ibid, p. 9). Because of problems such as these, logical positivism eventually lost much of its appeal in philosophy of science.
scriptive: it also presented a method that science should follow. Popper maintained that although confirmatory evidence never completely verifies a theory, a finite number of negative observations can falsify a theory. Science should therefore aim at laws that can be tested empirically and that will survive repeated attempts to gather negative evidence that could falsify them. The best theories are those that appear easy to falsify (e.g., ones that provide surprising explanations or predictions) yet gain corroborating evidence, that is, survive repeated empirical attempts that seek to refute them. This definition was also Popper’s answer to the demarcation problem: theories should be falsifiable—formulated in a manner that provides a means for empirical examination that may lead to the theory’s refutation. Unfalsifiable claims, in turn, would not be considered scientific.

Magee’s summary of falsificationist method was following (1973, p. 56): 1. problem (usually rebuff to existing theory or expectation); 2. proposed solution, in other words a new theory; 3. deduction of testable propositions from the new theory; 4. tests, i.e. attempted refutations by, among other things (but not only among other things), observation and experiment; 5. preference established between competing theories. Lee (1989), in a similar manner, has summarized the falsificationist research with four requirements: consideration of predictions that would prove a theory being wrong, internal consistency of those predictions, corroboration and confirmation of the predictions with empirical data, and ruling out the rival theories. Gregor (2006) additionally notes the requirements of explaining or predicting phenomena when discussing this type of research.

Falsificationism provided a new viewpoint on the scientific progress. In this view, scientific progress, following the falsificationist method, emerges from propositions of theories, their refutations, and replacement with corrected theories that better explain the phenomena of interest. Falsification is therefore a productive method and accelerates the growth of knowledge. It helps scientists improve their theories by pointing out those theories and their aspects that are empirically problematic. It also motivates the necessity of searching for existing theories’ boundary conditions by encouraging researchers to examine domains where theories may not be valid. Clarification of existing theories’ boundaries reveals “gaps” for which new theories can be proposed.

Today, Popper’s falsificationism is often considered as an ideal that is not, and should not be, always followed in its naïve form. If interpreted superficially, Popper seemed to claim that falsifying evidence would be infallible: that it should always be trusted. However, similarly with confirming evidence, falsifying evidence is, in fact, also fallible: it may, for instance, inadvertently focus on a non-critical aspect of a theory or be a result of a faulty empirical research design. The naïve approach would also fit poorly to probabilistic research (e.g., statistical testing) where an arrival at a false conclusion is an admitted possibility.

To demonstrate falsificationism’s deeper importance, Lakatos (1970) developed Popper’s ideas further and showed how falsificationism has played a part in actual success stories in the history of science. Lakatos presented sophisticated falsificationism that moved the attention away from individual theoretical claims to series of theories. These series, called research programmes, mature over time following a progressive logic by which theories expand to address larger scopes of empirical content. In so doing, they may also integrate in themselves the falsifying findings that emerge as the evolving theory is tested. Because of such programmes’ evolving structure, it becomes clear that a single falsifying finding cannot refute an entire programme. Only an entire programme can falsify another programme. For that it must explain all the content of its rival as well as predict novel findings that the rival deems improbable or impossible (p. 116). Such a falsification is never conclusive, however, since the rival programme may survive from the falsification by adjusting its theory.

One must note that Lakatos does not discard the principle of falsificationism itself. Rather, he proposes changes to how it should be adopted. He allows research programmes to sidestep falsifying evidence temporarily in two ways: by redefining their theories’ auxiliary conditions (in Lakatos’s terminology, their “protective belts”) or by postponing adjustments by promising that newer theory versions will eventually explain also the anomalies. However, sooner or later all the falsifying evidence must be addressed and the theory must be made compatible with it.
Lakatos’s view can be compared with that of Thomas Kuhn (1962) who introduced the concepts of normal science and scientific revolutions to the philosophy of science. In Kuhn’s view, much of science resembles inductivist research that is called normal science: application of widely accepted theories and ways of doing research (i.e., paradigms) to new phenomena, so as to generate new scientific knowledge, however not necessarily new scientific theories. Falsificationism emerges in Kuhn’s thinking only occasionally, when established theories are replaced with new ones in scientific revolutions (e.g., Magee, 1973, p. 41). However, Kuhn’s theory is only descriptive and does not therefore justify why scientific progress should proceed through distinct revolutions. Lakatos’s view, in contrast, is prescriptive: it recommends that theories are continuously challenged, enabling the research programmes to improve themselves and compete against each other. This may sometimes result in large-scale revolutions, but a gradual transformation is also possible. Without falsificationist research, when faltering to inductivist research only, scientific progress halts (Lakatos, 1970, pp. 179).

