A FRAMEWORK FOR VALIDATING INFORMATION SYSTEMS RESEARCH BASED ON A PLURALIST ACCOUNT OF TRUTH AND CORRECTNESS

John Mingers
University of Kent, j.mingers@kent.ac.uk

Craig Standing
Edith Cowan University, c.standing@ecu.edu.au

Follow this and additional works at: https://aisel.aisnet.org/ecis2019_rp

Recommended Citation
https://aisel.aisnet.org/ecis2019_rp/8
A FRAMEWORK FOR VALIDATING INFORMATION SYSTEMS RESEARCH BASED ON A PLURALIST ACCOUNT OF TRUTH AND CORRECTNESS

Research paper

Mingers, John, University of Kent, Canterbury, UK, j.mingers@gmail.com
Standing, Craig, Edith Cowan University, Australia, c.standing@ecu.edu.au

Abstract

Research in information systems includes a wide range of approaches that make a contribution in terms of knowledge, understanding, or practical developments. In these days of “fake news” and vast amounts of spurious internet content, scholarly research needs above all to be able to demonstrate its validity – are its findings true, or its recommendations correct? However, empirical studies show that discussion of validity in research is often weak. In this paper we examine the nature of truth and relatively correctness in order to construct a validation framework that can potentially encompass all the varied forms of research. Within philosophy, there has been much debate about the nature of truth – is it correspondence, coherence, consensual or pragmatic? In fact, current debates revolve around the idea of a pluralist view of truth – that there may be different forms of truth depending on context or domain. Related to truth is the wider concept of correctness – propositions may be true (and therefore correct) but correctness can also be applied to actions, performances or behavior. Based on these two concepts, we develop a framework for research validity and apply it to a range of research forms including positivist, mathematical, interpretive, design science, critical and action-oriented. The benefits are: i) that a greater and more explicit focus on validity criteria will produce better research; ii) having a single framework can provide some commonality between what at times seem conflicting approaches to research; iii) having criteria made explicit should encourage debate and further development.

Keywords: Action Research, Correctness, Critical Research, Design Science, Interpretive research, Positivist Research, Simulation, Truth, Validation

1 Introduction

Information systems is a wide-ranging discipline involving varied forms of research each with different purposes. There is research aimed at producing knowledge, from a variety of perspectives – positivist (Dubé & Paré, 2003; Straub et al., 2004), interpretive (Klein & Myers, 1999; Walsham, 2006b), critical (Klein & Huynh, 2004; Mingers, 2004) and more; research that aims at producing software or IT/IS artifacts – design science (Hevner et al., 2004); and research that hopes to bring about improvements to organizational problems – action research (Chiasson et al., 2009). These heterogeneous forms of research are carried out in many different ways; based on different and sometimes conflicting assumptions; and often use fundamental concepts such as “information”, “theory”, “causality” or “knowledge” incompatibly.

We are not against the idea of pluralism in IS research at all (Mingers, 2001a; Mingers, 2001c) but we do agree with Lee (Lee, 1991; Lee & Hubona, 2009; Lee et al., 2014; Lee & Hovorka, 2015) that there needs to be some degree of coherence or rigor underlying these multifarious approaches in order to justify and validate the results that end up being published in our journals and used as a basis for affecting peoples’ lives.

In this paper we will investigate one, crucial, element of research – that of truth or correctness which in many ways underlies all the others. Scholarly research in any field, not least IS, aims to produce
knowledge. This immediately begs the question of what exactly is knowledge, and how does it differ from mere belief or opinion? Traditionally, within philosophy, knowledge is said to be “justified true belief” (Gettier, 1963; Pritchard, 2006); that is, it is a kind of belief or opinion about how the world is but one for which we have evidence or warrant, and, essentially, that is actually true whether or not we can in fact determine its truth. This leads to the further question, what exactly is truth for unless we know what truth is, we cannot understand what knowledge is. And, following from this, what is the relation between truth and correctness?

That truth is indeed a goal of IS research has been expressed, for example, by Straub et al (2004) “The purpose of validation is to give researchers, their peers, and society as a whole a high degree of confidence that positivist methods being selected are useful in the quest for scientific truth” (p. 383). However, in most published papers, including that one, the actual nature of truth and how it might be discovered is little discussed. This provides the starting point of this paper – what is the nature of truth?

There is a traditional view within philosophy - the correspondence theory of truth (Lynch, 2001) - that there is a relation of correspondence between beliefs or statements about the world and the way the world actually is. But, to clearly articulate a theory of truth we need to specify its elements: what is it that can have this truth property (known as the “truth bearer”); what is it that could make the truth bearer true (known as the “truth maker”); and what is the nature of the correspondence relation?

There have, however, been many criticisms of the correspondence view of truth, particularly in terms of its realist view of the external world, and this led to a number of alternatives. For example, coherence theory which evaluates a belief in terms of its coherence or consistency with other well-attested beliefs (Walker, 1989); pragmatism which focusses on long-term success in practice (James, 1976; Peirce, 1878); or consensus theory (Habermas, 1978) which sees truth as that which a relevant community of enquirers agrees about. A more radical approach, known as deflationism (Quine, 1992; Strawson, 1950), suggests that actually truth has no substantive nature to be explained, and that it is really just a linguistic pseudo-problem.

