This is the Accepted Manuscript of the paper published at European Journal of Information Systems (2019). DOI: https://doi.org/10.1080/0960085X.2019.1685737

Building on Shaky Foundations? Lack of Falsification and Knowledge-Contestation in IS Theories, Methods, and Practices

Antti Salovaara[a b c d] *, Bikesh Upreti[d], Jussi Ilmari Nykänen[d e], and Jani Merikivi[d]

[a] Department of Design, School of Arts, Design and Architecture, Aalto University; Espoo, Finland. [b] Department of Computer Science, School of Science, Aalto University; Espoo, Finland. [c] Department of Computer Science, University of Helsinki; Helsinki, Finland. [d] Department of Information and Service Management, School of Business, Aalto University; Espoo, Finland. [e] VXT Research Oy; Espoo, Finland.
[f] Department of Management, Technology and Strategy, Grenoble Ecole de Management - Univ Grenoble Alpes ComUE, France.

* Corresponding author. Email: antti.salovaara@aalto.fi. Postal address: Department of Design, P.O. Box 31000, FI-00076 Aalto University, Finland. Phone: +358503360143.

Abstract

Among the defining characteristics of a healthy research discipline is ability to correct its knowledge if more recent evidence creates grounds for this. Studies that reveal errors in earlier theories demonstrate, in line with Karl Popper's thinking, an approach called *falsificationism*. They complement approaches aimed at developing and expanding knowledge by generalising empirical observations or postulating new contributions and testing them. The paper presents analysis that applies this categorisation to abstracts of research papers (N = 5,202) in the eight leading IS journals. Machine-learning-based classification determined that only 7.0% of the papers manifested any clear form of knowledge-contestation, such as falsification, in the approach or findings presented. In light of this, we call on IS researchers to increase the falsification and knowledge-contestation in their research, to nurture more valid theories, methods, and practices, thereby achieving greater societal impact. We present two suitable IS research designs accordingly: knowledge-contesting comparisons and knowledge-contesting replications. We also discuss how these designs, exemplifying opportunities to increase the number of knowledge-contesting studies in the field, can be applied in both positivist and interpretivist research epistemology.

Keywords: methodology, epistemology, falsification, philosophy of science, scientometric research, machine learning

Introduction

When setting out to plan a new study or report on findings, a researcher can take any of various approaches to claim that this work contributes to scholarship. In information systems (IS) research, the researcher may seek to develop a new model, one with higher predictive power than its predecessors; address a poorly understood IS-related phenomenon and thereby expand knowledge in the field; present a new theoretical claim; or make other impacts (Gregor, 2006).

Common to all these contributions is an either exploratory or confirmatory tack in seeking to expand the horizons of knowledge in our field. At the same time, they leave previously presented claims largely untouched. Although exploratory and confirmatory findings are valuable for both scientists and practitioners, they cannot properly exist in isolation – they call for accompaniment by a stream of research that critically identifies, analyses, and contests mistakes, theoretically or empirically unfounded claims, and unwarranted generalisations. Without such research, our knowledge is not always built on a solid foundation so much as a shaky one.

Through this paper, we attempt to engage IS researchers in reflecting on our field's research approaches, asking whether we are applying the most progressive means to refine and advance our knowledge. We start by reviewing the most significant approaches for scientific progress, with an observation that most of them demand critical evaluation of previous findings and competition between rival theories. Following terminology common in philosophy of science, we call this approach *falsificationist*. Later in the paper, we proceed from this notion to propose a generalised concept of *knowledge-contesting research*, which we define as a research approach that criticises and exposes problems in existing knowledge, methods, or practices in research or applied work. We also present a typology of forms of knowledge-contestation in IS research.

We address two research questions (RQs), asking, firstly, what percentage of publications in the IS field can be considered knowledge-contesting (RQ1) and, secondly, whether knowledge-contesting research has an impact with regard to citations when compared to research that is not knowledge-contesting (RQ2).

Our core premise, derived from philosophy of science, is that research disciplines need knowledge-contesting research if they are to progress and improve their theories, methods, and practices. Against this backdrop, we find a seeming paucity of knowledge-contestation in the IS field alarming, and we set out to address RQ1 by classifying 5,202 publications from the AIS's so-called Senior Scholars' Basket of Journals (referred to below as the 'basket of eight'; see https://aisnet.org/?SeniorScholarBasket), which the AIS college of senior scholars deems to encompass the most important journals of our

field: *MIS Quarterly* (MISQ), *Information Systems Research* (ISR), *Journal of Management Information Systems* (JMIS), *Journal of the Association for Information Systems* (JAIS), *Information Systems Journal* (ISJ), *European Journal of Information Systems* (EJIS), *Journal of Information Technology* (JIT), and *Journal of Strategic Information Systems* (JSIS). For our enquiry into RQ2, in turn, citation analysis suggests that knowledge-contesting papers do not differ from other types of publications in their research impact.

The implications of our results are both thought-provoking and able to inspire improvement. We argue that progress in IS research currently relies too much on seeking exploratory and confirmatory evidence (rather than falsificationist or knowledge-contesting evidence). The paper presents methodological suggestions for making faster progress: contributions to knowledge-contesting comparative research and knowledge-contesting replication work. Finally, we will describe how these approaches are highly suitable for the positivist and interpretivist tradition of IS research alike.

Knowledge-Expanding and Knowledge-Contesting Avenues to Scientific Progress

We begin with a brief summary of literature on scientific progress in the field of philosophy of science. Using the terminology introduced above, we group the research approaches into the knowledge-expanding and the knowledge-contesting, and we describe how both of them are crucial in any scientific discipline. Historically, philosophy of science has been influenced greatly by positivist research epistemology, a view in which science is aimed at uncovering objective facts and in which theories are treated as claims that can be empirically evaluated, as in the natural sciences. In the historical summary below, we use the term ‘theory’ in this positivist sense; however, we will later expand it also to encompass interpretivism and the corresponding, different approach taken therein to objectiveness of research knowledge.

Knowledge-expansion-related goals are most evident in *inductivism*, credited to Francis Bacon. This approach was the first to dominate empirical research and remains invaluable today. In inductivism, observations are generalised to ‘laws’ that the researcher seeks to test and confirm. If confirmed, laws can be expanded via more observations (e.g., Gilles, 1993). This emphasises generation of theories representing increasingly accurate knowledge about observed reality.

In the 19th century, thanks largely to a debate between William Whewell and John Stuart Mill, *hypothetico-deductivism* was developed. That important complement to inductivism was shaped further in the next century by the logical positivists. While inductivism involves developing theories after seeing the data, in the hypothetico-deductive method the researcher first postulates a hypothesis and then evaluates it in

light of data. While this hypothesis may constitute a challenge to a theory, a confirmatory approach appears to be more commonplace, just as in the natural sciences (Hansson, 2006) and psychology (Sanbonmatsu et al., 2015). Such a confirmatory knowledge-expanding hypothesis might predict the state of affairs in a context that the existing knowledge does not describe sufficiently. If the data support the hypothesis, the theory behind that hypothesis gains inductive support (Andersen & Hepburn, 2016; Hempel, 1996, p. 18; Snyder, 1997), and knowledge expands to cover the hypothesised state of affairs.

Knowledge-expanding research is saddled with a logical puzzle, though. Presented by 18th-century philosopher David Hume and hence referred to as Hume's problem, the issue is that no amount of evidence can give us full certainty that an inductively derived law is actually true: even if one has only ever seen white swans, there is no guarantee that other swans would not exist. Over the course of centuries, Hume's problem has led to philosophers and scientists becoming increasingly convinced that knowledge-expanding research requires a counterpart for seeking not just to generate claims or confirm them but also to test them critically. This is what approaches drawing from *falsificationism* provide.

The approaches subscribing to falsificationism are *knowledge-contesting*. In a contrast to inductivism, the underlying principle of *falsificationism* is an asymmetrical relationship between proof and disproof: Although confirmatory evidence can never prove a theory, even a small finite number of negative observations can, in principle, falsify it. This follows from the logic of *modus tollens* – if a theory claims that 'all *p* are *q*', finding an instance wherein a *p* is not *q* suffices to disprove that theory's universal validity.

Those subscribing to falsificationism call upon science to generate the sort of knowledge that one could falsify. Accordingly, the strongest scientific knowledge is the kind that survives repeated falsification attempts and provides informative explanations or predictions (Ackermann, 1976, p. 37; Magee, 1973, p. 26; Popper, 1959, pp. 286–287). This said, researchers should propose theories, seek to refute (i.e., disprove) their alternatives, and search for corrected theories that better explain the phenomena of interest. Falsification is therefore a productive method and accelerates the growth of knowledge. It complements inductivism by identifying existing theories' boundary conditions and directs researchers' focus to domains where the theory is insufficient.

Falsificationism appears in different versions. The first and perhaps the best-known, sometimes called *naïve falsificationism* (Lakatos, 1970) or *dogmatic falsificationism* (Ackermann, 1976), was presented by Karl Popper on the basis of ideas proposed by several 19th-century philosophers and scientists, among them Spencer, Mach, Boltzmann, and Mill (Feyerabend, 1981, pp. 192–195). In this approach, even a single falsifying empirically obtained piece of evidence can serve as a decisive refuting blow

to a theory. Naïve falsificationism is focused on seeking 'crucial experiments': empirical studies that would, even with a single experiment, determine whether a theory could be accepted (Ackermann, 1976).

A more elaborate version of falsificationism was presented by Imre Lakatos, who proposed *sophisticated falsificationism* to solve three problems of naïve falsificationism. Firstly, naïve falsification failed to acknowledge properly that falsifying studies themselves are fallible. That is, under the so-called Duhem–Quine thesis, every claim is inextricably tied to methods and other scientific beliefs. If data seem to falsify a theory, a possibility always remains that the outcome stems from external factors and auxiliary assumptions instead of the theory itself (e.g., Gilles, 1993, pp. 98–116). Secondly, many theories are probabilistic, so both confirmatory and falsifying results are possible by pure chance through type I (false-positive) and type II (false-negative) errors (e.g., Howell, 2002). Finally, naïve falsificationism would, in effect, rule out the development of any new theories, because the earliest versions of a theory are usually immature and hence would be too easy to falsify.