2 The Nature of Research in IS Science

To our knowledge, the philosophy-oriented publications to date in IS science have not paid exclusive attention to the distinction between induction and falsification in IS research. Many reviews have focused on the positivist–interpretivist distinction, however. In these papers, it has been common to contend that positivistic/interpretivist research tradition must not be regarded as the only valid model of scientific research in IS science (e.g., Chen and Hirschheim, 2004; Klein and Lyytinen, 1985). Another group of papers has charted the various research designs (e.g., case studies, experiments, etc.) in IS research (Chen and Hirschheim, 2004; Farhoomand, 1987). Of these, Farhoomand (1987) introduces Popper’s and Kuhn’s models of science but applies only Kuhn in an analysis of the IS research tradition. Lyytinen and King (2004) present Popper’s falsifiability criterion as a possible (but imperfect) means by which IS science has not been able to identify a core theory around which the field could based itself. Since the paper’s focus is on the field’s identity as a research enterprise, the use of falsification as a research approach is not in the focus of that paper.

The writings by Allen Lee and his colleagues (1989, 2004, 2009) are closest to our focus. In his first paper, Lee (1989) discusses the scientific rigour in case study research and concludes that also theories derived from single case studies may fulfil the criteria of scientific theory. Most importantly, replicability can be achieved by carefully specifying the conditions within which the researcher wishes to apply and test an existing theory. Presenting a single case study with those conditions enables later replication by others. The conditions do not need to be exactly the same as in the original theory that is being applied: they only need to be in the scope for which the original theory has been promised to be generalizable. These conditions can be chosen critically, this way allowing for partial refutation of the original theory. In the second paper, Lee (2004) discusses the general benefits that knowledge of philosophy of science can bring to IS researchers. He expresses his puzzlement of the persistence of traditional, inductive positivism in IS research and suggests that the reason may lie in the general lack of awareness of the criticisms of logical positivism (pp. 15–17). The third paper (Lee and Hubona, 2009), in turn, develops a general scientific model for both quantitative and qualitative IS research that follows the falsificationist logic and owes significantly to Popper’s thinking. Lee and Hubona explain with examples how its logic can be applied also to interpretivist reasoning.

Finally, Shirley Gregor’s (2006) typology is probably the best known philosophy paper in IS science. It attends to the purposes of theories within IS science. Gregor identifies the following uses: 1) analysing (such as classification), 2) explanation, 3) prediction (without an attempt for providing an explanatory mechanism behind the phenomenon), 4) explanation and prediction, and 5) theory for design and action. Gregor’s classification is orthogonal to ours: independently of the types of use that a researcher adopts in Gregor’s typology, the researcher can choose either an inductivist or a falsificationist research orientation.

In the present paper, similarly with Lee (1989), we apply concepts that traditionally have been associated with only positivist research and extend them to a more general scope of IS research. In our case
this concerns the definitions on inductivist and falsificationist methodology. By *inductive research* we refer to studies that present confirmatory evidence to a theory, either in a form of replication, application to a new domain, or through extension or synthesis of several theories. By *falsificationist research* we refer to two kinds of studies. They may present refuting evidence and problems in one or more theories, most often through replications purposely designed for generating negative evidence or through empirical comparisons between competing theories. Such studies contribute to science by showing limits in the existing theories, by arguing with empirical evidence where new theories are needed and by identifying theories that are stronger than others. Alternatively, they may also present new theories that explain or predict phenomena that their rivals deem improbable.

