A CRITIQUE OF GENERALIZABILITY IN INTERPRETIVE RESEARCH

Mike Metcalfe & Maureen Lynch
University of South Australia

mike.metcalfe@unisa.edu.au

ABSTRACT

This paper is about generalizability in interpretive systems research. The authors are concerned that, in its search for appropriate inquiry methods, the discipline of Information Systems (IS) does not slip into the errors of other social disciplines and become a dismal science. Many systems thinkers have repeatedly argued that the purpose of IS research needs to be to produce what Argyris calls actionable knowledge and Ulrich, critique heuristics. That is, rules of thumb which managers can use to solve design problems, not too vague and not too detailed. This is a different concept from generalizability, which aligns with the scientific notion of seeking universal objective laws. The paper uses the argumentative inquiry in its reflexive capacity to critique knowledge claims on a chapter by Baskerville and Lee that discusses generalizability to bring out the differences between this cornerstone of universal truths and the more systems thinking concepts of actionable knowledge or critical heuristics.

THE PROBLEM

Baskerville and Lee (1999) wrote a chapter entitled “Distinguishing Among Different Types Of Generalizing In Information Systems Research” in which they refine the definition of “generalization” for the benefit of interpretive empirical inquiry. My concern is that they may have inadvertently led new researchers to think that generalizability is in some way central to interpretive systems research when this richer research tradition has a very different set of priorities. At least five different research perspectives provide examples of these priorities, notably in the now extensive argumentative (Rehg, 1999), systems inquiry (Churchman, 1971), perspectival thinking (Linstone, 1999; Haynes, 2000), actionable knowledge (Argyris, 1996) and critical social theory (Alvesson and Skoldberg, 2000) literature. Moreover, any discussion of generalizability might want to address Popper’s (1963) critique of induction.

The alternative priorities in interpretive systems research can be loosely clustered into one single claim that social inquiry is best seen as the justification of a perspective (insights, falsifiable conjectures). This justification may or may not use measured empirics, as empirics are understood to be actions to assist thinking, or merely illustrations intended to improve the human condition. Generalization is not central to this view of empirics. So, more specifically, this paper will explore the conjecture that interpretive inquiry offers alternative priorities than generalizability that are more appropriate for the “wicked” problem of Information Systems (IS) related inquiry.

Rather than merely list the attributes of these alternatives, they will be used as the lens with which to critique Baskerville and Lee’s chapter constructively. It is hoped this will serve two purposes. The first is to reveal how the work of these five perspectives might be used systematically to critique any social inquiry and so reveal underlying assumptions. Encouraging critique methods lies at the core of interpretive inquiry. The secondary purpose is to give further airing to the evaluation criteria these five perspectives suggest for interpretive inquiry, empirical or not.

In line with most inquiry methods, some upfront working definitions may be helpful to establish common ground. My understanding of generalization is that it is the concept of being able to extrapolate from a sample onto the population from which it is drawn, as a whole. Interpretive IS inquiry, whether empirical or not, is the search for justified interpretations (perspectives) of events or things, maybe distilled from the multiple interpretations of others in order to improve the appreciation of the researcher (system designer). By argumentation is meant the process of making a one line claim, conjecture, conclusion or proposition that has to be justified (well reasoned) to a knowledgeable cynical audience using convincing evidence. It includes anticipating the counter evidence. Objective knowledge is defined in terms of Popper’s (1971) cultural world 2.

The concept is further confounded in the IS literature by its conflation with distinct concepts like generality and generalizing. Untangling this confusion reveals a variety of obscured, and sometimes legitimate forms, of generalizing. Many intensive research ventures could satisfy one of these forms, yet the authors self-flagellate over their inability to found their generalities on the one particular form that IS researchers have mindlessly idealised.