Lakatos's suggestion was to consider theories to be instances of larger theoretical developments called research programmes (Lakatos, 1970). As long as the programme manages to maintain positive knowledge-expanding momentum in its theory development, it can successfully resist the falsifying evidence. Weak programmes, in contrast, would fail to do so. Also, probabilistic theories thereby become scientific and falsifiable: while negative evidence is itself fallible, its accumulation finally speaks against the theory in question and leads the scientific community to abandon it in favour of stronger ones. In a Bayesian interpretation, evidence adjusts researchers' confidence in a theory (upward or downward), yielding their current confidence as a product of prior and posterior confidence levels (Earp & Trafimow, 2015; Papineau, 1996).

A third version of falsificationism can be found in the writings of Larry Laudan (1977), who discarded the idea of science as truth-seeking. He suggested that science should be *problem-solving* in seeking to maximise the scope of the empirical problems tackled and attempting to minimise anomalies and conceptual problems (p. 66). Here, theories' truthfulness is of secondary importance to the utility of the answers produced by them. Instead of truthfulness, research traditions should compete in terms of problem-solving capacity.

Fourth and finally, falsificationism plays an important role in Thomas Kuhn's description of scientific progress as alternation between *normal science and revolutions* (1962). The former refers to stable periods of knowledge-expanding science during which consensus (a particular 'paradigm') prevails in the scientific community, while revolutions are times when the accumulated empirical evidence has started to highlight mismatches within the paradigm and rival theories start to emerge. Response to the state

of anomaly ultimately leads to a new paradigm becoming widely accepted, and a new state of normal science is established.

Expansion and Contestation of Knowledge in IS Literature

Looking at IS literature, one finds, a wealth of philosophical texts notwithstanding, little discussion of falsification. For example, Hassan's (2014) analysis of the relevance of Kuhn's work for IS theory cites 30-plus IS publications (p. 600) and classifies them into eight categories, but no form of falsificationism can be found among these. The recent EJIS special issue on philosophy and the future of the discipline addresses other topics, especially values in IS research and their evaluation / critical reflection on them (Cheikh-Ammar, 2018; Chiasson et al., 2018; Ngwenyama & Klein, 2018; Rowe, 2018). Falsification is touched upon in two articles. Hassan et al.'s editorial (2018) problematises our field's hypothetico-deductive research practice, which typically combines deductive reasoning with adoption of theories from other disciplines. The authors maintain that this combination leads to non-ampliative research wherein knowledge is mostly confirmed, not expanded (p. 268). Second is the article in which Williams and Wynn (2018) point to a 'dominant script' (Grover & Lyytinen, 2015) that involves identifying knowledge gaps, 'recruiting' theory from outside the discipline, and testing it. They suggest an alternative script that, similarly to what Hassan et al. describe, is tuned to legitimising new forms of knowledge-expanding research. In their call for increased knowledge-expansion in place of increased knowledge-contestation, these two papers represent the flipside of our work.

Relative to falsification and contestation of knowledge, many more papers in IS literature have been devoted to the field's positivist–interpretivist distinction (e.g., Chen & Hirschheim, 2004; Goles & Hirschheim, 2000; Klein & Lyytinen, 1985; Orlikowski & Baroudi, 1991) and to research designs (Chen & Hirschheim, 2004; Farhoomand, 1987). Among the authors of these, Farhoomand (1987) introduces Popper's and Kuhn's models of science but applies only Kuhn's paradigm concept in an analysis of the IS research tradition. Of the few papers on falsification, that by Salovaara and Merikivi (2015), examining the 31 IS publications cited most, reports that only six of the works assessed refer to any kind of falsifying findings. While Klein and Lyytinen (1985) and Lyytinen and King (2004) present Popper's falsifiability criterion as a possible (but imperfect) means by which IS research could identify core theories, those authors choose to focus on the IS field's research identity, so falsification does not receive primary focus. Reflecting on IS as a research discipline without dedicated emphasis on falsification, Banville and Landry (1989), Hassan (2014), and Hassan and Mingers (2018) have all addressed knowledge-contestation in their evaluation of IS research through the lens of Kuhnian paradigms and revolutions, but they come to different conclusions. Banville and Landry find these concepts poorly applicable to IS, while the

latter two papers present them as useful and adequate.

Coming closest to our work in their focus are three writings by Allen Lee and his colleagues. In the first of these, Lee (1989) discusses how case studies are able to fulfil the criteria for rigorous scientific theory and can falsify earlier theories. He recommends studies in contexts whose scopes match the original theory's generalisability claims. If these contexts are chosen in a falsificationist manner, the studies permit partial refutation of the original theories (e.g., new boundary conditions). In the second paper, discussing the general benefits that philosophy of science can bring IS researchers, Lee (2004) expresses puzzlement at the persistence of traditional, inductive positivism (a knowledge-expanding approach) and suggests that the reason for its popularity may lie in a general lack of awareness of logical positivism's shortcomings (pp. 15–17).

The third paper in the set (Lee & Hubona, 2009) owes much to Popper's thinking. It uses the above-mentioned *modus tollens* logic in presenting a falsificationist scientific model for both positivist and interpretivist IS research. In the former, *modus tollens* drives both the statistical null hypothesis and higher-level falsifications of theoretical claims and models (pp. 247–250). To extend the remit of the *modus tollens* rule beyond positivism, the authors describe how it can inform interpretivism too, noting a parallel with the principle of the hermeneutic circle in interpretivism – i.e., continuous reflective comparison and mutual adaptation between individual observations and larger interpretations (Klein & Myers, 1999).

Finally, Grover and Lyytinen's (2015) review of IS theories features a call for increased falsification. The authors conclude that IS studies too often end up in a 'middle range' where, rather than engage in indigenous theory-building, they adapt theories from other disciplines. Grover and Lyytinen recommend appreciating other forms of theorising also, even ones deviating from the typically accepted standards of rigour. One of the many avenues mentioned is falsificationist research designs (p. 288).

The few articles that have discussed falsificationism in IS so far have been unanimous in pointing to a lack of this approach. One may note that, with regard to lower-level analysis processes, IS research does actually apply falsificationist logic. Positivist statistical tests can be described as falsification attempts in that, technically, they constitute attempts to falsify null hypotheses (Lee & Hubona, 2009, pp. 247–248). The hermeneutic cycle of interpretivist data analysis, in turn, consists of series of trials wherein the researcher formulates interpretations and critically compares them with data so as to correct them (Klein & Myers, 1999; see also the constant comparative method per Glaser & Strauss, 1967). The stuff of our paper, however, is 'higher', contribution-level falsificationism and knowledge-contestation in IS research. This level involves the claims that scholars present as their findings: whether these are knowledge-expanding or knowledge-contesting. That is, while findings may be arrived at on a lower level via

falsification of a null hypothesis or through a reflective, critical interpretive process, the results may be incremental, not contesting existing research knowledge, and therefore be knowledge-expanding.

Our review thus highlights that the importance of falsificationism on this higher level has been brought up only in some philosophy- or method-oriented IS papers. Furthermore, we are not aware of any attempts to evaluate the extent to which a falsificationist or knowledge-contesting research approach plays a part in actual IS research. A healthy research field should encompass these approaches (e.g., Lakatos, 1970), however. How close our field is to a healthy balance has hence remained open to question. To answer it, we employed scientometric analysis as presented below.

Bringing Knowledge-Expanding and Knowledge-Contesting Approaches to IS Research

We can now present two specifications that extend the foregoing general discussion of knowledge-expansion and knowledge-contestation for purposes of application in analysis of IS research.

The first specification pertains to the nature of theories and theorising. Research in IS is epistemologically more heterogeneous than the natural-sciences work that inspired the classic writings in philosophy. Theories in IS research must be considered in both their positivist and interpretivist connotations – positivist theories are often defined as claims about objective, empirically testable explanations or predictions where the phenomena are often observable only indirectly, via instruments (e.g., Risjord, 2014; Suppe, 1972), while interpretivism requires a different definition, since it assumes not an objectively observable reality but a more subjective, constructivist worldview. The definition for an interpretivist theory must allow room for multiple plausible accounts of a phenomenon (Gadamer, 1975), especially because theories are inextricably value-laden, perspective-imbued, and subjective (Risjord, 2014). In this sense, they can be ‘consultable records’ of social phenomena (Walsham, 1995) and ‘sensitizing devices to view the world in a certain way’ (Klein & Myers, 1999, p. 75). They can resemble interpretations by being researchers’ accountings of conceptual structures (e.g., meanings and norms) used by the people studied in their social living (e.g., Geertz, 1973, p. 321; Risjord, 2014, p. 53).

To accommodate both epistemologies, we focus on a characteristic we can identify in both: they make claims about matters/issues that are relevant for understanding certain settings. The claim might be an objectivist proposition such as ‘perceived usefulness predicts behavioural intention’ or an interpretivist statement such as that ‘to understand how users adapt their IS use, it is informative to use discrepancies in the workplace’s social structures as the analytical lens’ (as was done by Majchrzak et al., 2000, and Tyre & Orlikowski, 1994). A claim may also be critical, as with positing that insufficient

attention is paid to a particular feature or that theoretical work in a certain research stream has been blind to an important element. With this kind of unifying definition for theories, our analysis of IS publications can assess whether the claims are knowledge-expanding or knowledge-contesting, similarly across the two epistemologies.

The second specification involves research contributions that go beyond theories. Philosophical texts on scientific progress focus almost entirely on theory-building, while IS research attends also to methods (e.g., PLS) and practices (e.g., how agile methods are used in software development). IS field also covers both academic and practitioner-oriented matters. Hence, we define the (knowledge-expanding and knowledge-contesting) research approaches as comprising three types of contribution: theories, methods, and practices; also, they can incorporate both academic and practical viewpoints.

With these two specifications in mind, we can delve more deeply into the shapes and forms knowledge-expanding and knowledge-contesting studies take in IS research. The former may, for example, offer new observations, propositions, constructs, interpretations, concepts, and sensitising devices. The knowledge-contesting approach, however, requires more consideration. While claims that contest existing knowledge can be seen as falsifying statements in positivist research, pinpointing a similar operationalisable definition for interpretivism appears difficult at first. This is because traditional means of falsification, such as counterproofs, are not in interpretivism's primary interest (Klein & Myers, 1999, p. 75). A more general knowledge-contesting approach is frequent also in interpretivism, however: publications may present interpretive counter-arguments to earlier conclusions and engage in debate or critique. Klein and Myers's (1999) and Walsham's (1995) widely cited articles on interpretivism in IS cite several examples of such knowledge-contesting papers. Klein and Myers (1999) refer to a study in which Myers (1994) analysed an IS implementation failure and argued against the narrow, mechanistic theories for IS implementation that predominated at the time. Walsham (1995), in turn, highlights a study by Suchman (1987) that was knowledge-contesting from its very onset: its purpose was to show that user interface design based on script-like rigid interaction principles that was popular at the time would result in serious breakdowns in human–machine communication. Other examples are Orlikowski's (1992) criticism of the 'technological imperative' model, Trauth and Jessup's (2000) comparison of positivist and interpretivist approaches in an analysis of a group support system, and Wu's (2012) mixed-methods analysis of how different respondents may interpret the same technology acceptance model (TAM) survey items in contradictory ways. In addition, the latter two papers present arguments against reliance on solely positivist studies of IS use. These examples illustrate how interpretivist studies can be knowledge-contesting even when not conducted for traditional positivist falsification.