Although this distinction seems most readily applicable to quantitative research, it can be applied also to qualitative research. Also qualitative and interpretivist papers can be both inductive and falsificationist. Inductive qualitative papers adopt theories and present analyses on data that are aligned with the adopted theory. For example, several authors have applied structuration theory to information systems research (DeSanctis and Poole, 1994; Orlikowski, 1992). Falsificationist qualitative research, in contrast, uses empirical data to criticize existing research (Lee, 1989). In fact, qualitative research in many fields has a strong legacy of such theory critique, also in IS science.

Our opinion is, pending decisive confirmation, that falsification is rare in IS science. Very rarely a new theory is introduced as a direct competitor to an existing theory, or empirical data is gathered with an intention to better understand weaknesses in such theories. In other words, most of the theoretical work within IS science follows the inductive logic rather than falsificationist logic.

Treating falsificationism without due regard is disputable. This is because it poses a threat to the way of doing science. For example, if recently arrived theories do not go through critical examination, the risk is that the scientific knowledge is built on a shaky foundation. A higher percentage of falsificationist research would ensure that the flaws in the theories are corrected before the models and theories become widely adopted.

In the following section, we offer a brief review of IS acceptance and use research to evaluate the importance of falsificationist research in IS science. We chose IS acceptance and use as our example because of three reasons. First, it has preoccupied many IS researchers for several decades and is therefore a significant research domain. Besides its long tradition, it has received the most attention in IS use research, and hence, stands as a useful example for our purpose. Second, this vast field of research is based on a rather small number of established theories, and is therefore worth of a critical look. Third, our own research, presented later in the paper, is closely related to this research tradition.

We also considered alternative research traditions for our review, particularly publications on IS success (e.g., DeLone and McLean, 1992, and related studies) and organizational communication (e.g., Daft and Lengel’s media richness theory, 1986). The available space did not unfortunately allow for their inclusion in the review.

### 2.1 Falsificationism in the research on IS use

We sampled 177 studies from four extant meta-analyses on TAM research (King and He, 2006; Lee et al., 2003; Legris et al., 2003; Schepers and Wetzel, 2007). Of these, we first selected 30 empirical studies with the highest citation scores in Scopus. The papers in this sample accounted for 70% of the citations of the whole 177-paper sample. We tested this sub-sample against the sample that we would have got had we gathered 30 most cited papers from Scopus using keywords (e.g., “TAM” and “technology adoption”). The keyword-based sample contained much less cited papers, indicating that our decision to base the sampling on meta-analyses was superior to a keyword-based search. We did two changes to our initial top-30 sample: we removed Legris et al.’s (2003) meta-analysis because of its non-empirical content, and added Davis et al.’s (1989) paper that first introduced TAM in IS research. For reasons unknown to us this paper was not indexed in Scopus. We also validated the sample by creating a competing sample of top-30 papers using Web of Science. We found only one difference:
the Scopus-based sample contained a paper by Rai et al. (2002) while the competing sample had a paper by Dabhoklar (1996). We included the both in the analysis, resulting in a 31-paper final sample. We analysed each paper to learn whether its purpose was to extend an existing theory, suggest a new theory, apply it, or compare theories. Because papers may have both inductivist (confirmatory) and falsificationist findings, we gathered them separately, focusing on those aspects that the authors highlighted as their main findings. Based on this information, we classified each paper into one of the following types. We use the word “theory” to denote both theories and models:

• **Confirmatory** (18 papers): The authors present a new theory or an extension and provide evidence that supports their hypotheses. Alternatively, the authors apply a theory to a new domain where they expect to find confirmatory evidence to the theory.

• **Comparative + confirmatory** (6 papers): The authors compare existing theories of which some may be novel or contain novel elements. In the evaluation, they test which of the theories gets the highest confirmation in a given IS use context.

• **Partly falsifying** (5 papers): As in confirmatory papers, also here the authors present a new theory or an extension. In addition to confirmatory evidence they also highlight findings where expected relationships do not hold and argue why this finding is valid.

• **Falsifying** (1 paper): The authors test a theory in a use context which, in their expectation, will not provide confirmatory evidence to some or all of the elements of the theory. The falsifying attempt is included in the research design of the study from its very start.

• **Not applicable (n/a)** (1 paper): The theory is applied as a tool to learn about aspects in IS use. The purpose is not to evaluate or improve the theory.