4. Conclusion

The existing concept of generalizability is fastened on one particular form of generalizing. The concept is further confounded in the IS literature by its conflation with distinct concepts like generality and generalizing. Untangling this confusion reveals a variety of obscured, and sometimes legitimate forms, of generalizing. Many intensive research ventures could satisfy one of these forms, yet the authors self-flagellate over their inability to found their generalities on the one particular form that IS researchers have mindlessly idealised.

9 For style reasons, this paper is written in the first person.
The self-flagellation is not merely annoying; it is harmful. When researchers unnecessarily divest their right to claim generality, their research audience is denied their analysis of the utility of their theories. When researchers generalize and claim generality, they encompass the larger scope of phenomena, beyond those directly captured by their research, to which their findings and understandings apply... By renouncing their right to generalize and claim generality, intensive researchers lose the latitude to explain the wide field of uses for their findings...

ARGUMENTATIVE INQUIRY AND SYSTEMS THINKING

Popper’s work “Conjectures and Refutations” (1963) will need no introduction to scientific researchers. He solved the long induction versus deduction debate by calling for conjectures to be argued. To be “scientific”, these conjectures have to be risky or falsifiable rather than being inductions from empirics. Rehg (1999) usefully summarizes the now extensive literature (e.g. Perelman and Olbrechts-Tyteca, 1969; Walton, 1998; van Eemeren et al, 1987) on the use of argumentation as an interpretive inquiry method. He believes that argument, supported by convincing justification, provides a useful middle road between objectivism and relativism. It is the methodology of much of the critical social theory literature (Alvesson and Skoldberg, 2000) but one that asks that an argument has some surprise value or insight so as to emancipate.

As a critique method, argumentative inquiry encourages questions such as, “What was the chapter’s argument? Were terms well defined (common ground)? Was the supporting evidence convincing? Was the counter evidence dealt with well?” It also helps raise questions about the context of the argument such as, “Why is the author presenting this argument and what are the implications of accepting the argument?” Critical Social Theory prompts further questions such as, “What was their insight and how does that improve the future?” However, apart from relevance and possible counters, argumentative inquiry does not provide much advice on how to critique evidence beyond saying that it should be convincing. Here systems thinking (Ulrich, 1983) and Linstone’s (1999) multiple perspectives can be of some use. Systems thinking prompts questions like, “Where is the boundary, the inter-connectivity, the purpose, the transformation and the stakeholders” perspective?” The perspective approach, as developed by Churchman (1971), Checkland (2000) and Haynes (2000) can be used to complement argument in another way. This involves the explicit separation of the object from the perspective (lens, frame) of that object. So, a new system can be made the “object-under-study” and then alternative perspectives (efficiency, emancipation, interconnectivity) be used to bound the inquiry into that object. This perspectival thinking can be mapped onto an argument line that has an object and subject (Metcalfe and Lynch, 2002). For example, the argument, “generalizability is good” makes the object-under-study “generalization” and “goodness” the perspective. This separation then allows the two to be examined, defined, described separately prior to bringing them together to make up an argument that will need to be justified using convincing evidence.

For numerous other perspectives on how to critique an article see Metcalfe (2002). All these perspectives on human inquiry have been merged under argumentative inquiry into a pattern of six fairly broad questions (cues, themes) intended to assist in the critique of claims. These have been used to critique Baskerville and Lee’s chapter on generalizability. As the purpose of the questions is simply to prompt thought, the responses are intended to introduce new perspectives - not to be generalizable.

CRITIQUE PROMPTS

1) Do the authors explain why they are making their argument? Do they acknowledge any intent to persuade the audience of their conclusion?

I think they do. They believe there is a problem with some interpretive researchers feeling defensive that their findings are not generalizable because they were not based on the statistical assumptions of sampling. They do this by identifying different types of generalizability. For example, they define “generality” as generalizing to some subset of “the entire class of empirical referents”. That their advice is important is assumed as generalization is assumed to be a very important measure of the usefulness of interpretive research.

I think the authors did intend to persuade, but simply by presenting their own sub-definitions of generalizability. There are no empirics either of their own or of previous writers presented, nor are any alternative interpretations of the umbrella concept of generalizability considered.