A Typology of Knowledge-Contesting Research

When employing the adaptations above, one can identify several sub-types of knowledge-contesting IS research, as shown in Table 1. We developed this typology iteratively through the scientometric classification presented further on in the paper.

Table 1: Typology of knowledge-contesting research approaches (example papers for each sub-type are presented in Appendix A)

Type	Short name*	Pattern in the abstracts
Type of argument		
Contestation of an accepted belief	Contested belief	Presents a theory, method, or practice as accepted knowledge in the field, then presents a study or analysis undertaken to reveal an error in it.
Prescriptive comparison	Prescr. comp.	Compares existing theories, methods, or practices with each other in pursuit of finding a 'winner' that therefore should be accepted while the others should be discarded or held suspect.
Prescriptive improvement	Prescr. impr.	Presents a novel or modified theory, method, or practice, with the intent of proving its superiority over that theory, method, or practice's earlier versions, which are, therefore, to be discarded or held suspect.
Exposure of a problem or insufficient knowledge	Exposure of insufficiency	Presents a case or phenomenon that the relevant knowledge is unable to address.
Negative aposterioristic result	Apost. result	As an outcome of data analysis, presents a finding that disconfirms the expected result or a part of it.
Analysis of disagreements	Disagr. analysis	Analyses opposing theories, methods, or practices in the field.
Debate	Debate	Discusses errors in other work or participates in an ongoing discussion wherein theories, methods, or practices are presented as being in conflict.
Call for knowledge-contesting research	Call for contest.	Calls for increased falsification, knowledge-contesting research, or critical attitudes in the field.
Type of target		
Method	Method	The knowledge-contestation addresses the rules, principles, or steps of a procedure, a technique, or an approach.
Theory	Theory	The knowledge-contestation addresses a theory, claim, viewpoint, or belief in the field.
Practice	Practice	The knowledge-contestation addresses how researchers and practitioners apply a method or follow a tradition, an approach, a mindset, or a philosophy in their practice.

* The short forms are used in tables 4–8.

This typology can be used to understand the breadth of means by which existing knowledge can be contested and the spectrum of knowledge types addressed in doing so. In line with the discussion above, Table 1 presents knowledge-contesting research on

two dimensions: *how* the work contests the knowledge ('type of argument') and *what kind* of knowledge is contested ('type of target'): theory, method, or practice.

Of the sub-types presented, 'contestation of accepted beliefs' and 'prescriptive comparison' are most clearly Popperian in nature. They represent research designs that are initially oriented toward falsifying or criticising existing knowledge. An interpretivist example is found in Suchman's 1987 study contesting accepted beliefs about script-based user interface design. 'Prescriptive improvement' studies, in turn, are most clearly Lakatosian, in that they replace weak versions of knowledge with superior ones and thereby make the corresponding research programme stronger. The word 'prescriptive' here denotes that the improvements are presented in an imperative manner: the older research contributions should be abandoned in favour of new ones. The category of 'exposing problems or insufficient knowledge' is closest to Laudanian falsificationism, with emphasis on the importance of research problems that articulate the identity of the research tradition. Myers's (1994) criticism of mechanistic theories is an example of such exposition. The Kuhnian view is found particularly in papers presenting negative aposterioristic results and analyses of disagreements, in that these sub-types identify anomalies that may be starting points for scientific revolution. Finally, the 'debate' and 'call for knowledge-contesting research' classes are meta-level sub-types that do not necessarily present contestation themselves but sustain or encourage such approaches in other publications. Example articles are presented in Appendix A.

These sub-types are not mutually exclusive. The present paper, for instance, both exposes a problem or insufficient knowledge (we claim that IS research lacks awareness about the limitations in its body of knowledge) and advocates knowledge-contesting research. On the second dimension, this paper would be classified as practice-oriented, because we do not contest a theory or method but call for correcting the field's scientific practice. Also, though our typology is suitable for all the papers we eventually classified as knowledge-contesting in our study, we make no claim that our typology is complete. Other sub-types of knowledge-contesting research may be identified.

A Scientometric Study of Knowledge-Contestation in IS Literature

In a similarity to other disciplines, in IS research it is important to both expand our knowledge and critically evaluate what has been learnt. We should, therefore, expect to see both knowledge-expanding and knowledge-contesting research in our field's practice. We hypothesised that, to the contrary, only a small percentage of IS publications could be considered knowledge-contesting. Concerned about this possibility, we conducted the scientometric study described below to examine the field's amount of knowledge-contesting research.

The Research Problem

In essence, our task was to quantify the knowledge-contesting research and compare the amount to that of other ('not-contesting') research. Thus, our classification task was binary in nature, where the non-contesting publications comprised both knowledge-expanding papers and works that were not classifiable on the expansion–contestation dimension. For example, meta-analyses and literature reviews are sometimes (but not always; see Kepes & Thomas, 2018) neither knowledge-expanding nor contesting when they limit themselves to neutrally cataloguing what has been studied and ascertained.

Following the recommendations to make one's research falsifiable, we attempted to falsify our own hypothesis as to the scarcity of knowledge-contestation in IS literature. We therefore sought a finding that the IS literature, in fact, displays copious contestation of knowledge, and maximised the possibility for this finding by opting for an inclusive definition of knowledge-contesting research. Arriving at a low percentage even with an inclusive definition, then, would corroborate our hypothesis of a small percentage.

We embraced this inclusiveness in several ways. Our two-dimensional typology of knowledge-contesting research (again, presented in Table 1) covers several *argument types* in the knowledge-contestation domain that go beyond the most traditional, rather narrow approach of naïve falsificationism and encompasses three *target types* instead of only theory-oriented falsification. Secondly, we decided to classify as falsifying all of the papers that showed at least attempted falsification, even if the authors did not report success. Also, even efforts actually designed to be knowledge-expanding were deemed knowledge-contesting if the researchers reported at least one kind of knowledge-contesting aposterioristic finding (e.g., describing an unsuccessful attempt to find a significant path coefficient in a structural equation model).

Operationalisation

To estimate the percentage of knowledge-contesting research in the IS field, we used machine learning by training a classifier with a manually coded set of papers from the full corpus of basket-of-eight journals, thus analysing all the studies considered to be the most qualified in our field. Had we utilised not a classifier but manual coding, our sample could not have covered all the basket-of-eight journals, and questions could have arisen as to whether such a smaller sample would generalise to IS research overall. Our approach also rendered our study replicable and falsifiable: this paper and its supplements provide the details and data necessary for others to train their classifiers in the same manner and critically evaluate all the stages in our process as well as its results.

Our analysis focused on abstracts instead of papers' full text. There were two reasons for this choice. Most papers in our corpus were not available in plaintext or HTML form; there were only PDF files, whose full text was not usable for machine learning, especially when they consisted of scanned images. Secondly, we wanted to focus on the main findings. Given that abstracts summarise the most important contributions of the paper, they should communicate what the authors considered most valuable in their research. If a paper presents a significant falsificationist result, the abstract should reflect this, whereas the full text brings in extraneous information and greater difficulty in automatically learning the main points. We acknowledge that this criterion entails a possible source of bias: authors may, for example, avoid presenting negative findings in the abstract (the limitations of our operationalisation are addressed in the 'Discussion' section).

Following Lakatos (1970, p. 116) and Laudan (1977, pp. 148–149), we did not acritically deem incremental contributions to be knowledge-contesting research. According to both philosophers, in order to falsify its predecessors, a theory should both explain a larger amount of empirical content and explain new phenomena that earlier theories would find altogether improbable. Incremental improvements meet only the first of these two criteria. This demarcation was important for our decision on whether or not the so-called TAM extensions should be regarded as falsifications to the original TAM. In most cases though not always, we found them not to be knowledge-contesting, because they retained the original TAM core and usually only added more variables to the model. Similarly, when a TAM approach (whether modified or not) was applied to new contexts of use with a confirmatory intention instead of in efforts to find a case that would yield a negative result and thereby point to new boundary conditions for the model, we did not consider the study knowledge-contesting.

Creation of the Dataset for Training

We used Scopus to retrieve the data on papers and then discarded the items that lacked an abstract; these were primarily editor's notes and review articles printed in EJIS, JMIS, and JSIS. Also, for unknown reasons, 31 papers published in JIT in 1987–1989 lacked an abstract. In addition, Scopus did not contain some publications from 2016, indexing only 77 papers, as compared to the approximately 280 publications per year in preceding years. Finally, there were nine pieces that were duplicates in everything but citation count. We removed the duplicate version but retained both citation counts for the analysis for RQ2. The final sample, containing 5,202 papers, is characterised in Table 2.

Table 2: Summary of the papers analysed

Journal	Years missing	Years included	No. of papers covered	Papers per year	% of dataset
EJIS	1991–1994	1995–2016	671	30.5	12.9%
ISJ	1991–1993	1994–2016	421	18.3	8.1%
ISR	–	1990–2016	753	27.9	14.5%
JAIS	2000–2006	2007–2016	284	28.4	5.5%
JIT	1986	1987–2016	642	21.4	12.3%
JMIS	1984–1986	1987–2015	986	32.9	19.0%
JSIS	–	1991–2016	436	16.8	8.4%
MISQ	1977–1979	1980–2016	1,009	27.3	19.4%
Sums:	21 journal-years	205 journal-years	5,202 papers		

We developed the training set (provided in full in this paper's supplementary material) through several iterations. Independently, three members of the team of authors each carried out exploratory full-text analysis of approximately 50 papers from MISQ, ISR, and EJIS, to assess the difficulty of classifying the papers at knowledge-contesting vs. non-contesting level and evaluate whether analysis of abstracts would suffice or, instead, would such a task demand a full-text analysis. We saw informally that abstracts would provide the necessary information for decision-making. We noticed also that knowledge-contesting papers were difficult to find even under our highly inclusive criteria, with only a handful of the papers in the initial sample showing such characteristics.