Since two researchers took part in coding the studies we investigated the potential for coding bias. Before the actual procedure, we made sure we understood and agreed upon the evaluation standards by categorizing several example studies with each other. The actual categorization was performed separately. After the categorization we began our verification process. We passed majority of the studies as we both agreed on them. The studies with no agreement underwent a secondary analysis. This meant carefully justifying our categorization decision to each other. Nearly all remaining studies passed this stage. The ones that still lacked an agreement were categorized by the first author.

The results are presented in Table 1. They show that falsificationist research is rare among the most cited papers about IS use. Only Adams et al.’s study (1992) that examined TAM’s predictive power in several contexts has been falsifying by its nature on the very outset. The falsificationist goal is evident in the research design: Adams and colleagues did not expect to receive empirical support for TAM in all the contexts studied. The paper also presents several reservations related to TAM’s validity.

Outside of the top-31 papers some falsifications can be found, however. Many of these papers are focused on falsifications of research designs, instead of theories, by showing problems in methods’ validity. Marcolin et al. (2000) point out problems in measurements of users’ technical IS competence via self-report vs. objective measures. Straub et al. (1995) report that perceived usefulness (PU) and perceived ease of use (PEOU) predict only self-reported IS use but not objective IS use. They however do not use their result as evidence against TAM. Instead they provide several explanations why empirical evidence does not falsify TAMs but represents other kinds of phenomena.

Meta-analyses and conceptual papers on IS use theories are more openly critical. All of the 4 meta-analyses used when composing the top-31 sample present findings that could later be studied empirically and published as falsifications. Papers in JAIS special issue in 2007 criticised TAMs for non-falsifiability (Silva, 2007; see also Greve, 2001) and for intention–behavior gap, poor conceptualization of IS use and disregard of several potentially significant factors (Bagozzi, 2007). Seddon (1997) identified structural problems in DeLone and McLean’s (1992) IS success model and suggested reme-
### Table 1. Inductivist and falsificationist findings in the 31 most cited IS use papers.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>IS Research Would Benefit from Falsification Paper type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davis, MISQ 1989</td>
<td>9063</td>
<td>(1)</td>
</tr>
<tr>
<td>Davis et al., MgmtSci 1989</td>
<td>522</td>
<td>(2)</td>
</tr>
<tr>
<td>Venkatesh et al., MgmtSci 2003</td>
<td>4760</td>
<td>(3)</td>
</tr>
<tr>
<td>Venkatesh &amp; Davis, MgmtSci 2000</td>
<td>3328</td>
<td>(4)</td>
</tr>
<tr>
<td>Moore &amp; Benbasat ISR 1991</td>
<td>2374</td>
<td>(5)</td>
</tr>
<tr>
<td>Taylor &amp; Todd, ISR 1995</td>
<td>2181</td>
<td>(6)</td>
</tr>
<tr>
<td>Gefen et al., MISQ 2003</td>
<td>1726</td>
<td>(7)</td>
</tr>
<tr>
<td>Venkatesh, ISR 2000</td>
<td>1398</td>
<td>(8)</td>
</tr>
<tr>
<td>Mathieson, ISR 1991</td>
<td>1250</td>
<td>(9)</td>
</tr>
<tr>
<td>Bhattacherjee, MISQ 2001</td>
<td>1161</td>
<td>(10)</td>
</tr>
<tr>
<td>Venkatesh &amp; Morris, MISQ 2000</td>
<td>1131</td>
<td>(11)</td>
</tr>
<tr>
<td>Karahanna et al., MISQ 1999</td>
<td>1114</td>
<td>(12)</td>
</tr>
<tr>
<td>Pavlou, IJEC 2003</td>
<td>1032</td>
<td>(13)</td>
</tr>
<tr>
<td>Agarwal &amp; Karahanna, MISQ 2000</td>
<td>1013</td>
<td>(14)</td>
</tr>
<tr>
<td>Adams et al., MISQ 1994</td>
<td>1012</td>
<td>(15)</td>
</tr>
<tr>
<td>Moon &amp; Kim, I&amp;M 2000</td>
<td>988</td>
<td>(16)</td>
</tr>
<tr>
<td>Davis, IJMS 1993</td>
<td>986</td>
<td>(17)</td>
</tr>
<tr>
<td>Venkatesh &amp; Davis, DecSci 1996</td>
<td>939</td>
<td>(18)</td>
</tr>
<tr>
<td>Kourafas, ISR 2002</td>
<td>832</td>
<td>(19)</td>
</tr>
<tr>
<td>Thompson et al., MISQ 1991</td>
<td>829</td>
<td>(20)</td>
</tr>
<tr>
<td>Taylor &amp; Todd, MISQ 1995</td>
<td>757</td>
<td>(21)</td>
</tr>
<tr>
<td>Gefen &amp; Straub, MISQ 1997</td>
<td>746</td>
<td>(22)</td>
</tr>
<tr>
<td>Agarwal &amp; Prasad, DecSci 1999</td>
<td>694</td>
<td>(23)</td>
</tr>
<tr>
<td>Agarwal &amp; Prasad, ISR 1998</td>
<td>667</td>
<td>(24)</td>
</tr>
<tr>
<td>Hu et al., JIMS 1999</td>
<td>581</td>
<td>(25)</td>
</tr>
<tr>
<td>Szajna, MgmtSci 1997</td>
<td>577</td>
<td>(26)</td>
</tr>
<tr>
<td>Igbaria et al., MISQ 1997</td>
<td>540</td>
<td>(27)</td>
</tr>
<tr>
<td>Venkatesh, MISQ 1999</td>
<td>527</td>
<td>(28)</td>
</tr>
<tr>
<td>Agarwal &amp; Prasad, DecSci 1997</td>
<td>522</td>
<td>(29)</td>
</tr>
<tr>
<td>Rai et al., ISR 2002</td>
<td>595</td>
<td>(30)</td>
</tr>
<tr>
<td>Dabbahkark, IJRM 1996 **</td>
<td>45B</td>
<td>(31)</td>
</tr>
</tbody>
</table>