More recently, in the face of the stand-off between these competing positions, a new approach has been developing within philosophy that aims to retain the idea that truth is a substantive concept, and some form of realism about the relation to the external world, whilst accepting the criticisms of standard correspondence theory. This approach involves a pluralist view of truth: “truth is one and truth is many” (Lynch, 2009) – there are generic characteristics of truth but these may be realized differently in different domains (Lynch, 1998; Pedersen & Wright, 2013b). For example, in the physical domain one might hold a correspondence view while in the mathematical domain one might have a coherence view.

Whilst truth may be a defining characteristic of knowledge, as we saw above not all IS research concerns purely knowledge - design science, for example, aims to produce effective software or artifacts, and action research aims to solve problems in organizations. In these domains it may not actually be appropriate to talk about truth but rather the related term correctness (Engel, 2013; Thomson, 2008). It seems more satisfactory to say that a computer system works “correctly” rather than “truly”. In many areas truth and correctness are equivalent – if a belief is true then it will also be correct, while a belief that is incorrect would thereby be false. But correctness is a wider term than truth in that it applies to things other than beliefs or propositions, for example actions or procedures.

The purpose of this paper is to develop a general framework, based on the concepts of pluralistic truth and correctness, which can be applied across all areas of research in IS. Essentially, this will specify criteria for evaluating the rigor and validity of the research whatever its particular paradigm or method. This is akin to the proposal of Lee and Hubona (2009) that the fundamental logical laws of modus ponens and modus tollens can be applied across many research methods to produce more rigorous research.

In the first section of the paper we explain the pluralist view of truth and its relation to correctness. In the next section we develop a general model of truth and correctness and their links to research validity. Then in the third section of the paper we apply the framework to a variety of research approaches –
positivist statistical analysis, mathematical modelling and simulation, interpretive research, critical research and finally action research. In the final section we discuss the benefits of this framework.

2 Theories of Truth and Correctness

Pluralist theories represent a new development in response to the stand-off between the different truth theories. Generally, many philosophers do wish to maintain a substantive version of truth and do see correspondence theory as the most intuitive approach and so, in response to the criticisms of correspondence, they have developed the general idea that there may be different versions of truth dependent on the domain of knowledge concerned. There are three possible approaches (Pedersen & Wright, 2013a) – strong pluralism which sees only many versions of truth with no overarching unity to them, a position not held by many. Weak pluralism, which holds that truth is one and many – there is a general conception of truth, often characterized in terms of a number of properties that all forms of truth must have (called plattitudes or truisms), which is realized differently in different domains. And, third, what could be called correspondence pluralism which maintains that there is only correspondence theory but this itself can be differentially realized.

Putnam (1994) was perhaps the first to suggest that there are many ways in which propositions can relate to reality and that therefore the word “true” may be realized differently depending on whether we are talking about physical reality, mathematics or morality. Lynch (1998) followed up with a functionalist approach asking what are the functions of truth – e.g., objective, correct to believe, and aimed at facilitating enquiry – suggesting that these functions could be met in different ways. Pedersen and Wright (2013b) provides a state-of-the-art view of alethic pluralism (Smith, 2015).

The main point for this paper is the argument that truth is a vital component of knowledge, and thus of great importance for IS research, and further that a pluralist view of truth is most compatible with the many domains and approaches of IS research.

While truth may be seen as the constituting feature of valid knowledge, there are domains of IS in which it does not seem so appropriate, at least in any but its pragmatist forms. Examples are design science, which is concerned with the design of effective and efficient computer systems, or action research where the aim is effective problem-solving in organizations. It does not seem appropriate to talk of a computer system being “true” but we can say that it is working “correctly”. Correctness is clearly related to truth: Horgan (2001) talks of truth as “semantic correctness” and Floridi (2011b) discusses a “correctness theory of truth”. In this section we will explore the notion of correctness with a view to seeing if it might be a more appropriate term for, at least some, information systems. From its dictionary definition, correctness can mean three things – true or conforming with the facts; in accordance with accepted standards; and free from error. The third is essentially the obverse of the first two so we are left with two, the first as a synonym for truth, at least as correspondence, and the second wider meaning as conforming to some accepted or agreed standards or norms (Finlay, 2010).

Many things may be said to be correct or incorrect: mental states such as believing or knowing; actions or performances such as a statistical analysis or a logon procedure; representations such as a map or a computer model; information; a move in a game; or an English sentence. What is it these all have in common and how does this relate to truth?

Thomson (2008), in a major work on norms, claims that correctness has two aspects. First, correctness is always relative to the kind of thing it is applied to. A map of England may only be correct as a map of England, not just correct in general. Correct is an attributive adjective like “good”, it is always relative to a particular kind, K. The kind, K, fixes what properties the thing needs to have to be correct, essentially an exemplar of what the kind (or set) K would be. Note that some kinds of things do not have such exemplars, e.g., pebbles or shades of grey, and so cannot be correct or incorrect.

Second, with some kinds there are, in fact, two different ways in which something may be said to be correct. Consider an action such as asserting a proposition, “the computer system is down this morning”. The first form of correctness concerns the manner in which the assertion is made – is it correct English such that it makes sense and is comprehensible? This Thompson calls “internal-correctness”
or “i-correctness”. Many other actions can be said to be correct or incorrect in this sense, for example performing a symphony or spelling a word.