We accelerated our analysis by using Scopus's keyword-based search functionality. From the abstracts already analysed and classified as knowledge-contesting, we manually identified words with a falsificationist tone (e.g., 'inadequate', 'incorrect', or 'misspecified'), then used them to find more candidates from MISQ, ISR, and EJIS. We continued until the size of our dataset of manually classified papers reached 250, after which each of us classified each new candidate separately and we cross-evaluated our classifications. Upon meeting and reaching agreement, we drafted the first version of the codebook and typology of knowledge-contesting papers. At this point, we had 72 abstracts classified as knowledge-contesting from MISQ, ISR, and EJIS.

Because of the multiplicity of sub-types in our typology, each needing to be represented by several examples in our training set, we set a target of finding at least 200 abstracts

of knowledge-contesting papers, alongside the same quantity of non-contesting-paper abstracts, for training of the classifier.

To expedite the very time-consuming search for examples, we used an intermediate classifier to identify more candidate knowledge-contestation abstracts: we trained a radial-kernel-based support vector machine (SVM) classifier (Cortes & Vapnik, 1995) with the 250-sample dataset we had at this point. Using the intermediate classifier, the fourth member of the team sampled another 250-abstract dataset, within which the classifier predicted approximately 50% of the abstracts to be knowledge-contesting. The other three co-authors analysed the corresponding abstracts, first independently and then together, without knowing the predictions. At this stage, we also verified the inter-coder reliability of our manual coding with a sample of 25 abstracts, achieving a Cronbach's alpha that was above the recommended threshold of .70. After we had evaluated the intermediate classifier's output for 525 (250+250+25) abstracts jointly, we recognised that the intermediate classifier had found six papers – editor's notes and introductions to special issues – that should not have been included in the corpus. After the accordant removals, our final training set consisted of 519 abstracts, with consensus that 220 of them represented knowledge-contesting work.

Our final step, to finalise the typology of knowledge-contesting research (i.e., to complete Table 1), was to classify the knowledge-contesting examples from the dataset in more detail by labelling them with exact sub-types from our codebook. This was done by splitting the material into two portions and assigning each portion to one pair of coders from our four-author team. The two people working on a given portion examined the abstracts independently and then jointly, reaching consensus on each paper's sub-types. Then, the first author reviewed the other pair's work and discussed it with them until consensus was reached about all the sub-types.

Training of the Classifier

From a machine learning perspective, identification of knowledge-contesting papers can be regarded as a document classification problem wherein the documents are either 'knowledge-contesting' or 'not knowledge-contesting'. When compared to typical use of classifiers, our study had an unusual goal. Usually, training is aimed at minimising so-called false positives (erroneous determination that a document falls in the target category). If we had optimised for this, we would have attempted to find only the most salient cases of knowledge-contesting research and leave out those papers showing very little sign of knowledge-contestation. Our study, however, was dedicated to the opposite goal. The presence of false positives in our results would not pose a problem, since this only increased the percentage of knowledge-contesting papers found, thereby rendering it *harder* for us to defend our hypothesis. For the validity of our study, it was better to

obtain an over-estimated percentage of knowledge-contesting research. Therefore, we used sensitivity (Ting, 2010) as our performance measure. What we wanted to minimise was false negatives; we did not want to fail in detecting actual knowledge-contesting papers.

In our choice of classifier, we considered only popular algorithms, to rule out the possibility of the findings being an artefact of a bespoke choice of classifier rather than stemming from the data. Given the almost universal superiority of SVMs in classification of text-based documents (Joachims, 1998; Yang, 1999), we trained three SVM variants, with linear, radial, and polynomial kernel functions, and compared their performance with that of a simple regularised logistic regression classifier that has likewise performed well in text classification (Zhang & Oles, 2001). In all the training and classifier evaluation, we used the statistical software environment R, the 'caret' package for SVM training (Kuhn, 2008), genetic algorithms (Scrucca, 2013) for SVM parameter tuning, and 'glmnet' (Friedman et al., 2010) with ridge regression within binomial distribution to train the regularised logistic regression classifier.

Before the training, we pre-processed the abstracts with the 'tm' package (Feinerer, 2015) to convert the text input into a document–term matrix and remove pronouns, auxiliary verbs (e.g., 'has', 'will', and 'does'), articles ('a', 'an', and 'the'), and other frequent English stopwords. In further efforts to ensure retaining only meaningful words, we removed all short words (words with fewer than three characters) and rare ones (words seen in fewer than five documents). Finally, we applied TFIDF weighting (Salton & McGill, 1983), commonly used in document classification: a term (here, a word) in a document receives a high relevance score if it appears often in that document but rarely in other documents.

Tuning of parameters, although necessary, always carries a risk of overfitting. When fed with training data, a classifier tends to optimise prediction performance for the given data and hence learns rules that are tailored for the given training data but perform poorly with new samples. To avoid this pitfall, we followed the industry standard of carrying out tenfold cross-validation in which we used random slicing of the training data into 10 equal portions such that nine slices were utilised for the SVMs' parameter tuning and the tenth was employed for evaluation. To achieve reliable parameter estimates, we carried out this process five times. Finally, we evaluated the competing classification models for their ability to predict class labels for the data not included in the training phase.

In Table 3, we present the classifiers' performance levels as judged by commonly used metrics. To our surprise, the simplest of the classifiers, the regularised logistic regression function, outperformed all three SVMs for accuracy and in most other respects. Therefore, we selected regularised logistics regression as our final classification model.

Table 3: Performance figures for various classifiers, where high values indicate better classification ability (the best value for each metric is shown in bold)

	SVM			
Metric	Linear kernel	Radial kernel	Polynomial kernel	Regularised logistic regression
Accuracy	0.70	0.47	0.68	**0.75**
Accuracy, lower	0.61	0.38	0.59	**0.66**
Accuracy, upper	0.78	0.56	0.76	**0.82**
Sensitivity	**0.85**	0.65	0.70	0.75
Specificity	0.67	0.43	0.68	**0.75**
Precision	0.34	0.19	0.30	**0.38**
Recall	**0.96**	0.86	0.92	0.94
F1	0.49	0.29	0.42	**0.50**
Balanced accuracy	**0.76**	0.54	0.69	0.75

Results

We have now laid the groundwork to present the results for both research questions: what percentage of publications in the IS field can be considered knowledge-contesting (RQ1), and does knowledge-contesting research have an impact (RQ2)? That is, are such studies cited at least as often as non-knowledge-contesting ones? All the data – including the list of papers used for classifier training and the full results discussed here – are provided as an online supplement.

RQ1: How Many IS Papers Are Knowledge-Contesting Ones?

The classifier initially found 466 papers (9.0% of the dataset) that it deemed knowledge-contesting. That said, given that our training of the classifier was permissive with regard to false positives, we suspected the actual percentage to be lower than this. Using the typology in Table 1 as our classification codebook, we performed two-person evaluation (with independent judgements, then cross-checking to reach consensus) and found that *the validated number of knowledge-contesting papers was 365 (7.0% of the full sample)*.

Tables 4–7 present the details of the results. Table 4 shows the journal-level differences in publishing of knowledge-contesting papers, where ISR, JAIS, and MISQ are the journals with the largest percentages. Approximately 10% of the papers they have published showed knowledge-contesting characteristics. At the other extreme, only about 4% of the papers in ISJ, JIT, JMIS, and JSIS were knowledge-contesting. Between these groups lies EJIS, with 7.3%. Also, the table reveals a discontinuity in

2006: the bottom row shows that the number of knowledge-contesting papers rises to approximately 20 per year at this point, then remains at this level (with the last year in our dataset, 2016, being an exception; Scopus did not provide complete data for that year). Part of the discontinuity results from a new journal (JAIS) being added to the counts in 2006.

Table 4: Numbers of knowledge-contesting papers, by year and journal

Count

Journal	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Sum	%	Sparkline
EJIS																0	1	1	0	1	1	2	3	3	2	1	3	5	3	1	2	5	3	1	5	6	0	49	7.3%	
ISJ															1	2	3	1	0	0	0	0	1	1	0	0	0	1	3	2	0	1	1	0	0	0	1	18	4.3%	
ISR											1	3	0	1	5	3	5	1	1	5	0	2	0	2	1	3	8	1	2	6	4	5	11	11	8	4	2	95	12.6%	
JAIS												0	0	0	1	0	0	1	0	0	0	0	2	0	0	0	0	1	0	0	1	2	0	1	2	0	0	11	10.2%	
JIT																												7	3	4	3	2	5	0	1	2	2	29	3.9%	
JMIS								0	1	0	1	0	1	0	2	0	0	0	1	1	1	1	3	1	1	0	0	1	0	1	1	2	0	1	2	3	0	25	4.4%	
JSIS								0	0	1	0	0	1	3	1	0	1	4	3	4	2	3	1	2	1	0	1	4	4	2	2	1	0	0	1	1		43	2.5%	
MISQ	0	0	0	0	1	0	1	1	2	1	0	1	3	1	0	1	2	5	4	2	4	0	2	3	2	2	4	4	3	8	5	6	9	3	6	9	0	95	9.4%	
Average	0	0	0	0	1	0	1	1	3	2	2	4	5	5	10	6	12	13	9	13	8	8	12	12	7	6	16	24	18	24	18	24	29	17	25	25	5	365	7.0%	

Table 5 shows the same journal–year breakdowns as percentages. The average percentages shown here (see the second-to-last column) differ from those in Table 4 because of the calculation formula used: while the percentages in Table 4 come from the relationship between the total number of knowledge-contesting papers and the total number of papers in the journal, we calculated the values in Table 5 by obtaining annual percentages and then averaging them. This gives more weight to earlier years, when publication volumes were lower. Table 5 shows large fluctuations in the percentages of knowledge-contesting papers from one year to another. While this pattern is manifested by all the journals, it is most evident for ISR, with nearly 30% (or more) of the papers being knowledge-contesting in 1991, 1993, and 2006 while no such papers were published in 1992, 2000, and 2002.