* Citation count not found from Scopus. ** Included in the sample due to popularity in Web of Science based analysis. Citation count obtained from Scopus.


**Other abbreviations.** TAM = Technology acceptance model; SN = Subjective norm; TRA = Theory of reasoned action; TPB = Theory of planned behavior; BI = Behavioral intention; CI = Continuance intention; SE = Self-efficacy; PU = Perceived usefulness; PEOU = Perceived ease of use; U = Frequency or amount of use

**Table 1. Inductivist and falsificationist findings in the 31 most cited IS use papers.**
dies. These papers have not, at least as of now, reached a recognition that would have brought them to the group of most cited papers in the field.

To summarise, it appears that IS research community has an emphasis of positive inductivist results among their most cited papers. Even the comparative studies, which could take a critical stance on the theories, have been carried out with a goal of finding which models best fit a given dataset. Such comparisons are “inclusive” in a sense that their results tend to confirm that all the models are more or less truthful in the face of data. They do not allow any of the theories be falsified. In the Discussion we discuss whether “exclusive” comparisons could be carried out in IS research. In such studies one theory’s success would imply the other theories’ falsification.

We are worried that by focusing efforts on extensions and other inductivist findings the IS researchers will miss the opportunity to better understand the weaknesses and boundary conditions of their models. We agree with Adams et al.’s (1992) opinion that “Although the psychometric properties of the two measures developed by Davis appear to have been robust across studies and user groups, they should not be considered as the final chapter in the validation and refinement of these scales” (p. 245).

3 Examples of Falsificationist Studies on IS Use

Although falsificationist research was rare among the most cited papers in IS use research, we are aware of studies that have followed this scientific method. We present two such studies to exemplify this research approach and its value.

3.1 Example 1: Necessity of including individual differences in models

Burton-Jones and Hubona (2005) had paid attention to the claim that individual differences (e.g., age, staff seniority, and educational level) would be fully mediated by perceived usefulness and perceived ease of use in acceptance models. If full mediation were true, individual differences would not be needed in the models as separate constructs. Burton-Jones and Hubona evaluated this assumption by measuring how much direct effect would ordinary individual difference constructs have on the frequency of use and the volume of use across several information systems in a TAM-like model. They found out significant direct effects in all of the studied systems. Evaluating the effect of individual differences would therefore be necessary for the validity of any IS use model.

