



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Harley, B;Cornelissen, J

Title:

Reframing Rigor as Reasoning: Challenging Technocratic Conceptions of Rigor in Management Research Abstract Acknowledgments

Date:

2019-03-27

Citation:

Harley, B. & Cornelissen, J. (2019). Reframing Rigor as Reasoning: Challenging Technocratic Conceptions of Rigor in Management Research Abstract Acknowledgments. Zilber, TB (Ed.). Amis, JM (Ed.). Mair, J (Ed.). *The Production of Managerial Knowledge and Organizational Theory: New Approaches to Writing, Producing and Consuming Theory*, (1), Volume 59, pp.59-76. Emerald Publishing Limited.

Persistent Link:

<https://hdl.handle.net/11343/233331>

**Reframing Rigor as Reasoning: Challenging Technocratic Conceptions of Rigor in
Management Research**

**Bill Harley, The University of Melbourne & Joep Cornelissen, Erasmus University
Rotterdam**

We would like to thank the special issue editors and the reviewers for their very helpful comments. We are also grateful to Emma Bell and Tine Koehler, as well as participants in seminars at Cardiff, Lancaster, Bath, Glasgow and Durham and Erasmus Universities, for their comments on earlier versions of this paper. Any weaknesses in this paper are entirely our own responsibility.

Reframing Rigor as Reasoning: Challenging Technocratic Conceptions of Rigor in Management Research

ABSTRACT

In this essay, we critique dominant technocratic conceptions of rigor in management research and elaborate an alternative account of rigor that is rooted in methodology and involves a concern with the quality of scientific reasoning rather than a narrower focus on methods or measurement issues per se. Based on the proposed redefinition, we conceptualize how rigor, as an essential quality of reasoning, may be defined and we in turn qualify alternative methodological criteria for how we might assess the rigor of any particular piece of research. In short, with this essay our overall aim is to shift the basis of rigor to an altogether more legitimate and commensurable notion that squarely puts the focus on reasoning and scientific inference for quantitative and qualitative research alike. We highlight some of the benefits that such an alternative and unified view of rigor may potentially provide towards fostering the quality and progress of management research.

As scholars of management and organizations, our rationale for conducting empirical research is to produce theoretical knowledge. Empirical data and analysis are means to an end, therefore, in the sense that they form the basis for knowledge creation. This implies that the data and the analytical techniques which we employ must be of high quality, however quality is defined within the research paradigm which we employ. It also implies that the process of reasoning, through which we make truth claims and develop theory about the phenomena we study, must be based on solid grounds if the knowledge we produce is to be taken seriously. One way that we can think about the ways that we attempt to ensure the value of our data and analysis and the knowledge we produce from them is in terms of *rigor*.

We will develop the concept of rigor in more depth below, but for the moment Donaldson, Qiu and Luo's (2013: 154) definition as "the logical pursuit of an idea" captures the broad meaning of the term in most types of management research. This would seem an eminently sensible principle to apply to the collection and analysis of empirical data, as well as in relation to the processes leading to knowledge generation. Among mainstream management researchers, it appears, however, that rigor has been defined much more narrowly and specifically with for the most part an emphasis on technical precision in measurement and a broader concern with reliability and exactness in the resulting knowledge claims (e.g., Wright 2016).

In this essay, we argue that this widespread adoption of a specific notion of rigor is problematic as it operates from a limited conception of rigor as a technical concern at the level of empirical measurement and observations. We argue that this is problematic because it effectively reduces the philosophical notion of rigor from being a *methodological* concern around inferential reasoning forming the basis for defensible knowledge claims to a rather *technocratic* conception concerning measurement and measurement terms.

We aim to rebalance this unfortunate situation and set out to reframe and reclaim rigor as a methodological characteristic, or virtue, of inferential reasoning that thus applies

regardless of whether the research is in the positivist or constructivist tradition and employs either a qualitative or quantitative approach. In doing so, we challenge the root understandings of rigor in our field, elaborate an alternative framing and understanding, and offer provisional suggestions for how processes of reasoning in management research can be made more rigorous. In doing so, we offer an alternative basis for defining and judging rigor, which we believe is applicable to both quantitative and qualitative research and may, when it is subsequently applied, bring real benefits to the management field. Besides offering a commensurable framework for quantitative and qualitative researchers to work from - allowing them to converse more directly and on equal terms at the level of reasoning – the general account of rigor that we propose is helpful in improving the quality of our inferential processes as researchers, potentially aiding the quality of our research and of the knowledge we collectively produce (cf. Ketokivi Mantere, and Cornelissen, 2017).