2) What is the object or thing the chapter is studying? Did they discuss, define this object in terms of its boundaries, relationships, how it changes over time and its purpose?

Due to the well-structured and clear style of writing, there is no doubt that the object-under-study is generalizability. The authors want to re-define this more carefully into four different sub-types. Argumentation theory, starting from Aristotle, would commend this as definitions of words to be used in an argument are best agreed upon front so as to avoid semantic misunderstandings. Put another way, common ground needs to be
established prior to presenting evidence. In social inquiry arguments, a lot of time has to be spent on defining (describing, constructing) terms. A classic example is the: “All men are mortal, Socrates is a man, therefore Socrates is mortal” argument of Aristotle. As a social inquiry argument, this is very ambiguous, as “mortality” and “Socrates” need careful definition. Does mortality include in memory as well as the body, was Socrates a person or a personification, and are ideas mortal?

However, the authors present their sub-definitions of generalizability with little recourse to alternatives. It appears that they are not aiming to establish common ground prior to presenting evidence; rather it is their evidence. The authors do provide one phrase to define generalizability. They write “generalizability is a potential that a theory has, …” and draw a parallel with falsifiability as also being a potential that a theory has. There is no discussion of what a theory is or how essential it is to inquiry, methodologically speaking. They seemed to be simply telling the reader some unquestioned and unquestionable objective facts about generalizations.

Systems thinking encourages thought on interconnectivity of the object-under-study that makes me ask why “searching for objective facts” styles of inquiry place so much importance on generalizability. I think it is an outcome of that cornerstone of objectivity, repeatability. Results that can be repeated across time and space are generalizable. Clearly, repeatability is a problem with inquiry into human behaviour. This is because we, unlike physical objects like molecules, learn, making repeatable complex social situations difficult to achieve and study. It is because of this that social inquiry, such as IS research, is designed to alter the future behaviour of the objects under their study, people. This underlines the important differences between objectivist and interpretive (seeking interpretations) research and calls into question the use of an objectivist evaluation tool for interpretive inquiry. If you heat up molecules three times without explaining your purpose, they respond in exactly the same way. If you do the same to people, they are likely to abuse you.

Thinking about interconnectivity also leads to thoughts about the connection between interpretive research and statistical sampling. The idea behind sampling is that, if a sample has external validity, then averages that are true of the sample can be said to be true of (are generalizable to) the population. Averaging and sampling are two concepts that need to be used with care in interpretive inquiry. So, for example, with stakeholder analysis, which is so central to systems design, the averaging of the perspective of stakeholders would not be helpful. Sampling, and averaging, is therefore a concept not really compatible with seeking powerful stakeholders” interpretations. If one powerful stakeholder has a perspective, it is relevant, and perspectives can spread as stakeholders talk to each other. Interpretive research seeks perspectives, in particular unusual ones, not average perspectives.

3) What was the authors’ perspective (lens, intellectual frame…) on generalizability? Did they discuss this and appreciate the need to seek a range of perspectives in order to undertake interpretive research into generalizability?

The chapter authors seemed concerned that “intensive” IS research be generalizable because, if it were, that would make it scientific and therefore the results would produce good knowledge. They are concerned to produce objective knowledge. This, coupled with their use of words like “theory” and “the class of empirical referents”, forces me to assume they think research can be defined only as the search for objective facts, therefore their perspective is an empirical objectivist one. Their advice is all about how to add rigour to empirical measurement methods. They group what they call positivist, qualitative inquiry with interpretivist (qualitative or quantitative) inquiry. Put another way, it seems predictable that an objectivist would give well-meaning advice to interpretivists (as well as objectivists) that excludes attempts to explore alternative interpretations of the issue of the generalizability.