At first look Burton-Jones’s and Hubona’s finding may appear similar to the numerous “TAM extension” studies that argue for augmenting the previous TAMs with new constructs. However, Burton-Jones and Hubona’s finding is different. TAM extension papers contribute to research by telling how a model’s fit can be incrementally increased by an addition of a new construct. They do not, however, claim that models without, say, self-efficacy should be considered hereafter less valid—they are just less complete and unable to explain the same amount of dependent variable’s variance. In contrast, Burton-Jones and Hubona presented such a validity-decreasing argument as their finding: adding individual differences to the model does not only improve the model’s fit; it is also necessary for its validity, and should have been part of earlier studies as well. This makes their study falsificationist.

3.2 Example 2: Necessity of having a separate model for each use purpose

In this study (which is our own earlier work; see Salovaara et al., 2013), we were motivated to investigate the implied assumption that the individual variations in the purposes of use can be ignored in model development. We had paid attention to the two competing views within IS science concerning the use of information technologies. We contended that since 1990s, research has shown that ISs are used for multiple different purposes that advance different individual, organizational and political objectives. Users’s interpretations of an IS’s purpose and its actual use will therefore undergo changes and will differ from one point of time and one user to another. We based this claim in particular to qualitative IS research (e.g., Lassila and Branchseau, 1999; Majchrzak et al., 2000; Orlikowski, 1996;...
Robey and Sahay, 1996) and studies on computer-mediated communication (Isaacs et al., 2002; Nardi et al., 2000). We inferred from this that IS use is therefore heterogeneous, and different users will have different beliefs about its purposes and usefulness for those purposes.

Importantly for falsification, we pointed out that the view on heterogeneity is at odds with the IS use studies that deal with quantitative models, particularly with research on technology acceptance and continued IS use. In these studies the beliefs about IS are measured on a general level, without taking into account that perceived usefulness, for example, is likely to refer to different purposes of use by different users. Also the dependent variable—usually behavioural intention (BI) or a frequency of use—is measured with a single variable, although it is likely to represent a mix of uses for different purposes. Heterogeneous use implies that the dependent variables—BI and frequency of use—are both multi-valued and cannot be reduced to single scalar variables.

We gathered empirical data to further examine the discrepancy and tested whether neglecting the heterogeneity in IS use is a threat to these models’ validity. Using Microsoft Excel as an object of study, we hypothesized that 1) respondents use Excel for different purposes; 2) they have different interpretations about its primary purpose of use; and that 3) their perceptions about Excel’s usefulness vary across different possible purposes for which it can be used. All the hypotheses were confirmed. We concluded that the theories and models of IS acceptance should be specified separately for each purpose of use because empirical data shows validity problems in the general-level models.

The studies by Burton-Jones and Hubona (2005) and us show the power of falsificationist research. By falsifying aspects of existing models, they have implications to all subsequent models in the domain. The picture of the generalizability of the IS use models gets sharper in this type of research, and we are better able to understand what the IS use models actually do model.

4 Discussion

In this paper we have argued in favour of increasing empirical criticism within IS research. Falsificationist entail two methods for science: first, in theory building, formulation of falsifiable theories (i.e., presenting statements that can also be shown to be false) in a manner that allows their verification through falsification attempts, and second, in theory improvement, using falsification attempts as the method of theories’ verification. Our focus in this paper has been more on the latter—falsificationist theory improvement—rather than in the former, theory generation. We have suggested that falsificationism would be a valuable research approach also in IS science. If a theory is becoming widely adopted because of mounting inductivist confirmatory evidence (as has happened with TAM), falsification attempts should be particularly valued. This paper exemplified this approach with two such studies.

It is tempting to consider falsificationism as an applicable method only in quantitative research. However, our intention is not to suggest that all IS research should be quantitative and that they would only present falsifiable theories. Instead, we would like to present falsificationism as a general research approach that is applicable both in quantitative and qualitative research. Already in 1985, Klein and Lyytinen (1985) argued that Popperian scientism must not be regarded as the only valid method for IS science. They advocated a pluralist view that allows also other forms of inquiry, and we agree with them. Inductivist and falsificationist research are both valuable. Were all studies required to be falsificationist, existing theories would never be used for accumulation of knowledge (i.e., in “normal science” sense; Kuhn, 1962) because mere application of a theory would qualify neither as theory building nor its improvement. Rejection of inductivist research would therefore be futile. In addition, also in theory building, inductivist research is needed for substantiation of new theories and expansion of theories to new domains.