The paper begins by reviewing the concept of rigor in management research, starting by classifying and characterizing dominant understandings of rigor. We then demonstrate that most of these understandings converge around a technical understanding of rigor as chiefly concerning questions of measurement and the proper application of methods. We problematize the specific and limited conception of rigor that it implies and offer an alternative, methodological framing of rigor that is rooted in works on reasoning in the philosophy of science (Lipton, 2004). We end the paper by elaborating how such an expanded conception might fruitfully be used in evaluating whether a piece of research and the knowledge which it generates is to a greater or lesser extent “rigorous”.

RIGOR IN MANAGEMENT RESEARCH

In recent years, there have been several debates about rigor in the field of management studies. In many of these debates, the notion of rigor is rhetorically positioned against a supposed alternative, such as “imagination” or “creativity” (Alvesson and Sandberg, 2013; Donaldson et al (2103) or “relevance” (Anderson, Herriott and Hodgkinson, 2001;

Kieser and Leiner, 2009), the implication being that rigor is hard to realize in combination with such other ends. What is furthermore noteworthy about the debates about rigor in management studies is that the term appears almost universally to be defined in terms of rigor in the application of methods, and at times even more specifically as confined to the application of particular kinds of methods.

The Oxford English Dictionary defines rigor as: “Extreme strictness or stringency (of standards, conduct, etc.); scrupulousness or rigidity in the application of rules, principles, or precepts...Strict sense or interpretation; precision, exactness” (<http://www.oed.com/>). It seems, however, that this broader, common sense understanding of rigor in reasoning and interpretation is not reflected in how rigor is oftentimes defined in our field as being primarily about technical issues concerning methods and measurement. In much of the writing on rigor, it is assumed that rigor must be of the kind applied in the physical sciences, with a particular emphasis on accurate measurement and the correct application of statistical techniques that supposedly form a robust basis for inferential reasoning (Alvesson & Sandberg 2013: 133-4; Anderson et al., 2001: 396; Harley 2015). We label this the technocratic view of rigor.

The dominance of this narrow conception is problematic. For quantitative researchers, it allows an unreflexive focus on issues of technique. Moreover, it facilitates the assumption that if a study is well executed in a technical sense, then it must be a rigorous study, shifting attention away from more fundamentally methodological questions. In a sense, the implicit assumption is that if a study meets criteria of precise measurement then *ipso facto* it also qualifies as a rigorously executed study in general, including how broader theoretical generalizations and patterns of causality were inferred. However, because this narrow definition of rigor is associated with positivism and apparently largely taken for granted, it is unsurprising that it appears to be rarely if ever challenged by quantitative social scientists. Instead, management researchers operating in this tradition implicitly assume that the precise

measurement of observations in survey or phenomena based research and within experimental tests equates to, and thus can indeed be taken as a sign of, rigor in general.

For qualitative researchers, it poses an even more fundamental challenge. Harley (2015) recently argued that

the conflation of ‘rigor’ with ‘statistical rigor’ has led to at least an implicit view that the mainstream approach is rigorous and thus credible, and other approaches are less rigorous and less credible. Thus, qualitative research appears, implicitly at least, to be a less acceptable approach than quantitative work (2015: 402).

To this extent, the slippage in the use of the term rigor has significantly narrowed our understanding of the very concept, and has promoted a very specific notion of rigor. A number of scholars have therefore challenged this narrow definition of rigor (see Alvesson & Sandberg, 2013; Anderson et al., 2001; Bell, Khothiyal and Willmott 2016; Harley 2015). Concurrently, many qualitative researchers also appear to have accepted the dominant definition, although how they have responded to it appears to vary.