The perspective of those numerous epistemologists who present a more inclusive definition of knowledge are not presented (see Ulrich, 1983; Perelman, 1969; van Eemeren et al, 1987 and Rehg, 2002), to say nothing of the postmodernists. For example, the issue of generalizability has long been an issue in historical inquiry. What do they have to say on the issue? How generalizable are the causes of the outbreak of the First World War? Historians have long argued that to dismiss their inquiries as not producing good knowledge, because some interpretations of historical facts may not be generalizable, is a little simplistic. One of the seminal philosophers in IS, Churchman, as long ago as 1971 recommended and demonstrated that inquiry benefits from seeking these multiple perspectives (worldviews). His ideas have inspired two generations of IS inquirers. These include Mitroff, Mason, Argyris, Linstone and Checkland. These writers suggest that any inquiry into generalizability would seek to be inclusive of some alternative perspectives on knowledge. This work is not mentioned in the chapter.

My perspective, having been influenced by the works of the above writers, is that generalizability is not the most pressing performance measure that editors of interpretive inquiry should be looking for. More pressing is ensuring all stakeholders in an inquiry are given a voice; their justified perspectives need to be appreciated. Moreover, Argyris’ (1996) call to judge inquiry by its production of “actionable” knowledge is also very relevant. Rehg’s (1999) justified argument or Alvesson and Skodberg’s (2000) “justified insights” seem similar

---

10 For a good critique of falsifiability see Martin Gardner (2001), *Skeptical Inquirer* 25(4), pp13-14
calls. One of the consequential differences of undertaking inquiry into “objects” (subjects), often more intelligent and experienced than the inquirer, is that the unique opportunity exists to not only seek your interpretation of their behaviour but theirs also. To provide the researcher’s explanation and prediction confines inquiry into physical objects. However, that is about all you can do when studying inanimate objects (molecules), or people as inanimate objects. When studying people as intelligent and experienced equals, inquiry should seek their prediction of how they would like to do things in the future. The critical social theorists call this “social action”; Argyris calls it a “theory for action”. Using the systems design language for this, the inquiry should result in recommendations for the design of a better system. Again, this is an option barred to inquiry into physical objects. We cannot design nature.

4) What does the chapter argue/conclude?

Neither the introduction of the chapter nor the conclusion is presented as an argument. Rather, it states in the introduction that the purpose (a good systems term) is “to clarify the different processes of generalizing …” so, the chapter is perhaps a description of the objective concept of “generalizability”. However, using the argumentation perspective introduces some caution in accepting writings as neutral descriptions. I interpret the chapter as having an implicit argument that generalizability is an important performance measure for interpretive empirical research. Moreover, if interpretive researchers try hard, their inquiries could be acceptable to the objectivist. Indeed, this appears to be a “given” not worthy of any justification. Their suggestions do not appear to be open to debate - they are telling it like it really is, explaining it carefully to the “molecules”.

5) What evidence is brought to support the argument/conclusion? Was this evidence convincing? Was the counter argument fully considered? Were there any empirics? If so why, if not why not?

The chapter uses reasoning evidence to justify the authors’ argument that interpretive empirical results may be generalizable. No empirics are used to create this knowledge. Knowledge is often created this way; indeed the rationalists demanded it and Einstein provided an excellent example in his Relativity Theory. The reasoning presented in the chapter, as is often the case in social inquiry, comes in two parts: definitions of generalizability and definitions of standards of what constitutes good knowledge. First consider the provided definitions of generalizability. A quick check in the dictionary confirms that the authors’ definitions are not in common usage, and they explicitly say “we define” referring to their variants of the “general” concept. So, the only evidence provided is their un-argued sub-definitions of the possible different types of generalizability. Put another way, no attempt is made to anticipate the counter evidence or multiple perspectives. Supporting evidence for their four types of generalizability could have drawn on how it was used in the past, its roots in language or in its association with statistical testing. Do their sub-definitions of the types of generalizability align with statistical testing theory? If something is generalizable from some subset of the sample, then is it necessary to demonstrate this statistically? Also, “classification” is a large subject in library science, what theoretical justification is there for their particular sub-division of generalizability? What other types are possible?