Falsificationism as a general research approach applies both to quantitative and qualitative theory improvement. Lee (1989) has described how this can work out in case study research in IS science. The general logic, applicable to all kinds of empirical research, is based on the asymmetry of confirmatory
and falsifying evidence—that a theory can never be fully proved but a single observation may suffice for falsifying it. Following this logic, the researcher may consider a theory’s predictions in different initial conditions and the different observation predictions that will follow. A case study can show where these predictions do not hold, thereby pointing out a need and ways to improve the theory (Lee, 1989, p. 40–41). A single case study, when properly conducted and rich with observations, can both serve as an evidence of a theory’s weaknesses and provide direction for suggestions that will improve the falsified theory or replace it with a new one.

Epistemologically, inductivist and falsificationist research follow different logic. We find that IS science would benefit if researchers better identified which approach each of their studies primarily follows. Our impression is that positivist research in IS science remains mostly inductivist (already noted by Lee, 2004, p. 15) while interpretive research presents a wider variety, ranging from explorative studies to refutations and corrections of existing theories. Deliberation on one’s approach, especially in positivist research, would increase the range of methods by which IS science theories are developed. A likely effect would be their increased quality.

The following two sections demonstrate the benefits that may be capitalized by being more conscious of one’s approach. We present the ways in which two typical research designs—replication and comparative research—can be performed in inductivist and falsificationist manners. In IS research, usually only the former is pursued, usually without discussing the falsificationist alternative.

4.1 Inductivist vs. falsificationist replications

A typical inductivist replication study aims for evaluation of a theory with new confirmatory evidence. This may mean a complete replication, with the same methods and research context, or application of the theory to a new domain. In our review, papers by Koufaris (2002), Gefen et al. (2003), Hu et al. (1999), Moon and Kim (2001) and Pavlou (2003) were inductivist replications of the original TAM study in new domains, without a dedicated effort on new theory development.

A falsificationist alternative would be a replication in a domain to which the original theory may be potentially but not straightforwardly generalized and where it has not been empirically tested. This is important especially because IS researchers have been criticized for over-generalizing their findings (Seddon and Scheepers, 2012). For the best effect, the domain for replication should be chosen in a manner where a falsifying result reveals theoretically interesting aspects about the theory. Such a replication will provide valuable information also if the study yields confirmatory evidence, because then it corroborates the theory in a domain where theoretical reasons suggest that confirmation would not be expected.

4.2 Inductivist vs. falsificationist comparisons

In our review, we classified all the comparative studies as “comparative + confirmatory”, indicating that they followed the inductivist research approach. Confirmatory comparative studies investigate which theory has the best fit with the data. The theories that manifest poorer fit are usually not declared being false but are instead regarded as theories of lesser predictive power but which nonetheless hold some truth as well. Some papers may also join theories in the face of mutually reinforcing comparative evidence. Because theories are compared based on the amount of received confirmation, these comparisons do not decisively prune out bad theories from good ones and are therefore not falsifying.

Falsificationist comparisons, in contrast, seek to examine mutually contradictory theories. Contradictions may arise, for example, from the assumptions regarding underlying causal mechanisms that these theories postulate. A comparative study measures evidence for these underlying mechanisms, and generates this way evidence against one theory in favor of another. If the research design is fair, the result can be used for falsification of competing theories. Also other research designs, in addition to measurement of underlying mechanisms, are possible. The key is in the search for research designs that put the theories against each other.
Falsificationist comparisons are common in psychology. The competition between filter theories in attention research is a classic example. The problem was to explain why people cannot attend to two streams of auditory information at the same time: attending to one (e.g., speech in left ear) made it impossible to pay attention to another (e.g., speech in the right ear). Broadbent’s (1958) theory held that sensory information is unconsciously processed by a selective filter, allowing only one channel’s information enter attention and short-term memory. Treisman (1964) pointed out that people nevertheless may notice their own name being mentioned in the unattended channel, which falsified Broadbent’s theory. Treisman’s theory suggested that information goes through an attenuator instead of a filter after which it is pushed through a limited capacity channel before entering the short-term memory. This explained why information in one channel sometimes “broke through”. Finally, Deutsch and Deutsch’s (1963) theory maintained that there is no filter and that all the perceptual stimuli are fully processed but only the ones deemed important excite a response. The competition between these theories was based on falsificationist comparative studies. A similar line of falsificationist research has examined the effect of knowledge on insightful problem solving. Gestalt psychologists have maintained that existing knowledge leads to functional fixedness that hinders the solvers from seeing problems in new ways. The so-called ordinary thinking school has maintained that knowledge, on the contrary, helps the solver see alternative paths of action that novices are unable to see and that functional fixedness is an artefact effect (see, e.g., Weisberg, 2006).