Table 5: Percentages of knowledge-contesting papers, by year and journal

Percentage (%)

Journal	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Average	Sparkline
EJIS																0	4	4	0	4	5	11	14	14	9	3	6	9	6	3	5	12	8	3	14	17	0	6.8	
ISJ															7	13	19	6	0	0	0	0	6	6	0	0	0	6	12	9	0	5	5	0	0	0	7	4.3	
ISR											5	33	0	8	28	19	19	5	5	24	0	9	0	11	5	14	35	5	8	21	8	11	15	19	17	8	17	12.8	
JAIS																												23	13	13	10	7	15	0	3	7	17	10.5	
JIT								0	10	0	5	0	4	0	8	0	0	0	4	4	5	6	14	5	4	0	0	3	0	4	4	11	0	5	10	13	0	3.9	
JMIS								0	0	5	0	0	3	9	3	0	2	11	10	11	6	9	3	6	3	0	3	10	10	5	5	3	0	0	3	2		4.2	
JSIS												0	0	0	6	0	0	8	0	0	0	0	15	0	0	0	0	6	0	0	5	7	0	5	10	0	0	2.4	
MISQ	0	0	0	0	7	0	4	3	6	4	0	4	12	6	0	5	11	24	18	8	20	0	13	14	8	7	10	13	9	21	13	12	17	5	11	21	0	8.2	
Average	0	0	0	0	7	0	4	2	5	3	2	4	4	5	8	5	7	9	6	8	5	6	8	8	4	3	8	10	8	10	7	9	10	6	9	9	6		

Interestingly, the sharp growth in the number of knowledge-contesting articles that Table 4 renders visible for 2006 cannot be detected in the bottom row of Table 5. The peak in Table 4 may have arisen as a combined result of a new journal's appearance in the data and a general increase in our field's publishing volume.

Going into more detail, Table 6 presents a breakdown of the 365 knowledge-contesting publications by sub-type. The totals at the bottom of this table show the most Popperian type of knowledge-contesting research – contestation of an accepted belief – to be the most commonplace (27%), followed by the other sub-type that identifies errors in existing knowledge (i.e., exposing problems or insufficient knowledge, at 21%). Comparisons were not as frequent, whether neutral ones between competing alternatives (i.e., prescriptive comparisons: 14%) or between authors' novel contributions and the existing alternatives (i.e., prescriptive improvements: 16%). However, they did outnumber reports on negative aposterioristic results (9%), cases in which the abstract presents findings not consistent with the expected outcomes.

Table 6: The knowledge-contesting papers in each journal broken down by type of knowledge-contestation

	Type of argument								Type of target		
Journal	Contested belief	Prescr. comp.	Prescr. impr.	Exposure of insufficiency	Apost. result	Disagr. analysis	Debate	Call for contest.	Method	Theory	Practice
EJIS	22	7	5	11	4	3	10	2	12	25	14
ISJ	7	1	2	7	3	0	0	0	6	7	5
ISR	31	15	30	12	12	0	8	0	34	40	24
JAIS	8	2	6	12	2	2	7	0	10	15	4
JIT	7	3	3	10	1	3	6	0	8	8	9
JMIS	15	15	5	6	8	0	0	0	11	19	16
JSIS	5	2	1	4	0	1	1	0	2	7	3
MISQ	22	17	17	30	11	1	12	1	33	40	26
Sum	117	62	69	92	41	10	44	3	116	161	101
Total								438*			378*
% of total	27%	14%	16%	21%	9%	2%	10%	1%	31%	43%	27%

* Because a paper could be classified as of multiple types, the totals sum to more than the number of knowledge-contesting papers (365).

Among the more conceptual 'meta-level' publications discussing contestation of research knowledge at a higher level, calls for an increase in falsification and knowledge-contesting research proved extremely rare, with only three papers (1%:

Grover & Lyytinen, 2015; Paul, 2002; Stahl et al., 2011) found in the 365-paper sample. Also, the other sub-types of meta-level knowledge-contesting research (analyses of disagreements and debate) proved rare. The small percentages may attest to the general rarity of non-empirical publications in IS literature. Another possibility is that abstracts in the IS publishing tradition rarely communicate such aspects – a paper might call for knowledge-contestation, engage in a debate, or devote itself to analysing disagreements without emphasising this in its abstract. In particular, a work that is knowledge-contesting in several ways may highlight only one or two such aspects in its abstract. Therefore, one might expect full-text analysis to identify more papers that encompass a meta-level knowledge-contesting contribution. That would increase these sub-types' representation but not the overall percentage of knowledge-contesting publications.

With regard to contribution sub-types – methodological, theoretical, or practice-oriented – the papers were quite evenly distributed. This is interesting in that the literature in philosophy of science has emphasised falsification of *theories* and paid very little attention to how research methods or practices have been found insufficient and been replaced with better ones.

While highlighting such patterns, Table 6 is not suited to comparisons between journals, because the journals differed in their representation in the Scopus database. For example, while data for MISQ have been indexed there since 1980, information about JAIS has been available only since 2007. Accordingly, while Table 4 indicates that in excess of 10% of the papers in JAIS were knowledge-contesting, this is not reflected in Table 6. The latter suggests that very few knowledge-contesting papers have been published in JAIS, because Scopus covers only 2007–2016 for JAIS.

To compare the eight journals, we created a table of percentages of knowledge-contesting research for each sub-type and journal. We limited the analysis to 2007–2015, the years with complete data from every journal in our dataset, comprising, all told, 204 knowledge-contesting publications. Table 7, which presents the results, reveals that the three journals identified in Table 4 as leaders in knowledge-contesting research (ISR, JAIS, and MISQ) are prominent also in percentagewise comparison. However, in Table 7 they are joined by EJIS, whose percentages for contestation targets (method, theory, or practice), for example, are at or slightly above the levels of JAIS. In its sixth issue of 2010, EJIS published a widely recognised special section dedicated to 'Contrarian Information Systems Studies'. The growth in that journal's role as a publisher of knowledge-contesting papers since 2012 may be a result of a lasting effect of that issue on EJIS editorial policy. This possibility merits more in-depth study for confirmation, however.

Table 7: Percentages of knowledge-contesting papers in each journal for each type of knowledge-contestation, from 2007–2015 (for which the dataset is complete)

	Type of argument								Type of target		
Journal	Contested belief	Prescr. comp.	Prescr. impr.	Exposure of insufficiency	Apost. result	Disagr. analysis	Debate	Call for contest.	Method	Theory	Practice
EJIS	27%	13%	8%	15%	9%	0%	25%	50%	13%	15%	18%
ISJ	5%	3%	2%	6%	4%	0%	0%	0%	3%	3%	5%
ISR	25%	25%	44%	9%	26%	0%	6%	0%	25%	22%	29%
JAIS	11%	6%	10%	23%	9%	33%	19%	0%	15%	15%	7%
JIT	2%	3%	2%	9%	4%	33%	19%	0%	4%	3%	9%
JMIS	7%	16%	6%	2%	17%	0%	0%	0%	6%	7%	9%
JSIS	9%	0%	0%	6%	0%	17%	3%	0%	1%	6%	4%
MISQ	14%	34%	27%	30%	30%	17%	28%	50%	31%	28%	18%
Sum	100%	100%	99%	100%	99%	100%	100%	100%	98%	99%	99%
No. of papers	56	23	53	48	32	6	32	2	67	86	55
Total								252*			208*
% of total	22%	13%	19%	21%	9%	2%	13%	1%	32%	41%	26%

* Since any given paper could be of multiple types, the totals are greater than the number of knowledge-contesting papers (204).

Analysis of the various contestation sub-types across all eight journals reveals that EJIS accounts for a larger percentage of the papers addressing contestation of beliefs (27%) than its overall proportion of knowledge-contesting papers (7.3%; see Table 4) would suggest. A similar observation can be made about JMIS with respect to comparisons (16% as compared to 4.4% overall) and JIT for analyses of disagreements (33%, relative to 3.9%) and debate (19%, relative to 3.9%).

These comparisons suggest that particular editorial policies may affect journals' tendencies in publishing of knowledge-contesting research. To investigate this possibility, we reviewed the journals' 'Aims and Scope' descriptions (as of 2019). Most journals explicitly welcome knowledge-contesting contributions: critical evaluations (ISR); critical reviews, opinions, and debate (JIT); 'critical views' (EJIS); contributions that challenge or clarify existing theories (both MISQ and JAIS); and work on paradoxical or controversial research issues (MISQ). The journals that do not mention critical contributions (ISJ, JMIS, and JSIS) indeed display the lowest percentages of knowledge-contesting publications (see tables 5 and 7); however, the percentages were low also for JIT, notwithstanding its solicitation of reviews, opinions, and debate.

We answer RQ1 thus: 7% of research published in the eight leading IS journals is knowledge-contesting. This lends support to our hypothesis that this kind of research is scarce in our discipline.

RQ2: Do Knowledge-contesting Studies Have an Impact?

Considering the scarcity of knowledge-contesting studies in our field, accounting for only 7.0% of the IS research examined, one might expect those few studies to have received more attention and citations than others. The question arises of whether some sub-types of contestation, per the typology in Table 1, may be particularly valued in the field.

To examine the possibilities, we compared citation counts between knowledge-contesting and non-contesting papers in our corpus. To take into account two factors that could confound our comparison – journals' prestige and the passing of time since an article's publication – we compared only between knowledge-contesting and non-contesting papers from the same year and journal, examining citation counts for the 139 instances in the dataset where a journal published both contesting and non-contesting papers in a given year. We calculated the average citation counts for both the knowledge-contesting and non-contesting papers and used these as measurements for pairwise comparison.

The averages calculated for the 139 instances had left-skewed non-normal distributions, with the median citation counts for knowledge-contesting and non-contesting papers, respectively, being 42.0 and 50.7 but the corresponding means being 243.3 and 92.5. Also, two extreme outliers existed in the knowledge-contesting class: Davis's (1989) TAM paper in MISQ, with 15,394 citations, and Venkatesh et al.'s (2003) paper on a unified theory of acceptance and use of technology (UTAUT) in MISQ, with 8,754 citations. All other papers had 4,772 or fewer citations. In response to the non-normal distributions and outliers, we employed Wilcoxon's matched-paired signed-rank test (e.g., Howell, 2002, pp. 713–717) to perform two distinct comparisons: comparing the knowledge-contesting papers with the non-contesting ones as a whole and, on a more detailed level, comparing the individual sub-types of knowledge-contesting articles with non-contesting papers. Because of the 12 repeated tests, we used Bonferroni correction and defined $p = .05 / 12 \approx .0042$ as our level for statistically significant results.