But with an assertion there is a second form of correctness, that is, whether the proposition it asserts is correct, i.e., is the computer system actually down? Thompson calls the “external-correctness” or e-correctness. In the case of an assertion, e-correctness is the same as truth – is the proposition true? This is also the case for other speech acts such as describing, answering, reporting or explaining.

“In short, an instance of one or other of these speech act kinds is internal-correct just in case the speaker carries out the speech enterprise correctly, and external-correct just in case the propositional content of his act is true” (Thomson, 2008, p. 99)

One further aspect of these forms of correctness is that i-correctness is normative – it is the manner in which something ought to be done; but e-correctness is descriptive – that something is true or not, or successful or not. Also, the two are independent of each other – we can have an i-correct assertion of something false, or an i-incorrect assertion of something true (if we could understand it).

There are many other kinds of actions that have both forms, especially purposeful actions. Here i-correctness is the manner in which the action is performed and e-correctness is whether the goal or purpose is in fact achieved. However, there are actions that do not have some external relation, such as performing a symphony or spelling a word which can only be i-correct.

We can also consider mental states such as believing, expecting or assuming which can be said to be correct. In these instances, Thompson argues that there can only be e-correctness as believing or expecting is simply a mental state, it cannot be performed in a more or less correct way. And again, with these kinds the e-correctness is essentially simply truth. Engel (2013) considers the case of beliefs (which could of course be manifested in terms of propositions or statements) and agrees that the e-correctness is in fact truth – beliefs aim for truth and are e-correct when they are true. However, he argues that they do have i-correctness which is our evidential reasons (warrants) for believing them, which becomes normative in the sense that we should believe things for which we have strong evidence, whether or not they are in fact true.

3 A Framework for Correctness and Truth

This section will construct a framework for combining correctness and truth based on the ideas in the above discussion. Thomson’s distinction between e- and i-correctness is fundamental and is the same as distinctions in other fields. In particular, it is essentially the same as Horgan’s (2001) distinction between the semantic standards of a domain (i-correctness), and the correspondence relations between assertions and the way the world is (e-correctness). For Horgan, all (discursive) domains have i-correctness, assuming they are coherent and well-formed domains, while some have direct e-correctness and some only indirect e-correctness.

It is also the same as the distinction made within the context of model and research validation (Boudreau et al., 2001; Kleijnen, 1995; Lee & Baskerville, 2003; Lee et al., 2014; Lee & Hovorka, 2015; Lukka & Modell, 2010; Sargent, 2013; Venkatesh et al., 2013). Although different terms may be used, essentially there are two distinct stages to model validation which we will call verification and validation. Verification concerns the internal structure of the model, whether it is a statistical model, a simulation model, or indeed a piece of interpretive research. Validation concerns the external aspect of the model – whether it adequately represents that which it a model is of. Thus, verification is i-correctness and validation is e-correctness. Within statistics and measurement theory, these are often termed precision (the degree of replicability of repeated measurements) and accuracy (the closeness of the measurement to the quantity’s true value).

We will briefly explain the framework here and then look at particular parts in more detail in the context of information systems. Here, we look at how e- and i-correctness can be applied to different kinds of things or actions that are relevant to IS such as representations, procedures, information and artefacts.
Truth and Correctness

Representation cover a range of things that may have different purposes – a picture or photo may just describe something; a theory may explain why something happened; a simulation model may try and replicate behavior. Whether they are e-correct depends both on their correspondence but also the purpose – a map of the London underground is not correct in terms of walking the streets.

Procedures are specified steps that need to be undertaken to achieve a result. A mathematical proof does not correspond to anything but, starting from axioms and following logical rules it can generate a conclusion. It is i-correct in terms of adhering to the rules and potentially other criteria such as elegance or simplicity. It is e-correct in demonstrating the result. Logging into an account is e-correct if it succeeds. It may not be i-correct if done in the wrong way, for example by hacking.

Information is a disputed phenomenon with different conceptions (McKinney & Yoos, 2010). Some, such as Floridi (2011a) and Mingers and Standing (2014), argue that information is both objective and true in which case its i-correctness is its semantic meaningfulness and its e-correctness is its truth in some form. Others (Checkland & Holwell, 1998b) argue that it is subjective and not necessarily true in which case it is difficult to understand what might be correct information.

There are many artefacts, especially humanly produced, which may have the property of “being a good K” if there are standards or properties which exemplars of such a kind exhibit. A good toaster produces evenly browned toast; a good information system produces accurate, timely, relevant information in an easy to use way. For these, the e-correctness involves meeting the specified goodness criteria, which may often be in terms of functions. The i-correctness concerns the form of the artefact – is it aesthetically pleasing? Is it robust and easy to use? Is it efficient? Is it economical? In these cases there could be disagreements about which properties were part of the function and which the form.

In summary, correctness has two aspects – internal and external – with internal being normative, some example of a K is as it should be; external is descriptive, some example of a K bears an appropriate relation to an external state of affairs or a goal. Correctness is an attributive adjective that may be applied to a wide range of entities or actions. In some cases, e-correctness is essentially the same as truth, although not necessarily the correspondence theory, in other cases it is not. Truth is thus a subset of correctness.

4 Applying the Framework to Information Systems Research

There are two distinct questions concerning truth – what is it, i.e., what is its nature? And how do we discover it, i.e., how do we tell true theories from false ones? We may call these the definitional and the justificational questions. The first part of the paper has been concerned with the first question, but we now move to the second, practical question – how do we justify our research conclusions?