There appear to be two common responses from qualitative researchers to the dominant definition. The first, which is perhaps more common among those who work within interpretivist or social constructionist traditions, is apparently to acknowledge it, but then to reject it as being meaningful for qualitative research. The second is to attempt to demonstrate that in fact qualitative research can be made consistent with the precepts of technocratic rigor (Merriam, 1995). As an illustration of the first approach, Bell and Thorpe argue that:

Rigour is a term used in the hard sciences like physics, where its application as a value is so consensually widespread that you would not need to say that someone’s work was rigorous, because this would be assumed to be the case. In a discipline like sociology, the precise opposite applies; use of terms like...rigour are likely to be seen as betraying the researcher’s limited understanding of methodological issues...In management research, the term rigour is frequently used, but the extent to which it is seen as an appropriate measure of quality, particularly when talking about qualitative research, is contested (Bell & Thorpe 2013: 25).

In other words, as this quote suggests, a specific notion of rigor appears to be accepted on faith and is then generalized to all kinds of management research. Qualitative researchers

may on those grounds reject it, but then in turn risk not getting their work published (Wright, 2016).

The second kind of response appears to be to accept the dominant definition of rigor as legitimate but, rather than rejecting it, to embrace it and to seek ways to make qualitative research more consistent with it. This would appear to be characteristic of an approach to qualitative research which leans towards a positivist epistemology and has largely taken the form of attempts by qualitative scholars to orient their work towards what Cornelissen (2017) calls “factor analytical” approaches. This involves using qualitative data to build causal linear models, analogous to the outputs of quantitative research, and which can be applied in, and generalized to, other contexts. This second approach is also problematic, in the sense that there is a risk that by pursuing this version of rigor qualitative scholars will limit the kinds of theoretical knowledge which they can produce. As Cornelissen puts it:

...when we reduce the breadth of methods and theorizing styles of the qualitative tradition to mimic the quantitative tradition, the gains in perceived methodological rigor may be offset by a negative impact on the richness and variety of theories in our field. (Cornelissen 2017: 13).

The two main responses from qualitative researchers to the dominant conceptualization of rigor both seem to have problems attendant upon them. To reject rigor per se, seems likely to lead to qualitative research being further delegitimized in mainstream management research. On the other hand, embracing the dominant version of rigor runs the risk of undermining precisely the characteristics of qualitative work which make it so valuable to the field (Cornelissen, 2017). At the same time, for quantitative researchers, simply accepting the dominant view militates against any attempt to ask fundamental questions about rigor which go beyond the technical concern with measurement.

The discussion above suggests that both acceptance and rejection of the dominant conceptualization of rigor are problematic, albeit for different reasons in each case. This implies that the problem lies not with rigor in a general sense – as a quality to be pursued in research (Amis & Silk, 2008), but with the specific way that rigor is commonly

conceptualized. We submit that empirical researchers need rigor. The alternative situation, where we do not think of quality standards such as rigor through which to judge empirical research are a looseness, sloppiness, carelessness and general laissez faire approach towards the research that we do and the knowledge claims that we make from such studies. Few if any scholars would support a lack of a concern with rigor, or indeed with quality considerations regarding research more generally (Amis & Silk, 2008), given the detrimental effect that a lack of such considerations has on the credibility and value of our knowledge claims. We certainly do not. We do, however, recognize that the dominant technocratic view of rigor is for the most part informed by the specific ontological, epistemological and methodological assumptions of positivism; yet, at the same time, this view is more broadly disseminated in the field as mentioned. In short, what we are thus not advocating is sloppy, careless, lax or loose approaches to research. Rather, in the remainder of this paper we argue for a conception of rigor which can accommodate a plurality of approaches to management research.

RIGOR IN PROCESSES OF REASONING

We argue that if we accept that rigor in research is about the logical pursuit of knowledge, and if we restrict our concerns to empirical research – research which analyzes empirical data to derive theoretical conclusions and thus to advance knowledge – then the pursuit of rigor must necessarily involve a focus on reasoning. This implies a focus on the logical chains which link steps in research method and theoretical claims which are made on the basis of data collection and analysis. This broader methodological concern with rigor in logic and reasoning is the prevalent view within the philosophy of science, intimately combining methods with issues of theory and inferential reasoning. However, within management research – as indeed in many other areas of the social sciences – a concern with technique often trumps a methodological concern with reasoning (cf. Hammersley, 2011).

Drawing on the work of Toulmin (2003), Ketokivi and Mantere (2010) argue that the process of reasoning comprises three elements: grounds, claims and warrants (see also

Ketokivi et al., 2017). By grounds they refer to the empirical component of research, while claims are the theoretical conclusions which are informed by the analysis of the data.