Table 8: Comparisons of citation averages between non-knowledge-contesting papers and knowledge-contesting papers, of various types, in the same year and same journal

Paper type	No. of pairs compared	Median of citation averages	p[a]
Non-knowledge-contesting papers		50.7	
Comparisons:			
All knowledge-contesting papers together	139	42	.223
Contested belief	77	42	.028
Prescr. comp.	47	54	.193
Prescr. impr.	46	36	.031
Exposure of insufficiency	65	69	.169
Apost. result	33	72.5	.458
Disagr. analysis	10	36	.203
Debate	27	17	.001*
Call for contest.	2	22	.180
Method	75	34	.182
Theory	91	52	.410
Practice	63	48	.002*

[a] Significance of the difference between papers published in the same journal and year, based on Wilcoxon's matched-pairs signed-rank test.
* $p < .05 / 12 \approx .0042$ (Bonferroni-corrected level of significance for 12 tests).

Table 8 indicates that *knowledge-contesting papers are not cited more often than non-contesting ones*. In fact, the opposite appears to be prevalent: knowledge-contesting papers that feature debate or focus on practices have statistically significantly fewer citations than non-knowledge-contesting papers published in the same journal in the same year. A closer look reveals that only four sub-types of knowledge-contesting publication (prescriptive comparisons, exposing of problems or insufficiency of knowledge in existing scholarship, negative aposterioristic results, and work focusing on theory contestation) yielded higher citation counts than non-contesting papers. None of these differences is statistically significant.

For RQ2, we state in conclusion that knowledge-contesting contributions to IS research do not receive more attention than non-contesting ones, such as knowledge-expanding studies. In fact, some types of knowledge-contesting contributions are cited less than non-contesting papers in general.

Discussion

We have presented arguments for a greater presence of knowledge-contestation, both empirical and conceptual, within IS research. Our scientometric study supported the troubling hypothesis that knowledge-contesting research is extremely rare in our field: even though we used highly inclusive criteria for classifying papers as knowledge-contesting, our abstract-based analysis showed only 7.0% of the papers published in the eight leading IS journals to be knowledge-contesting. In addition, our citation analysis indicated that knowledge-contesting research does not receive more attention from researchers than other work and in some cases is cited less. This is cause for concern about whether the theories, methods, and practices within IS research are really built on a solid foundation. Should we trust the body of knowledge accrued in academic publications if it is seldom explicitly challenged beyond the peer-review process?

Our findings are consistent with what Rowe and Markus (2018) encountered in their struggles to publish an updated version of a classic yet partly outdated framework. They faced unexpected resistance from reviewers in the process of working through the revisions. In light of this experience, Rowe and Markus recommended that the IS field offer more opportunities for fair critique and revisiting prior contributions. Similarly, we call on our research community to carry out more knowledge-contesting research, accept such work for publication, read it appreciatively, and cite it when doing so is appropriate. Reasoning developed in philosophy of science underscores that knowledge-contesting research is essential for rigour. If, as our study suggests, it is overlooked in the IS field, the knowledge-contesting research that does exist is not used to its fullest and, hence, loses impact.

We find a knowledge-contesting approach to serve the two main epistemological traditions of our field (the positivist and the interpretivist) equally well. To encourage more of this kind of research, we will discuss two types of research design below: knowledge-contesting replications and comparisons. While knowledge-expanding forms of these are widespread in IS research, knowledge-contesting variations are rarely seen. We believe that they would offer important contributions.

Not surprisingly, these two research designs are most readily applicable to positivist studies, with those utilising qualitative methods being no exception (Eisenhardt, 1989). This is because falsificationism was developed through analysis of the natural sciences and since our concept of knowledge-contesting research, in turn, draws ideas from it. That said, knowledge-contesting replications can be carried out within an interpretivist epistemological setting, following Lee and Hubona's (2009) *modus ponens – modus tollens* framework.

Knowledge-Expanding vs. Knowledge-Contesting Replications

We begin our suggestions for knowledge-contesting replications by describing their complement: knowledge-*expanding* replications. With the latter, which are rather common within IS, researchers seek to strengthen existing contributions with new confirmatory evidence. That might mean complete replication, using the same methods and research context, or application of the earlier contribution to a new domain. Studies of TAM in new domains, without dedicated endeavours toward new model development, are typical of this category.

A knowledge-contesting alternative, in contrast, would be a replication in a domain to which the original theory may be generalisable but wherein its validity has not yet been empirically tested. Studies of this nature are sometimes called conceptual (Dennis & Valacich, 2015; Earp & Trafimow, 2015; Olbrich et al., 2017) or far replications (LeBel et al., 2017). Replications in new domains are important especially in response to criticisms of IS researchers as over-generalising their findings (Seddon & Scheepers, 2012). For the best ultimate outcome, we posit that the context of replication should be chosen in such a manner that a falsifying result in particular would yield insight into theoretically interesting aspects of the earlier contribution, such as a theory, method, or practice (cf. Table 1). A falsifying result would suggest a new boundary condition for the contribution and thereby limits to its initially assumed generalisability. Even if it transpires that the replication yields confirmatory evidence, the work would still be valuable, in providing corroboration in a domain in which reasons would suggest *a priori* that confirmation is not to be expected.

To be defendable, a knowledge-contesting replication should meet certain requirements. If researchers obtain a falsifying result by replicating a study in a domain different from that in which the original study was conducted, the original contribution can be defended via the above-mentioned Duhem–Quine thesis (e.g., Gilles, 1993), by appealing to inappropriate execution in the replication's methodology or attributing the conflicting result to ancillary conditions, or the 'protective belt' of the theory (see Lakatos, 1970; Earp & Trafimow, 2015). For precluding claims of erroneous methodology, it has been recommended that the replication encompass direct replication (Dennis & Valacich, 2015; Earp & Trafimow, 2015; Olbrich et al., 2017) in the original domain, where the same finding should emerge as in the original study. This would prove the researchers' mastery of the method, and *ceteris paribus* ('all else equal') logic would entail greater confidence that the falsifying result in the new domain can be attributed to the factor identified (Earp & Trafimow, 2015).

Knowledge-contesting replications can be employed also in qualitative positivist work, wherein a researcher may consider an earlier contribution's predictions for the given conditions and postulate that certain other outcomes should emerge in that setting if said

contribution, such as a theory or method, is valid. The researcher can then demonstrate via a case study that these predictions do not hold, thereby pointing out a need to improve the contribution (Lee, 1989, pp. 40–41). Thus, even a single-case study can supply evidence of an earlier research contribution's weaknesses and point to suggestions for rectifying them.

It may be harder to envision aligning knowledge-contesting replications with interpretivist research: in interpretivism's worldview, replication of this kind cannot be employed to rule out truthfulness or a good match between external reality and the theory or method. That said, because it embraces the possibility of several plausible interpretations for the same phenomenon, replications of a different kind can be carried out as attempts to identify areas of bias, where research has not considered all possible interpretations or explanations. Indeed, there is a history of such critique with regard to positivist theories of IS use. Suchman's (1987) studies of photocopier use and Wu's (2012) mixed-methods TAM study, referred to in the review above, exemplify this. The replications were carried out with prevailing empirical methods but with stated purposes of analysing the observations from knowledge-contesting perspectives.

Knowledge-Expanding vs. Knowledge-Contesting Comparisons

In a parallel with the foregoing discussion, we begin considering knowledge-contesting comparisons by presenting their counterpart. In the most typical knowledge-*expanding* comparative studies, investigating which of several alternative pieces of knowledge shows the best fit with the data, scientists do not seek to compare options that are mutually exclusive. Our conclusion from reading several hundred abstracts of IS articles is that most comparisons in IS research are knowledge-expanding. Rather than falsification, the alternatives found less suitable in these comparisons are regarded only as contributions of lower predictive power while remaining valid. Since such comparisons are not decisively aimed at pruning out bad contributions from among good ones, they are not knowledge-contesting.

With truly knowledge-contesting comparisons (prescriptive comparisons in our typology), scholars examine existing mutually contradictory research contributions, where the contradictions may stem, for example, from the assumptions, posited causal mechanisms, or worldviews. A comparative study may analyse the support for the underlying mechanisms cited and, accordingly, articulate grounds for dismissing one contribution in favour of another (Ylikoski & Aydinonat, 2014). If the comparison is fair, the results can be used for knowledge-contestation.

Positivist knowledge-contesting comparisons are epistemologically unproblematic, by the same logic used for replications, and are commonplace in such fields as psychology. For example, creativity research has seen numerous knowledge-contesting studies

examining whether knowledge affects insightful problem-solving. Gestalt psychologists have sought to prove that existing knowledge leads to functional fixedness that hinders people from seeing problems in new ways. The so-called ordinary thinking school, meanwhile, has produced evidence that knowledge helps people find alternative paths of action that novices are unable to see. They maintain that lack of creative insight does not need to be, nor should be, explained via a concept of functional fixedness (e.g., Weisberg, 2006).

Interpretivist knowledge-contesting comparisons require a different approach. While denial of objective measurable yardsticks practically rules out empirical comparisons, a possibility remains for conceptual critical comparisons. Here, the purpose is to analyse whether alternative approaches, frameworks, or methods entail conceptual weaknesses that could motivate adoption of one framework rather than another. In search of examples of interpretivist knowledge-contesting comparisons, we reviewed the papers labelled as prescriptive comparisons in our 5,202-paper corpus. As expected, we did not find any empirical interpretivist comparisons, but we did note one interpretive conceptual comparison: Seddon and Scheepers (2015) compared four distinct types of generalisation in IS research to learn whether any of them might be suitable for positivist and interpretivist epistemology alike.

Overall, our study determined that, of the 7.0% of IS research found to be knowledge-contesting, only 14% appears to involve knowledge-contesting comparative studies (i.e., prescriptive comparison). We would find it valuable for these to be more commonplace.

Limitations

Our study had several limitations. Firstly, we analysed only articles' abstracts, with the possibility remaining that these did not fully reflect the papers' true content. Hence, the validity of the resulting high-level analysis relies on the assumption that the papers' abstracts express the findings that the authors found most valuable in their research. For example, an abstract may be written to highlight a knowledge-expanding contribution while the actual content is more knowledge-contesting. Word limits imposed on abstracts (e.g., 200 words for EJIS) may force researchers to direct attention to knowledge-expanding rather than knowledge-contesting contributions. For instance, authors of a TAM study who detect an unpredicted weak link between two constructs may decide not to report that in the abstract. Such exclusion of unexpected and therefore interesting negative aposterioristic findings (see Table 1) from abstracts would run counter to our assumption that the abstracts reflect the most valuable findings. If such exclusion is common practice in the IS discipline, the percentage of knowledge-contesting studies would, in reality, be higher than 7.0%.