The second arm to the reasoning in the chapter is that practitioners have lower standards of what constitutes knowledge than scientists. I figure that someone can be both a scientist and a practitioner and have these two standards. I am a bit nervous of any suggestion of elitism with knowledge; Popper’s language of knowing within a culture seems preferable. It seems just a power thing to assume practitioners have an impure version of knowledge that only scientists can purify. Scientists get their respect from having detailed knowledge, usually from precise measurement (Cohen, 1994). Practitioners can spend their every working minute among some complex social problems, which are not measurable or repeatable. They acquire the most intimate experience of the detail, knowing more than scientists will ever be able to measure from a site visit.

The definition of knowledge is something that has plagued philosophers for some time. Popper (1963) presents the conjecture that some idea like the law of gravity or evolution is simply the best argument yet, waiting for a Kuhnian revolution when a better argument is presented (Kuhn, 1970). Einstein’s relativity did this to Newton’s attraction argument. This conjecture might be about an objective fact or a Kuhnian perspective. Perspectival knowledge includes how to “see” organisations (Morgan, 1986), how we should treat other people, where best to invest and what defines scientific knowledge. The role of argument is to allow people who claim to know something (have an interpretation) to attempt to justify their position.

If empirical objectivist knowledge is the best and only true form of knowledge, then it seems reasonable to use its standards back on the generalizability chapter. The chapter does not apply the empirical objectivist standard on itself so should we assume it has produced unacceptable knowledge? If objective empirics are the only source of good knowledge, how could the authors have tested their knowledge that interpretive results can be generalizable? Rather than suggest some trivial and unending empirical program to test interpretive results for generality, an alternative might be to again turn the test on those who claim its usefulness. How generalizable are the results of social scientific research that has used the recommended statistical sampling? I suspect not at all. I struggle to think of any truly generalizable rule I have learnt from reading maybe thousands of so called social scientific articles. This includes research I have designed myself aimed at showing objectively and
empirically that small groups made up of self-selected members outperformed (statistically significant at 5%) those groups that had members allocated based on criteria like ability, cognitive style and inclusion needs. I could claim, as the tradition demands, that these results were generalizable, but my life experience stops me. Concerns over external validity were created by the very demands for empirical rigour of experiments. It became a “catch 22”. If I made the experiment realistic, then I lost the independence of variables. If I increased independence, I lost external validity. Either way, generalizability suffered. I do not believe that anything I could have done would have produced generalizable results, using the word as a practitioner would.

Moreover, by defining types of generalizability using only their own reasoning, Baskerville and Lee have silenced the perspective of interpretive researchers. This raises emancipatory concerns. Is power to be the judge of good knowledge; are there to be “insightful experts” whose opinions are the truth? Or rather, should there be silenced the perspective of interpretive researchers. This raises emancipatory concerns. Is power to be the judge of good knowledge; are there to be “insightful experts” whose opinions are the truth? Or rather, should there be a more democratic process for setting the criteria of good research? As an Australian, supportive of a culture that is very wary of tall poppies, the democratic option sounds preferable. The apparent impartiality of repeatable actions that provide the same outcome sounds attractive. But there are problems, as Popper (1963) points out. Who says the repeated results are to be classified together, the results still need to be interpreted, and what if repeatable actions are not possible? Popper’s solution is argumentation where everybody is given a voice to refute.

Moreover, by not treating the definition of generalizability as worthy of empirical evidence, the authors have missed the opportunity to learn from action – a key facet of systems thinking. Rather than thinking of empirics as an exercise in precise measurement, it can be thought of as an opportunity to justify an insight (Alvesson and Skodberg, 2000). Inquiry starts with an idea (perspective, intellectual frame) that can be learnt by actions that provide specific input into our senses. This may be done through experiments, measurements or by seeking the perspectives of those who have seen and heard things you cannot, by learning from them. In action research, this is extended to getting a group to experience a novel idea in actions and then reflect upon the idea. Again this perspective does not see generalization as the most important criterion for empirics.