Warrants are the rules of reasoning which are applied to justify making theoretical inferences (claims) based on the data (grounds). When scholars present a claim based on analysis of data, the challenge they face is to convince others that their claim is warranted – that the process of making an inference from data is a logical and complete one. To put it another way, scholars seek to convince others that their reasoning process is *rigorous* (Ketokivi et al., 2017).

A crucial challenge emerges in empirical research, however, because the reasoning process can never be complete in a logical sense (Ketokivi & Mantere 2010). This is because all attempts to make theoretical claims based on empirical studies encounter the problem of induction, regardless of the specific method employed. This recognition sits somewhat uneasily with contemporary positivistic thought based on the Popperian principle that the inductive logic which characterized the work of the Vienna circle and their progenitors was fatally flawed in its reliance on verification. Popper's subsequent principles of deductive research based on falsification rested on the argument that we cannot infer universal statements from singular events and that this applies regardless of how numerous the singular events might be (Popper, 1934). Ketokivi and Mantere (2010), however, make a compelling argument that in fact all empirically-based theoretical conclusions are inductive in a logical sense, because researchers are moving from the specific and singular (their data) to the general (their theoretical claims), and that positivist research is therefore unable to overcome the problem of induction. All empirically-based theoretical claims are in fact based on an inductive logic according to this argument.

The reason that this is so important for questions of rigor is that inductive reasoning is logically incomplete. From any dataset there are multiple possible theoretical inferences that could be made. That is, there is no logical necessity that a particular piece of empirical analysis should lead to a particular theoretical conclusion; theoretical conclusions are

‘underdetermined’ by data (Lipton, 2004). The role of warrants – the rules of the game(s) – takes center stage in this respect as a way of justifying linking particular claims with particular grounds. This in turn means that what we can legitimately claim based on our data is always dependent on what set of rules we apply to reaching and evaluating theoretical inferences.

This realization introduces a further point which is important in thinking about rigor. If it was possible to identify one set of rules as inherently more rigorous than others, then knowledge generated using those rules could be regarded as more rigorously grounded than that based on other rules. The problem, however, is that there appears to be no universal basis for choosing between sets of rules or what Ketokivi and Mantere (2010) call ‘epistemic virtues’. By epistemic virtues, these authors refer to the rules which we apply in judging the value of the knowledge we produce. They argue that:

There is no consensus on universally applicable epistemic virtues. Instead, they must be understood as “standards around which a cognitive community can coalesce, standards that its members adopt as theirs, but not standards that hold universally” (Longino, 2002: 185). The choice of virtues is “a function of both the aims of research and inherited tradition”; they are indeed “features of a local epistemology” (Longino, 2002: 186, 187). (Ketoviki & Mantere, 2010: 310).

What this means is that ultimately any choice of rules – for example, privileging technical precision, measurement accuracy, novelty, ‘interestingness’ or any criterion over any other criterion – is ultimately arbitrary in a logical sense. If we accept this argument, then the logical conclusion is that we cannot regard any particular set of rules governing how we derive knowledge claims as superior to any other in terms of inherent rigor. This point can be illustrated by considering the process of reasoning which underpins the dominant technocratic version of rigor which we mapped out earlier in the paper. In the analysis that follows, we first deconstruct the roots of the dominant conception of rigor in our field, and highlight the limited basis that it provides for guiding management research. We also argue that the way in which this particular conception has been generalized as a yardstick for the entire field is based on a logical fallacy of hasty generalization. We then elaborate in the

subsequent and final sections of the essay how the alternative account that we provide may offer a roomier and legitimate basis for rigor in the context of the field of management research.

THE LIMITS OF RIGOR IN POSITIVIST SOCIAL SCIENCE

We have argued above that within management research there appears to be a widespread assumption that hypothetico-deductive quantitative social science in the positivist tradition offers an ideal-type template for thinking about as well as defining rigor. The emphasis here is chiefly on rigor in *method*, as a tradition that Hammersley (2011) characterizes as ‘methodology as technique.’ To put it another way, the presumed rigor that is typically associated with positivist social science would appear to reside in technique, but is then generalized beyond technique – a skilled use of technique becomes *ipso facto* the basis of rigor. If we shift our attention to rigor in processes of reasoning, however, it becomes clear that this approach faces challenges. Our aim here is not to denigrate positivist research, but simply to demonstrate that if we focus on processes of reasoning it is problematic to regard this tradition as exemplifying a kind of ‘gold standard’ for rigor in all kinds of management research.