Unfortunately, considering IS use theories specifically, similar falsificationist research designs may be impossible. The reason is that the competing IS use theories tend to be instantiations of theories of reasoned action (TRA) or planned behaviour (TPB) and therefore their belief–attitude–intention models are largely mutually compatible. However, these theories could be possibly successfully compared to the task-technology fit (TTF) model (Goodhue and Thompson, 1995) in a falsificationist manner. The challenge would be in the decision on the suitable dependent variable, however, since TTF focuses on performance effects while TRA/TPB models focus on the prediction of user’s intention and action. Developing and suggesting a research design is however outside the scope of this paper.

To increase the amount of falsificationist studies in IS research we would wish to see not only simple refutations but also constructive interpretations of those refutations in a manner that would increase general understanding of the evaluated theory. This is what in our opinion the two example studies in the previous section have done by presenting necessary conditions (Example 1) or boundary conditions (Example 2) to existing theories. Alternatively, more progressively, a falsification could accompany a competing theory that is at odds with the theory that was evaluated. The new theory would explain the previous empirical content as well as present new predictions for future research. This is the approach that Lakatos (1970) advocated in his sophisticated falsificationism. In either of the ways, the findings from falsificationist studies would not be only negative, but also theory-clarifying or theory-increasing, and would gather interest and benevolent responses from reviewers and readers.

5 Conclusion

We have presented two research approaches—inductivist and falsificationist—and demanded for more falsificationist research in IS science. We sincerely hope our view is not judged of too hastily. We do acknowledge the importance of inductivism and normal science to our field. Yet our concern is that falsificationism is underrated in the present-day IS research practice and, hence, deserves our call for a corrective turn.

Falsifications, although often seen as non-productive research that is less valuable than positive evidence, in fact may yield significant positive contributions. This viewpoint was superbly expressed by Nobel laureate Sir John Eccles, a friend of Karl Popper:

[…] we should even rejoice in the falsification of an hypothesis we have cherished as our brainchild. […] Science becomes an exhilarating adventure where imagination and vision lead to conceptual developments transcending in generality and range the experimental evidence.
The precise formulation of these imaginative insights into hypotheses opens the way to the most rigorous testing by experiment, it being always anticipated that the hypothesis may be falsified and that it will be replaced in whole or in part by another hypothesis of greater explanatory power. (Eccles, 1970, p. 107)

That one’s theory may eventually be falsified can be therefore relieving. In the inductivist approach, researchers easily become preoccupied with maximizing certainty and empirical support, in order to avoid other researchers’ criticism. When this happens, the incentive is to generate only incremental improvements to established theories, since this usually involves the least risk of failure. From a falsificationist viewpoint, in contrast, other researchers’ keen interest in falsifying one’s theory is a testimony of the theory’s value to the research field. The weaknesses that are identified provide an impetus for further creative work. The benefit is twofold. First, falsification increases knowledge of existing theories’ boundary conditions. Second, it also accelerates the exchange of opinions and interactions between researchers, thereby strengthening the research community. Together with inductivist research that expands the applicability of existing theories, falsificationism provides a basis and direction for new discoveries.

6 Acknowledgments

AS thanks Academy of Finland for funding through the Creative use of ICT project. JM is thankful for the financial support from the Foundation for Economic Education. The authors thank Nik Hassan and Alan Litchfield from ICIS2014 SIGPhil workshop for excellent comments on this paper’s earlier draft.

References


Appendix: Papers analysed in the review