Developing a framework that could potentially apply across all research approaches in IS is clearly a tall order, and within this paper we only have space to cover what we consider to be the most prominent approaches (Galliers et al., 2006; Galliers & Land, 1987; Grover & Lyytinen, 2015; Hirschheim et al., 1997; Hirschheim & Klein, 1989; Hirschheim & Klein, 2006; Lee & Hubona, 2009; Nissen et al., 1991; Orlikowski & Baroudi, 1991; Venkatesh et al., 2013). In particular, positivist, interpretive, critical, action research and design science – but we recognize that there are others which we have not been able to address in this paper, for example NeuroIS (Reidl, 2018) or feminist research (Wilson, 2016), and of course new areas may develop. As we have already said, we would not wish our framework to be seen as rigid and would welcome contributions from other approaches. Even within a particular approach, we have had to be selective in our choice of perspectives – a whole paper or book could be devoted to validity within a particular domain. What we are concentrating on is trying to show that there are potential underlying commonalities across varied research approaches, and that an awareness and consideration of these could potentially contribute to higher quality research.

The framework as a whole is presented in Table 1. The succeeding sections of the paper explain how Table 1 is developed from the various research approaches. Columns 2 and 3 concern i-correctness and e-correctness respectively. We can note from the discussion above that i-correctness is essentially the same as verification, precision or formative validity (terms used in various fields) and is norma-
tive. *e*-correctness is essentially the same as validation, accuracy or summative validity and is descriptive. As we consider each research area, we will look at the practice and terminology of that area and then endeavour to show that it can be incorporated within our correctness framework. Column four shows particular forms of truth that are relevant to the research approach. For reasons of space, some approaches will only be discussed briefly.

### 4.1 Justifying positivist research

Within IS, there is a considerable literature devoted to justifying empirical, positivist research. There is a strong line of good practice recommendations developing, in the main, from Cook and Campbell’s (1979) treatise on quasi-experimentation. This was picked up by Straub (1989) and further developed in theoretical (Bagozzi, 2011; Im & Straub, 2015; MacKenzie et al., 2011b; Shadish et al., 2002; Straub et al., 2004; Venkatesh et al., 2013) and empirical studies (Boudreau et al., 2001; King & He, 2005). In *theory*, this approach ties in directly with our correctness framework; whether it does in practice is more debatable. We critique it for being excessively narrow and technically focused, ignoring the more general issues of internal and external validity. We then describe an alternative approach developed by Lee and Hubona (2009).

The original work (Campbell, 1957) distinguished between internal and external validity (cf internal and external correctness above) while Cook and Campbell (1979) added construct validity and statistical validity. However, these forms of validity take on quite specific meanings. And only refer to specific forms of statistical research in which there are some underlying latent, subjective constructs, and relationships between them, which are then operationalized in terms of particular quantitative measures and an instrument to collect data. The instrument is assumed to be some form of questionnaire or perhaps experiment. The point is that there are other forms of positivist research beyond surveys and statistics, popular as they may be.
<table>
<thead>
<tr>
<th>Form of research</th>
<th>i-correctness</th>
<th>e-correctness</th>
<th>Relevant forms of truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positivist research: statistical</td>
<td>Content</td>
<td>Comparison of predictions with actuality,</td>
<td>Correspondence between constructs and concepts, and between results and actuality;</td>
</tr>
<tr>
<td></td>
<td>Construct</td>
<td>e.g., hold-out samples or cross-validation</td>
<td>Coherence of constructs</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>Elimination of alternative explanations</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Statistical validity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positivist research: simulation and mathematical modelling</td>
<td>Model comparison</td>
<td>Predictive validation</td>
<td>Correspondence between results and actuality;</td>
</tr>
<tr>
<td></td>
<td>Extreme conditions</td>
<td>Historical data</td>
<td>Coherence of model;</td>
</tr>
<tr>
<td></td>
<td>Degeneracy tests</td>
<td>Event validation</td>
<td>Pragmatism and consensus about operational validity</td>
</tr>
<tr>
<td></td>
<td>Sensitivity analysis</td>
<td>Face validation</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Replications</td>
<td>Graphical animation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Credibility</td>
<td></td>
</tr>
<tr>
<td>Interpretive research - emic</td>
<td>Confirmability</td>
<td>Credibility/ authenticity in the eyes of the</td>
<td>Correspondence between description and participants’ world;</td>
</tr>
<tr>
<td></td>
<td>Dependability</td>
<td>participants/ subjects</td>
<td>Consensus about authenticity of results</td>
</tr>
<tr>
<td></td>
<td>Consistency</td>
<td>Interpretive coherence</td>
<td>Coherence of interpretations</td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>Pragmatic coherence: consistency between talk and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transgressive</td>
<td>action</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Performativity – the stranger test</td>
<td></td>
</tr>
<tr>
<td>Interpretive research - etic</td>
<td>Theoretical validity – comparison of results</td>
<td>Plausibility in the eyes of the research community</td>
<td>Consensus about plausibility of theoretical interpretation</td>
</tr>
<tr>
<td></td>
<td>with theory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design research</td>
<td>Methodological validity</td>
<td>Efficacy that the system works</td>
<td>Pragmatism and consensus about operational success</td>
</tr>
<tr>
<td></td>
<td>Efficiency</td>
<td>Effectiveness that it does the right thing</td>
<td>Coherence of design method</td>
</tr>
<tr>
<td></td>
<td>Ethicality</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Elegance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action research</td>
<td>Declaration of theory and methodology</td>
<td>Effectiveness that the problematic issue has</td>
<td>Pragmatism and consensus about operational success</td>
</tr>
<tr>
<td></td>
<td>Active application of theory and participation</td>
<td>been alleviated</td>
<td>Coherence of design method</td>
</tr>
<tr>
<td></td>
<td>in situation</td>
<td>Justification of theoretical contribution</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Recoverability</td>
<td>Generalizability to other contexts</td>
<td></td>
</tr>
<tr>
<td>Critical research</td>
<td>Critical perspective and use of critical</td>
<td>Enlightenment of individual participant</td>
<td>Correspondence of theory to social mechanisms</td>
</tr>
<tr>
<td></td>
<td>theories</td>
<td>Change of social arrangements</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Participative research design</td>
<td>Judged by the participants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analysis of underlying, coercive mechanisms</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Comparison with other contexts</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Researcher reflexivity</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 1 The correctness framework for selected IS research approaches*
Moreover, most of the discussion concerns the fairly technical issues of instrument validity and statistical validity rather than the more general ones of internal and external validity. Again, these latter concepts are defined quite narrowly, and perhaps counter-intuitively, in this approach – internal validity only concerns the possibility of there being other causal relationships, i.e., explanations, that are not included in the model. In many ways this seems like an external factor since it makes direct reference to the external worlds beyond the model, and cannot really be dealt with from a purely internal perspective.