Hypothetico-deductive positivist research has a single logic, which can be traced to its roots in the experimental methods of the physical sciences. The starting point is existing theory, typically ‘correlational theorizing’ (Delbridge and Fiss, 2013) framed in terms of causal associations between variables, which are used to develop falsifiable hypotheses about associations between the relevant variables. Variables are typically theoretical constructs, which must be operationalized by developing valid and reliable measures so that data can be collected. When data have been collected, they are analyzed using a variety of statistical techniques which allow for measuring associations and thus the testing of (the attempt to falsify) the hypotheses. If a hypothesis is not disproven, then it is regarded as (for the

moment) true and to that extent the theoretical proposition about cause and effect is regarded as correct.

Even on its own terms, however, the process of reasoning within this tradition faces significant challenges in terms of strict adherence to its own principles. This is not about the fact that some research is done that is ‘sloppy’, introduces biases or fails to apply statistical techniques properly, but rather more fundamental problems with the internal logic of positivism which afflicts even the most well-executed research study.

An apparent problem with the assumed rigor inherent in this kind of research is that, while a hypothetico-deductive logic requires that theory should precede and dictate method, in fact the methods we use dictate to a significant extent the kinds of questions we seek to answer and shape the theory which we generate (see Van Maanen, Sorensen and Mitchell, 2007). The assumption of those who adopt the technocratic view of rigor appears to be that methods are simply neutral technical tools which are employed to discover facts, which are then brought back to bear on theoretical claims. If, however, it is the case that methods dictate the kind of questions which can be asked and the kinds of theory which can be developed, it is apparent that positivist research does not typically – or even ever – conform to the logical reasoning process which is held to underpin it (Mantere & Ketokivi, 2013).

The form of reasoning – the linking of data and theory – that is supposed to characterize hypothetico-deductive research is what Mantere and Ketokivi (2013) call ‘computational reasoning’, involving the application of strict technical rules or algorithms to bridge empirical findings to theoretical inferences, and which assumes that any researcher who applied the rules properly would reach the same conclusion. In practice, however, this assumption does not seem to be tenable. Several scholars have argued that it remains very difficult to describe with precision the rules which researchers apply in reaching inferences based on results and that in fact, intuition, emotion and guesswork are more common than a

form of reasoning which involves the strict application of algorithm-like rules (see for example Lipton 2005, Swedberg 2016).

A final limitation of the hypothetico-deductive model is that it “enables us to co-opt our accounts of deduction for an account of induction” (Lipton, 2004: 15), focusing the efforts of researchers on the observable logical consequences of a given hypothesis. But in doing so, it “leaves to one side the important question of the source of hypotheses” (Lipton, 2004: 15), or the inductive grounds for a particular line of reasoning. Management researchers oftentimes simply state their hypotheses as based on a set of “true causes” (Snyder, 2006: 201); causes and causal models that are known to have produced similar effects in other instances and by the same token should be accepted as legitimate grounds for the study at hand. But in doing so, the problem of induction (why these causes in this instance? What is this instance a case of?) is not directly addressed, so that in many cases the resulting inferences and explanations do not rest on solid grounds. In fact, the inferences may logically turn out to be not rigorous at all, as “being too permissive, finding support where none exists” as well as being “too strict, since data may support a hypothesis which does not, along with reasonable auxiliary premises, entail them” (Lipton, 2004: 16).

The preceding discussion suggests three important limitations to the apparent rigor of conventional hypothetico-deductive positivist research, which go beyond ‘sloppiness’ in any given study and which reflect inherent limitations in this approach. First, as noted previously, it suffers from the same limitation of other traditions in the sense that it simply cannot overcome the problem of induction. Secondly, it appears that theory does not necessarily or fully precede data, which suggests that no matter how technically rigorous a particular piece of research is, the logical chain from theory to data to theory is brought into question. Third, the apparent neutral and objective application of rules to drawing theoretical inferences from data appears in practice to be impossible to achieve, thereby undermining any claims to rigor in reasoning. Finally, there are limits to the ‘tightness’ of linkages between hypotheses and data.