COUNTERS

So what are the counters to generalizability? As mentioned in the introduction, there are many. Briefly, Argyris (1996) argues that what should be provided by social inquiry is actionable knowledge. If managers are really to alter their actions as a result of an inquiry, they need to be given a rule of thumb to help make decisions. System design seeks solutions to problems, solutions that can be enacted. Some solutions, like argumentation, are critique methods to find or seek or think through solutions to problems (Ulrich, 1983; Metcalfe, 2002). These need to be at the right level of conception. Too vague, like General Systems Theory, and the manager will struggle to apply it meaningfully to the many problems that present themselves on a daily basis. Too demanding, such as significance testing, and it cannot be applied quickly to all problem situations. The example Argyris provides refers to gatekeeper roles. His actionable knowledge is that gatekeeper roles bestow power on the holders, so think carefully about whether a gatekeeper role is involved in a problem’s solution and how it is to be managed.

Ulrich (1983, 2001) calls for the creation of critical heuristics (rule of thumb critique methods) to explore the boundary, a very systems term, of a proposal. He advocates systems thinking, seeking Churchman’s perspectives, and argumentation perspectives. Checkland (2000) seems to support the first two, to improve the appreciation of the systems designer. Argumentation (Perelman, 1969; van Eemeren, 1987; Walton, 1998; Rehg, 1999) sees the idea of providing a justified argument using convincing evidence as the equivalent to generalization. If the argument that systems X is best for situation Y is convincing, then the boundary (generalizability?) is explicit. Perspectival thinking (Haynes, 2000) seems to have a slightly different agenda. It feels the purpose is to identify and separate the object under inquiry from the perspective. This allows for appreciation that the object can be inquired about from different perspectives. For example, a new system (the object) can be inquired about from an economic, environmental and emancipatory perspective. Generalizability, if analogous, would be the search for new perspectives. Last, Critical Social Theory (Alvesson and Skoldberg, 2000) provides another counter to generalization being the dominant evaluation of inquiry. Rather, this perspective sees the extent to which the inquiry improves society as the most relevant evaluation.

6) Were the implications (the so-what) of accepting the authors’ argument explicit?

Accepting the authors’ explicit argument that interpretive empirics are generalizable to some extent rather lures systems interpretivists into being evaluated by objectivist criteria. This reinforces the impression that objective empirics borrowed from the physical sciences are the best and only way to produce useful knowledge. It also de-motivates discussion on alternative criteria to evaluate the search for knowledge to other than repeatable measurements. The interpretive literature has many suggestions about how to evaluate inquires - these need more voice. Examples such as convincing argument and actionable knowledge have been mentioned in the previous section.
CONCLUSION

I have argued against generalizabilty as a dominant measure of good interpretive system research, drawing on Popper’s solution to the induction versus deduction debate - that is to call for argumentative inquiry. This was elaborated using the system thinking problem solving concepts and of perspectival thinking as a critique lens on generalizability. Good knowledge is seen to be as much about appreciating justified perspectives as it is about seeking universal truths for human behaviour.

Rather than measuring research in terms of how well it can be generalized, it was suggested that preferable criteria might be how much it will improve the condition of mankind, or the problem appreciation by decision makers. Purported knowledge has to compete with other claims of knowledge, so a “Popperian” measure might simply be how well it out-argues opposition knowledge claims. A “Churchmanian” measure is whether the research provided people with more alternatives, and whether it altered people’s actions. Interpretive systems research needs to identify and have faith in its own measures of inquiry success and stop its rather sad pretence of “self flagellation” to appease journal gatekeepers.

REFERENCES

Metcalfe, M. (2002) “Multiple Methods For Critiquing Articles”, Working Paper University of South Australia {mike.metcalfe@unisa.edu.au}