Equally, the idea that external validity primary concerns the extent to which the results can be generalized to other populations and settings (King & He, 2005) seems mistaken. As Reichardt (2011) argues, the fundamental purpose of validation is to assess the truth of the inferences made in the model, it is not particularly concerned with how wide or narrow those inferences are. Whilst generalizability is an important and much debated (Lee & Baskerville, 2003; Lee & Baskerville, 2012; Seddon & Scheepers, 2015; Tsang & Williams, 2012), characteristic of a statistical finding, it is a separate issue from the question of validity.

Finally, this approach makes almost no reference to the fundamental issue of designing the study in the first place in such a way that the eventual results will form valid answers to the research questions. It takes for granted the development of appropriate constructs, hypotheses of the relationships between them, and the initial determination of the appropriate measures and data collection instrument and yet arguably these factors are much more important for overall validity or correctness of the research findings that is instrument validity (Johnston & Smith, 2010). As the empirical research shows (Boudreau et al., 2001; Jones, 2004; King & He, 2005; Straub, 1989), in many cases of papers published in leading journals even the most basic aspects such as describing and justifying the methods of data collection and analysis are absent.

Lee and Hubona (2009) provide an alternative approach to validation. Their primary aim is to produce a framework that can apply to both qualitative and quantitative research based on the logical forms of argument – *modus ponens* (p implies q; p; therefore q) and *modus tollens* (p implies q; not q; therefore not p) which they call the MPMT framework. They distinguish between formative validity and summative validity (taking these terms from education research) and suggest that much IS research involves formative validity but little summative validity. Formative validity is the process of forming or producing the theory or inference and so this type of validity concerns the extent to which the research has correctly followed an accepted procedure. Summative validity is a characteristic of the sum result or product of the process that has been followed. It involves comparing the consequences or predictions of the theory with observed evidence according to the logic of *modus tollens*. If a consequence or prediction of the theory cannot in fact be observed then the theory does not have summative validity and could potentially be rejected. Lee and Hubona show that this approach can apply to quantitative research, qualitative research, and even systems design – a system may be designed according to an accepted systems design methodology and yet still fail to meet its aims. They also argue that of the two, summative validity is more important than formative validity even though in practice it is seldom demonstrated, particularly in positivist research.

In order to generate summative validity in statistical-type research (which is the content of this particular section), Lee and Hubona argue that statistical validity in the sense of significance tests or confidence intervals for various fitted parameters which constitute the hypothesized relationship is not sufficient. This is actually part of formative validity. As well as this it is necessary to test the theory’s predictive capabilities on out-of-sample data points using hold-out samples or cross-validation. We should note, however, a very common problem pointed out by Lee and Hubona – the fallacy of affirming the consequent. If we find that the predictions are in fact correct, does that prove or confirm the theory? The answer is unfortunately no, since there could always be some other explanation which actually accounts for the results. This can be expressed in logic – p implies q; q; therefore p – which is not a valid inference. This point relates to Straub’s issue of internal validity which concerns alternative explanations. We would suggest that this is mis-named and is really external or summative validity –
as well as trying to confirm the predictions one also need to actively try and eliminate alternative explanations (cf. the section on critical realism below).

In comparison with the correctness framework, it seems clear that Lee and Hubona’s approach fits it very well. Formative validity is essentially the same as internal correctness, while summative validity is the same as external correctness and the two are related but independent. We would hope that formative validity (i-correctness) would lead to summative validity (e-correctness) but it is not guaranteed; while it would be possible to reach summatively valid conclusions even through research that was formatively weak. These conclusions are summarized in Table 1 where formative validity is shown in terms of four of Straub’s terms and summative validity is developed from Lee and Hubona. In terms of truth, the primary form is correspondence between the model and the actual situation or, less strongly, just between the results and the data if the model is primarily predictive rather than explanatory. There is also coherence truth between the various constructs.