Regrettably, examining this possibility would demand full-text analysis, which was beyond the scope of our research. We should note, however, that our finding is in line with work in which other researchers found only 3/70 of papers published in *Nature* in 2000, or 4.3%, to be knowledge-contesting (Hansson, 2006). Still, we encourage further research to engage in full-text analysis aimed at falsifying our finding.

The second limitation is related to our pairwise citation comparison for RQ2. While our analysis did answer the research question, we did not investigate why the knowledge-contesting papers were cited: was it for their knowledge-contesting claims themselves or for other, possibly knowledge-expanding content? This analysis would have required extensive full-text analysis of all the articles containing relevant citations. Such laborious analysis was likewise beyond the scope of our paper, though it may be carried out later.

A third limitation involves the possibility of the classifier not having recognised all the knowledge-contesting papers. Because classifiers' training is based on examples, the classification quality hinges on the quality of the training set. There is a chance that our approach failed to detect a class of falsificationist papers that, if included in the analysis, would have led to finding a higher percentage. This possibility exists because our use of keywords and another classifier in the search for examples cannot rule out some sub-class of falsificationist papers escaping our attention. We find it unlikely, though, that the set of papers missed would be substantial, given that we manually evaluated hundreds of non-knowledge-contesting papers over the various stages of analysis. Secondly, we applied inclusive criteria for falsificationism and trained our classifier in keeping with those principles. Stricter definitions of knowledge-contesting research, such as ones adhering more closely to the definitions in philosophy of science, would yield much smaller percentages in this kind of analysis.

Finally, the low percentage of falsification within IS might be explained by a tendency for our field to adopt theories from other fields in the absence of its own reference theories. Whether it does so has long been debated among IS researchers (e.g., Baskerville & Myers, 2002; Lyytinen & King, 2004; Hassan et al., 2018). If IS indeed largely takes its theories from other disciplines, one could claim that those theories may have undergone rigorous maturation processes and falsification attempts already, in the 'mother disciplines'. Were all the theorising, methodology, and practice within IS to be based on theories, methods, and practices adopted from other disciplines, there might indeed be little need for their falsification in IS. We do not believe this is the case, however. Firstly, some IS researchers (e.g., Baskerville & Myers, 2002) argue that our discipline has matured sufficiently and developed a substantial body of innate theories, with Davis (2000) having identified five bodies of knowledge unique to IS in his review of ICIS and MISQ publications, for instance. As the IS discipline is not entirely devoid of innate theories, falsification should be carried within our discipline. Secondly, our

concept of falsification covers methods and practices, not just theories. The need for falsification within our discipline is further accentuated accordingly, in that these two kinds of knowledge may be more innate than theories. Thirdly, even theories rooted in other fields need critical scrutiny, since they may not be straightforwardly transferable to IS contexts. Finally, not all theories developed in other disciplines have necessarily been properly validated in their source disciplines. As the so-called replication crisis within psychology and other fields shows (e.g., Aarts et al., 2015; Baker, 2016), the reference disciplines' findings should not be taken for granted. For all but the most mature, best-corroborated theories, methods, and practices, knowledge-contesting research must be carried out also within IS.

Conclusion

In our discussion of the two main research approaches, knowledge-expanding and knowledge-contesting, we have argued in favour of increasing knowledge-contestation in IS research. While some may see knowledge-contestation as less productive than confirmatory evidence, it may, in fact, yield significant progressive contributions. Such work can present refuting evidence and problems related to one or more theories and can enrich science by pinpointing limits in existing knowledge and showing where new contributions are needed. It may also present fruitful comparisons or contribute by explaining or predicting phenomena that do not mesh well with prevailing approaches.

We conclude, therefore, that researchers' keen interest in falsifying and correcting a particular research contribution is testament to that contribution's value to scholarship, and any identification of weaknesses gives impetus to further creative work. The benefit is twofold: Knowledge-contestation increases awareness of existing contributions' boundary conditions. Secondly, it accelerates exchange of opinions and interactions between researchers, thereby strengthening the research community. Together with knowledge-expanding research, falsificationism and knowledge-contesting efforts provide a rigorous basis and good direction for new discoveries.

Acknowledgements

Antti Salovaara received financial support for this work from the Academy of Finland (grants 259281 and 298879). Jussi Ilmari Nykänen was supported by the Foundation for Economic Education, the HSE Support Foundation, and the Jenny and Antti Wihuri Foundation. Some ideas in this paper were presented in an early form at the ICIS2014 SIGPhil workshop and at ECIS2015. Also, the authors want to thank Juho Pääkkönen for his critical observations and suggestions as to relevant literature.

References

Aarts, A. A., et al. (2015). Estimating the reproducibility of psychological science. Science, 349(6251), aac4716-1–8.

Ackermann, R. J. (1976). The Philosophy of Karl Popper. Amherst, MA: University of Massachusetts Press.

Andersen, H. & Hepburn, B. (2016). Scientific method. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University.

Baker, M. (2016). Is there a reproducibility crisis? Nature, 533(7604), 452–454.

Banville, C. & Landry, M. (1989). Can the field of MIS be disciplined? Communications of the ACM, 32(1), 48–60.

Baskerville, R. L. & Myers, M. D. (2002). Information systems as a reference discipline. MIS Quarterly, 26(1), 1–14.

Cheikh-Ammar, M. (2018). The IT artifact and its spirit: A nexus of human values, affordances, symbolic expressions, and IT features. European Journal of Information Systems, 27(3), 278–294.

Chiasson, M., Davidson, E., & Winter, J. (2018). Philosophical foundations for informing the future(S) through IS research. European Journal of Information Systems, 27(3), 367–379.

Chen, W. & Hirschheim, R. (2004). A paradigmatic and methodological examination of information systems research from 1991 to 2001. Information Systems Journal, 14(3), 197–235.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340.

Davis, G. B. (2000). 5. In R. Baskerville, J. Stage, & J. I. DeGross (Eds.), Information Systems Conceptual Foundations: Looking Backward and Forward (pp. 61–82). Boston, MA: Springer.

Dennis, A. R. & Valacich, J. S. (2015). A replication manifesto. AIS Transactions on Replication Research, 1(1).

Earp, B. D. & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. Frontiers in Psychology, 6, 1–11.

Eisenhardt, K. M. (1989). Building theories from case study research. Academy of Management Review, 14(4), 532–550.

Farhoomand, A. F. (1987). Scientific progress of management information systems. ACM SIGMIS Database, 18(4), 48–56.

Feinerer, I. (2017). Introduction to the tm package text mining in R. https://cran.r-project.org/web/packages/tm/vignettes/tm.pdf

Feyerabend, P. K. (1981). Problems of Empiricism. Cambridge, UK: Cambridge University Press.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. Journal of Statistical Software, 33(1), 1–22.

Gadamer, H.-G. (1975). Truth and Method. New York, NY: Continuum International Publishing Group.

Geertz, C. (1973). The Interpretation of Cultures: Selected Essays. New York, NY: Basic Books.

Gilles, D. (1993). Philosophy of Science in the Twentieth Century: Four Central Themes. Oxford, UK: Blackwell Publishers.

Glaser, B. G. & Strauss, A. L. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research. New Brunswick, NJ: Transaction Publishers.

Goles, T. & Hirschheim, R. (2000). The paradigm is dead, the paradigm is dead ... long live the paradigm: The legacy of Burrell and Morgan. Omega, 28(3), 249–268.

Gregor, S. (2006). The nature and theory in information systems. MIS Quarterly, 30(3), 611-642.

Grover, V. & Lyytinen, K. (2015). New state of play in information systems research: The push to the edge. MIS Quarterly, 39(2), 271–296.

Hansson, S. O. (2006). Falsificationism falsified. Foundations of Science, 11(3), 275–286.

Hassan, N. R. (2014). Paradigm lost ... paradigm gained: A hermeneutical rejoinder to Banville and Landry's `Can the field of MIS be disciplined?'. European Journal of Information Systems, 23(6), 600–615.

Hassan, N. R. & Mingers, J. (2018). Reinterpreting the Kuhnian paradigm in information systems. Journal of the Association for Information Systems, 19(7), 568–599.

Hassan, N. R., Mingers, J., & Stahl, B. (2018). Philosophy and information systems: Where are we and where should we go? European Journal of Information Systems, 27(3), 263–277.

Hempel, C. G. (1966). The Philosophy of Natural Science. Englewood Cliffs, NJ: Prentice-Hall.

Howell, D. C. (2002). Statistical Methods for Psychology, 5. ed. Belmont, CA: Wadsworth Publishing.

Joachims, T. (1998). Text categorization with support vector machines: Learning with many relevant features. In European conference on machine learning (pp. 137–142). Springer, Berlin, Heidelberg.

Kepes, S. & Thomas, M. A. (2018). Assessing the robustness of meta-analytic results in information systems: Publication bias and outliers. European Journal of Information Systems, 27(1), 90–123.

Klein, H. K. & Lyytinen, K. (1985). The poverty of scientism in information systems. In E. Mumford, R. Hirschheim, G. Fitzgerald, & T. Wood-Harper (Eds.), Research Methods in Information Systems, ch. 8 (pp. 131–161). Amsterdam, NL: North-Holland.

Klein, H. K. & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. MIS Quarterly, 23(1), 67–93.

Kuhn, T. S. (1962). The Structure of Scientific Revolutions. Chicago, IL: University of Chicago Press.

Kuhn, M. (2008). Building predictive models in R using the caret package. Journal of Statistical Software, 28(5), 1–26.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), Criticism and the Growth of Knowledge (pp. 91–196). Cambridge, UK: Cambridge University Press.

Laudan, L. (1977). Progress and Its Problems: Towards a Theory of Scientific Growth. Berkeley, CA: University of California Press.

LeBel, E. P., Berger, D., Campbell, L., & Loving, T. J. (2017). Falsifiability is not optional. Journal of Personality and Social Psychology, 113(2), 254–261.

Lee, A. S. (1989). A scientific methodology for MIS case studies. MIS Quarterly, 13(1), 33–50.

Lee, A. S. (2004). Thinking about social theory and philosophy for information systems. In L. Willcocks & J. Mingers (Eds.), Social Theory and Philosophy for Information Systems, ch. 1 (pp. 1–26). Chichester, UK: John Wiley & Sons.

Lee, A. S. & Hubona, G. S. (2009). A scientific basis for rigor in information systems research. MIS Quarterly, 33(2), 237–262.

Lyytinen, K. & King, J. L. (2004). Nothing at the center?: Academic legitimacy in the information systems field. Journal of the Association for Information Systems, 5(6), 220–246.