There are other forms of positivist research such as simulation or mathematical models (Galliers & Land, 1987) which are included in Table 1 but we do not have space to discuss.

4.2 Justifying interpretive research

Qualitative or interpretive research is a much more complex area in terms of validation and truth (Cole & Avison, 2007; Goldkuhl, 2012; Lee, 2018; Myers & Avison, 2002a; Myers & Avison, 2002b; Walsham, 2006a). First, there are a wide variety of methods that differ significantly in their ontological and epistemological assumptions from relatively objective post-positivist approaches such as grounded theory (Glaser & Strauss, 1967) or “subtle realism” (Hammersley & Atkinson, 1995) through textual analyses such as semiotics (Mingers & Willcocks, 2014; Mingers & Willcocks, 2017), hermeneutics (Cole & Avison, 2007; Myers, 2004; Ricoeur, 1981) or discourse analysis (Cukier et al., 2009) to highly subjectivist or constructivist ones such as phenomenology (Boland, 1985; Introna & Ilharco, 2004; Mingers, 2001b; Schutz, 1972) or post structuralism (Dreyfus & Rabinow, 1982; Foucault, 1980). Second, there is debate even within methods as to the possibilities of some form of external, or even internal, validation at all.

We would see our framework as potentially helping with this process by exploring underlying commonalities between research genres. Where we would differ, perhaps, is in believing that research in each genre can benefit from some guidelines as to what constitutes high quality in that genre.

Papers in IS that provide guidance on doing interpretive research generally fail to discuss validity. For example, Klein and Myers’ (1999) authoritative paper provides seven principles that should be applied in interpretive research (primarily limited to hermeneutics) but say little about validation principles. Similarly, Sarker et al (2013) review empirical studies and also offer guiding principles but do not discuss validation.

Interpretive research begins from the position that its object of study, whether it is actions, texts, beliefs or discourse, is socially constructed by the actors involved. Therefore, its primary task is to gain an authentic understanding (verstehen) of that meaning in the terms of the actors who produce it rather than in terms of theory, or the interpretations of the researchers. For some researchers, e.g., ethnographers, that is sufficient whereas others would want to go on and interpret the results and perhaps relate them to theory.

Moving to possible validity criteria, Maxwell (1992) suggested three forms of criteria based partly on different stages of the project. First is descriptive validity which solely concerns the quality of the data production process: that it is comprehensive, accurate and not subject to dispute (although the participants may themselves hold different and perhaps contradictory viewpoints, these should be faithfully recorded). The second is interpretive validity which goes beyond merely recording events, actions and discourse to generating interpretations of it, but still from the participants’ point of view not the researchers. This has been described as an “emic” viewpoint rather than an “etic” viewpoint (Headland et al., 1990), an insider rather than an outsider one. Interpretive validity involves the faithfulness or authenticity of the account to those involved, but even here the boundaries are blurred because actors are not always fully transparent to themselves and, as Giddens (1979) emphasizes, there are often un-
known conditions and motivations for action. The third form of validity is theoretical validity which does move away from an emic account to an etic one. The researcher aims to develop theories that may explain the particular observed behaviors. Theory could come in two directions, from within as in the case of grounded theory where the theory is developed internally from the research material, or from without as theory that already exists is applied to explain the situation.

In terms of validity, Sandberg (2005) proposes aspects: communicative validity, pragmatic validity and transgressive validity. Communicative validity aims for coherence – coherence of understanding between researcher and participants; coherence in a hermeneutic sense within the interpretation of the material; and coherence between the researcher and other researchers and participants within the practice. Pragmatic validity aims for consistency between what the participants say and what they actually do. This is necessary because participants accounts are often not open and honest, being mediated by politics, storytelling, social codes or impression management (Alvesson, 2003). Transgressive validity is orthogonal to the other two in searching for ambiguity, complexity and contradiction.

From our validity and correctness point of view we wish to have a classification which is quite general and compatible with many of the particular approaches. We would therefore make one main distinction, that between emic and etic research. In emic research (which must necessarily come before etic) the primary concern is with reproducing, in as authentic and rich a manner as possible, the way of life of the actors within a situation of interest, in their own terms. This includes both descriptive and interpretive validity in Maxwell’s model. Some research, for instance descriptive ethnography, may choose to stop there but increasingly there is a view that even ethnographic research should move towards some form of explanation (Kakkuri-Knuuttila et al., 2008; Lukka & Modell, 2010). This moves towards an etic account, which will either be based on or generate theory; will be expressed in the researchers’ language; and must be plausible to the research community. This is congruent with Cole and Avison’s (2007) use of the trustworthiness of the research process and the truthfulness of the results. Based on our correctness framework, both of these emic and etic research will have both i-correctness (formative) and e-correctness (summative) validity criteria as shown in Table 1.

In terms of truth criteria in Table 1, correspondence does play a part in that the research aims to authentically mirror the participants’ world even if that is to be judged by the participants. Coherence is also relevant in assessing the overall interpretation, and consensus as well in terms of the participants judging authenticity and other researchers in judging the plausibility of the results.

4.3 Justifying system design: design science and action research

We are considering these two somewhat different approaches together for two reasons. First, they share purposes that make them different from the research approaches we have so far considered – that is, they both aim to bring about beneficial change in organizations, one through the development of an IT artifact, the other through problem solving which might include developing artifacts. Second, because these similarities have already been noted in the literature (Baskerville et al., 2009; Järvinen, 2007; Lee, 2007; Sein et al., 2011; Wieringa & Morali, 2012) although Livari and Venables (2009) suggest the similarities may not be deep. But, from the point of view of validation they do have significant commonalities.