Magee, B. (1973). Popper. London, UK: Fontana Press.

Majchrzak, A., Rice, R. E., Malhotra, A., King, N., & Ba, S. (2000). Technology adaptation: The case of a computer-supported inter-organizational virtual team. MIS Quarterly, 24(5), 569–600.

Myers, M. D. (1994). A disaster for everyone to see: An interpretive analysis of a failed IS project. Accounting, Management and Information Technologies, 4(4), 185–201.

Ngwenyama, O. & Klein, S. (2018). Phronesis, argumentation and puzzle solving in IS research: Illustrating an approach to phronetic IS research practice. European Journal of Information Systems, 27(3), 347–366.

Olbrich, S., Frank, U., Gregor, S., Niederman, F., & Rowe, F. (2017). On the merits and limits of replication and negation for IS research. AIS Transactions on Replication Research, 3(1).

Orlikowski, W. J. & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumption. Information Systems Research, 2(1), 1–28.

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. Organization Science, 3(3), 398–427.

Papineau, D. (1996). Philosophy of science. In N. Bunnin & E. P. Tsui-James (Eds.), The Blackwell Companion to Philosophy, ch. 9 (pp. 290–324). Oxford, UK: Blackwell Publishers Ltd.

Paul, R. J. (2002): (IS)3: Is information systems an intellectual subject? European Journal of Information Systems, 11(2), 174–177.

Popper, K. (1959). The Logic of Scientific Discovery. London, UK: Hutchinson & Co.

Risjord, M. (2014). Philosophy of Social Science: A Contemporary Introduction. New York and London: Routledge.

Rowe, F. (2018). Being critical is good, but better with philosophy! From digital transformation and values to the future of IS research. European Journal of Information Systems, 27(3), 380–393.

Rowe, F. & Markus, M. L. (2018). Taking on sacred cows: Openness, fair critique, and retaining value when revising classics. European Journal of Information Systems, 27(6), 623–628.

Salovaara, A. & Merikivi, J. (2015). IS research progress would benefit from increased falsification of existing theories. In Proceedings of the Twenty-Third European Conference on Information Systems (ECIS 2015).

Salton, G. & McGill, M. J. (1983). Introduction to Modern Information Retrieval. New York, NY: McGraw-Hill.

Sanbonmatsu, D. M., Posavac, S. S., Behrends, A. A., Moore, S. M., & Uchino, B. N. (2015). Why a confirmation strategy dominates psychological science. PLOS One, 10(9), 1:1–13.

Scrucca, L. (2013). GA: A package for genetic algorithms in R. Journal of Statistical Software, 53(4), 1–37.

Seddon, P. B. & Scheepers, R. (2012). Towards the improved treatment of generalization of knowledge claims in IS research: Drawing general conclusions from samples. European Journal of Information Systems, 21(1), 6–21.

Seddon, P. B. & Scheepers, R. (2015). Generalization in IS research: A critique of the conflicting positions of Lee & Baskerville and Tsang & Williams. Journal of Information Technology, 30(1), 30–43.

Snyder, L. J. (1997). The Mill–Whewell debate: Much ado about induction. Perspectives on Science, 5(2), 159–198.

Stahl, B.C., Tremblay & M.C., Lerouge, C.M. (2011): Focus groups and critical social IS research: How the choice of method can promote emancipation of respondents and researchers. European Journal of Information Systems, 20(4), 378–394.

Suchman, L. A. (1987). Plans and Situated Actions: The Problem of Human–Machine Communication. Cambridge, UK: Cambridge University Press.

Suppe, F. (1972). What's wrong with the received view on the structure of scientific theories? Philosophy of Science, 39(1), 1–19.

Ting, K. M. (2010). Sensitivity and specificity. In C. Sammut & G. I. Webb (Eds.), Encyclopedia of Machine Learning (pp. 901–902). Boston, MA: Springer.

Trauth, E. M. & Jessup, L. M. (2000). Understanding computer-mediated discussions: Positivist and interpretive analyses of group support system use. MIS Quarterly, 24(1), 43–79.

Tyre, M. J. & Orlikowski, W. J. (1994). Windows of opportunity: Temporal patterns of technological adaptation in organizations. Organization Science, 5(1), 98–118.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. MIS Quarterly, 27(3), 425–478.

Walsham, G. (1995). Interpretive case studies in IS research: Nature and method. European Journal of Information Systems, 4(2), 74–81.

Weisberg, R. W. (2006). Creativity: Understanding Innovation in Problem Solving, Science, Intention, and the Arts. Hoboken, NJ: John Wiley & Sons.

Williams, C. K. & Wynn, Jr., D. E. (2018). A critical realist script for creative theorising in information systems. European Journal of Information Systems, 27(3), 315–325.

Wu, P. F. (2012). A mixed methods approach to technology acceptance research. Journal of the Association for Information Systems, 13(3), 172–187.

Yang, Y. (1999). An evaluation of statistical approaches to text categorization. Information retrieval, 1(1–2), 69–90.

Ylikoski, P. & Aydinonat, N. E. (2014). Understanding with theoretical models. Journal of Economic Methodology, 21(1), 19–36.

Zhang, T. & Oles, F. J. (2001). Text categorization based on regularized linear classification methods. Information Retrieval, 4(1), 5–31.

Appendix A: Example Abstracts from Knowledge-Contesting Publications

The following abstracts exemplify the use of the typology of knowledge-contesting papers presented in Table 1. We use boldface letters and italic type for those parts of the abstracts that demonstrate the relevant characteristics. The types' full names and definitions are found in Table 1.

Type of argument	Type of target	Full abstract and reference
Contested belief & ***Apost. result***	Theory	Research in face-to-face teams shows conflicting results about the impact of behavioral controls on trust| some research shows that controls increase the salience of good behavior, which increases trust while other research shows that controls increase the salience of poor behavior that decreases trust. The only study in virtual teams, which examined poorly functioning teams, found that controls increased the salience of poor behavior, which decreased trust. We argue that in virtual teams behavioral controls amplify the salience of all behaviors (positive and negative) and that an individual's selective perception bias influences how these behaviors are interpreted. **Thus the link from behavioral controls to trust is more complex than first thought.** We conducted a 2x2 experiment, varying the use of behavioral controls (controls, no controls) and individual team member behaviors (reneging behaviors designed to reduce trust beliefs and fulfilling behaviors designed to increase trust beliefs). ***We found that behavioral controls did amplify the salience of all behaviors however, contrary to what we expected, this actually weakened the impact of reneging and fulfilling behaviors on trust.*** We believe that completing a formal evaluation increased empathy and the awareness of context in which the behaviors occurred and thus mitigated extreme perceptions. We also found that behavioral controls increased the selective perception bias which induced participants to see the behaviors their disposition to trust expected rather than the behaviors that actually occurred. DENNIS AR, ROBERT LP JR, CURTIS AM, KOWALCZYK ST, and HASTY BK (2012) Trust is in the eye of the beholder: A vignette study of postevent behavioral controls' effects on individual trust in virtual teams. Information Systems Research 23(2), 546–558.
Prescr. comp. & ***Exposure of insufficiency***	Method	Many business process modeling techniques have been proposed over the last decades, creating a demand for theory to assist in the comparison and evaluation of these techniques. A widely established way of determining the effectiveness and efficiency of modeling techniques is by way of representational analysis**. This paper comparatively assesses representational analyses of 12 popular process modeling techniques in order to provide insights into the extent to which they differ from each other.** We discuss several implications of our findings. ***Our analysis uncovers and explores representational root causes for a number of shortcomings*** that remain in process modeling practice, such as lack of process decomposition and integration of business rule specification. Our findings also serve as motivation and input to future research in areas such as context-aware business process design and conventions management. RECKER J, ROSEMANN M, INDULSKA M, and GREEN P (2009) Business process modeling – A comparative analysis. Journal of the Association for Information Systems 10(4), 333–363.

Prescr. impr. & ***Debate***	Method	In "Generalizing Generalizability in Information Systems Research," ***Lee and Baskerville (2003) try to clarify generalization and classify it into four types. Unfortunately, their account is problematic.*** **We propose repairs**. Central among these is our balance-of-evidence argument that we should adopt the view that Hume's problem of induction has a solution, even if we do not know what it is. We build upon this by proposing an alternative classification of induction. There are five types of generalization: (1) theoretical, (2) within-population, (3) cross-population, (4) contextual, and (5) temporal, with theoretical generalization being across the empirical and theoretical levels and the rest within the empirical level. Our classification also includes two kinds of inductive reasoning that do not belong to the domain of generalization. We then discuss the implications of our classification for information systems research. TSANG EW and WILLIAMS JN (2012) Generalization and induction: Misconceptions, clarifications and a classification of induction. MIS Quarterly 36(3), 729–748.
Exposure of insufficiency & ***Call for contest.***	Practice	Is Information Systems an Intellectual Subject, or an Academic Subject, or an Academic Discipline? And does it matter? Taking the particular perspective of the author, ***this paper discusses these questions with a view to raising debate*** in the Information Systems (IS) community about some of the **current pitfalls in IS, such as inappropriate claiming of expertise in other disciplines, and the widespread malpractice concerning research methods and methodologies.** The issue of change is addressed from the perspective of misuses (Exaggerations, Explanations and Enigmas) of the issue and how this adds to the pitfalls. A particular definition of research is examined, whose inclusiveness provides some potential relief. The conclusions offer hope and a way forward. PAUL RJ (2002) (IS)3: Is Information Systems an intellectual subject? European Journal of Information Systems 11(2), 174–177.
Disagr. analysis	Theory & Practice	This paper offers a study of contradiction in the usage of mobile email. Using qualitative data, the paper identifies mobile email usage patterns that are dangerous, distracting, anti-social and that infringe on work-life boundaries. **Mobile email users were forthcoming in describing these dysfunctional usage patterns, but they made a convincing argument that their mobile devices are highly functional** and allow them to be efficient, to multitask without disruption to others, and to respond immediately to messages, as well as offering them the freedom to work from anywhere. **These dual perspectives on mobile email (dys)functionality are explored through a metaphorical lens**, showing how organisational cultures can reinforce the functional perspective while simultaneously suppressing the dysfunctional view. It is argued that it is important to understand and explore the dysfunctional perspective of mobile email adoption. The paper concludes with a series of questions that challenge organisations to reflect critically on their assumptions about mobile email usage. MIDDLETON CA and CUKIER W (2006) Is mobile email functional or dysfunctional? Two perspectives on mobile email usage. European Journal of Information Systems 15(3), 252–260.