Design science is concerned with producing new and innovative IT artifacts to solve organizational problems (Hevner et al., 2004) although Lee et al (2015) point out that it should be the IS artifact not just the IT artifact. As Hevner et al. (2004, p. 78) note, “design is both a process (set of activities) and a product (artifact) – a verb and a noun” and this concords with the two aspects of correctness in our framework – i-correctness (formative) as conforming to a process or methodology, and e-correctness (summative) as successfully achieving its goal or purpose (in this case in terms of its organizational stakeholders). Various proposals have been made for a design science methodology (Gregor & Jones, 2007; Hart & Gregor, 2010; Vaishnawi & Kuechler, 2008) that have been integrated by Peffers et al (2007) into the following general stages (we have included Hevner et al’s (2004) guidelines in parentheses where they fit in):

- Problem identification and motivation (2. Problem relevance)
• Define objectives for solution (6. Design as a search process)
• Design and development (1. Design as an artifact; 5. Research rigour)
• Demonstration (4. Research contributions; 7. Communication of research)
• Evaluation (3. Design evaluation)

Note that here the 5th step is actually evaluation, in particular evaluating whether the artifact does indeed meet the objectives that were required of it (summative or e-correctness).

Venable et al (2012) have developed a detailed and comprehensive framework of different methods for assessing both formative and summative validity. It is based initially on the 5E’s approach to evaluation (Checkland & Scholes, 1990) – Efficacy, Efficiency, Effectiveness, Elegance and Ethicality. Of these, Efficacy and Effectiveness primarily concern summative validity and the other three concern formative validity. Efficacy is the extent to which the artifact performs as it is designed to do whilst effectiveness is the extent to which performing those tasks is actually successful in the organizational context - does it do what it is supposed to do, and is that the right thing to do?

In terms of formative validity, first was the artifact designed according to a rigorous methodology of whatever kind? Then there are questions as to whether it was developed with an economical use of resources (efficiency), according to ethical principles, and ultimately elegantly and aesthetically (the Mac vs the PC)? One would like to think that formative validity would lead to summative validity but unfortunately the high number of IS failures that still happen (Dwivedi et al., 2014; Georgiadou & George, 2006) shows that this is not the case. These forms of validity are shown in Table 1. In relation to truth, clearly the primary one is pragmatism in terms of evidence that the system works effectively, and consensus about that (sometimes there can be disagreement about whether a system is successful or not. We can also look for coherence in terms of the design method used and its results.

Moving to action research, although as we have shown there are many who consider the two can be intimately linked, in terms of validation we will deal with them separately although Wieringa and Mohrah (2012) actually define the concept of “technical action research” as a specific method for evaluating design science. Action research (AR) (Checkland & Holwell, 1998a; Eden & Huxham, 1996) has a long history dating back to Kurt Lewin (1946) and comes in many varieties including action learning (Revons, 1993), action science (Argyris et al., 1985) and participatory action research (Whyte, 1991). It has been recommended for research in information systems (Baskerville & Wood-Harper, 1998; Baskerville, 1999; Chiasson et al., 2009; Davison et al., 2004). Given this variety we will have to consider a very broad description of AR as being constituted by several elements performed in a cyclical manner:

• Initial recognition of problematic issue and entry of researcher
• Declaration of theories and methodologies thought to be relevant
• Undertaking action to improve the situation as both participant and researcher (in participatory AR the actors are also seen as participant researchers)
• Evaluate results in terms of improvement to the particular organizational situation
• Evaluate results in terms of the theory/methodology used and disseminate the learning

Checkland and Holwell (1998a) also emphasize the importance of “recoverability”, that is explicit documentation of the process followed and decisions made which will help generate the theoretical lessons as well as allowing later critical scrutiny. In terms of the correctness or validation of the process (Baskerville & Wood-Harper, 1998; Checkland & Holwell, 1998a; Eden & Huxham, 1996), we can consider the i-correctness (in Table 2) in terms of the extent to which the AR process was followed – declaration of theory, application of theory and participation, and recoverability; and the e-correctness in terms of two distinct criteria – the success in terms of resolving the problem, and the learning and development of theory which may be applicable elsewhere. It is the latter which mainly distinguishes action research from pure consultancy. Forms of truth are similar to design research – pragmatism in terms of whether the action was successful; consensus about that success; and coherence between the different parts of the research and the results.
4.4 Justifying a critical approach

In this section we will cover a range of explicitly critical approaches mainly based on the work of theorists such as Bourdieu (Kvasny & Keil, 2006), Foucault (Willcocks, 2004) and Habermas (Brocklesby & Cummings, 1996; Howcroft & Trauth, 2005; Klein & Huynh, 2004; Myers & Klein, 2011) (Mingers, 1980); as well as critical realism (Johnston & Smith, 2010; Mingers, 2004; Mingers et al., 2013), and critical versions of interpretive approaches such as critical ethnography (Myers, 1997) and critical discourse analysis (Cukier et al., 2009). A critical approach, or the idea of critique, has two lineages one traceable to Kant and one to Marx (Cecez-Kecmanovic, 2011; Mingers, 2000). Kantian critique concerns the limits of our knowledge and research methods while Marxist critique concerns the oppressive nature of society. Generally, both are involved in a critical approach. However, a critical approach is not primarily about research methods but about attitude and values (Cecez-Kecmanovic, 2011; Morrow & Brown, 1994). In other words, there are not specific critical research methods, rather traditional methods, both quantitative and qualitative, are used but with a critical intent. Alvesson and Deetz (2000) provide perhaps the most general framework for doing critical research that involves three stages:

Insight – hermeneutic understanding and the archaeology of knowledge. This stage involves gaining knowledge and appreciation of the situation of interest using a range of ordinary research methods, both qualitative and quantitative. But it will be guided by explicitly critical attitudes and values and will view the subjects as active participants in the research rather than passive objects.

Critique – deconstruction and the genealogy of knowledge This stage involves using varied critical theories and constructs to uncover and reveal the often hidden or suppressed mechanisms that distort the participants’ understandings of the situation and act so as to maintain this power differential.

Transformative redefinition – enlightenment and emancipation. This stage aims at enlightening participants to the true nature of the situation and thereby helping them to bring about change. It also reflexively develops social theory. Final validity is in the judgement of the participants.

I-correctness concerns the process of research and analysis itself, and whether it has properly followed the research steps, primarily from the Insight and Critique stages. We have selected a subset of these, since many overlap, to form i-correctness in Table 1. E-correctness concerns the actual success of the critical analysis in terms of the change of consciousness of the participants, and change of oppressive social arrangements. This is ultimately to be judged by the participants themselves rather than the researchers. These criteria come from the Transformative redefinition and again we have taken a subset for inclusion in Table 1.

5 Conclusions

The implications of our analysis of truth and correctness are as follows. Research is often carried out and published with little explicit regard for its validation (Boudreau et al., 2001; Gonzalez & Sol, 2012; Jones, 2004; King & He, 2005; MacKenzie et al., 2011a; Straub, 1989; Wieringa & Morali, 2012). For research to make a genuine contribution, either to knowledge or to practice, and to be published in journals or lead to organizational change, every effort must be made to demonstrate that the results are valid, that is believed to be true or correct.

As we have demonstrated, there are two fundamental and distinct characteristics – internal correctness and external correctness, also known as verification and validation or formative and summative. The first is normative and concerns the way in which the research should be carried out; the second is descriptive and concerns the relationship of the research findings to the external context. We have shown within the framework criteria for both of these across a wide range of research approaches. We agree with Lee and Hubona (2009) that much less attention is paid to e-correctness than i-correctness and yet arguably the latter is more important.

It is tempting also to align these two with rigor and relevance. Certainly i-correctness concerns the rigor of the research, and e-correctness is at least related to its relevance although there may be very abstract research which does not, at the time, seem to have much direct relevance but one only needs
to think of the laser or prime number theory to see how such research may later come to have huge relevance. In some ways our work can be seen as similar to that of Rosemann and Vessey (2008) who are concerned with improving the relevance of IS research. They have produced applicability checklists to assess the practical relevance of research through focus groups and nominal group technique. Our framework can be used similarly to improve the rigor of research, but clearly the e-correctness component does already include relevance in methods like design science or action research that specifically aim for change but could also include this as a criteria in other forms of research if desired.

We believe that it is important that we have produced a framework that encompasses a wide range of methods. Too often, different research methods are seen to be in competition or even in conflict with each other. The framework demonstrates that they can all be seen as sharing some very basic characteristics, and are all ultimately part of the same human drive to better understand and improve the world. By focusing explicitly on both internal and external correctness we hope that the results of research will be more informative and effective.

The way in which research needs to develop is that it needs to explicitly consider both these aspects of correctness at all stages – the design of the research, its operationalization, and its description and dissemination. We hope that the framework can provide guidance for researchers to consider in designing their research, and for referees and editors to look for when evaluating submissions or grant applications.

In terms of limitations and further research, we note the following. The next stage is to apply the framework to a range of published empirical research to test its utility. This has been done but cannot be included here for space reasons. It is available in a working paper from the authors.

The framework could be developed to include further research approaches that we have not considered, for example theoretical computer science or neuroIS or feminist research or multimethodology. It could also be developed internally to provide a greater discrimination within approaches, especially the interpretive area where it may be found useful to have different criteria for, say, hermeneutics, phenomenology, textual analysis or semiotics. The advantage of a framework such as this is that it makes everything explicit (Klein & Myers, 1999) so that it can act as a trigger for debate. It may well be that proponents of particular methods may disagree with our validity criteria but at least there is now a target to be aimed at.

For a major research question it may be that all the validity criteria cannot be answered within a single study – there may need to be sequential studies, perhaps some formative, and then later ones summative; or different methods may need to be applied to different aspects of the situation, thus invoking different validity questions. These considerations clearly lead on to the possibility of mixed methods work. They also touch on the question of generalizability. In this paper we have distinguished the e-correctness (or validity) of a particular study from the extent to which it can be generalized to other contexts, but the two are clearly related, and the generalization question raises its own validity issues that we have not here addressed.

References


(Eds.), Information Systems Research: Relevant Theory and Informed Practice (pp. 121-142). Boston: Springer.


Reichardt, C. (2011). "Criticisms of and an alternative to the Shadish, Cook and Campbell validity typology". In H. Chen, S. Donaldson & M. Mark (Eds.), *Advancing Validity in Outcome...*
Truth and Correctness