As noted above, the point here is not to denigrate this approach to management research. Rather we use the case example to demonstrate that on its own terms this tradition appears to face significant, and inherent, challenges in terms of living up to its own rules of rigor when we shift the focus from technical issues to processes of reasoning. To put it another way, we should be careful not to accept uncritically any assertion that the common technical understanding of rigor provides the yardstick by which the rigor of other research should be judged. In addition, the example demonstrates the importance of not conflating rigor in measurement with rigor in reasoning, and of generalizing a specific attribute of measurement to a general quality or characteristic of a study. This is a fallacy of generalization, and as a field we should take these notions apart again and reclaim rigor as a general methodological concern, rather than a technical one. Doing so, as we will argue, will not only provide a roomier notion of rigor, but also one that we suggest could provide a basis for judging the quality of reasoning and the robustness, or strength, of knowledge claims regardless of the paradigm within which a piece of empirical research is conducted or the kind of method that is being used.

BASES OF RIGOR IN SCIENTIFIC REASONING

If we cannot regard any particular mode of reasoning as inherently more rigorous than any other, the obvious question is how we might decide what constitutes rigor within and across different approaches to research. Are there particular rules which should be applied within or even across particular modes of reasoning? That is, are there general criteria for rigor that we might invoke? According to Mantere and Ketokivi (2013), this is a gap in the extant literature. They write:

Conspicuously missing from the extant literature is a methodological—as opposed to rhetorical, psychological, or social—account of scientific reasoning. The missing piece is crucial, because the general understanding of how scientists reason and formulate explanations is surprisingly limited...and yet prescriptive norms are essential in defining criteria for methodological rigor. (Mantere & Ketokivi 2013: 70)

If we accept this argument, as well as the earlier argument that there is currently no objective basis for regarding one set of rules as inherently more rigorous than any other, then the field remains wide open in terms of proposing bases on which to assess the rigor of reasoning in management research (Ketokivi et al., 2017). As discussed above, reasoning involves the process by which a researcher proceeds from data, as a set of grounds, to a set of theoretical claims in a plausible manner. Reasoning involves the conscious and deliberate use of relevant assumptions, explanatory principles, and observations to form and assess new inferences towards theoretical claims and to reject existing ones (Brandom, 1994). Such a reflective ordering and reordering of inferential linkages between grounds and claims by management researchers is what in ordinary language is typically assumed to be reasoning proper (e.g., Sonenshein, 2007) – the experience of forming an inference and arriving at a conclusion because one has reasons to accept it beyond intuitive force (Mercier and Sperber, 2011). In other words, reasoning involves the conscious and reflective use of logical argumentation to choose and articulate the reasons (i.e., assumptions and explanatory principles) for explicitly entertained hypotheses or explanatory accounts (Brandom, 1994), and in such a way that they stand as robust “warrants” to justify such inferences (Toulmin, Rieke and Janik, 1979).

As discussed above, such reasoning can be judged on the basis of different epistemic and/or pragmatic virtues for management research. As no universally accepted and absolute principles govern such reasoning (Ketokivi and Mantere, 2010; Mantere and Ketokivi, 2013), some researchers have therefore instead highlighted the importance of the pragmatic virtues of a deductive, inductive, or abductive argument in context, such as its simplicity (Weick, 1989), interestingness (Alvesson and Sandberg, 2011), or usefulness (Corley and Gioia, 2011). Here, we argue that the question of rigor as part of reasoning in management research can best be judged based on the two epistemic values of coherence and of inference to the best explanation, the general idea being that the more a given study embodies these values the more rigorous it is.

We must make two points clear before we outline our argument for these two values. First, we are proposing principles which we argue can be used to assess rigor in reasoning processes regardless of the specific method used in conducting research. That is, we are explicitly rejecting the idea that when it comes to reasoning any particular tradition or technique is more rigorous than another. We make no judgement about what sets of ontological, epistemological and methodological assumptions scholars make, nor what techniques they apply, but rather our focus is on how they go about reasoning within and as part of their chosen approach. Second, we do not pretend that we have overcome the difficulties, discussed by many scholars, of explaining in precise detail how reasoning is, or should be, conducted (see Lipton 2005; Swedberg 2016). Rather, we put forward two principles which can be used to inform reasoning and can therefore be used to assess methodological rigor by authors, reviewers and readers of management research alike.

Coherence

Theoretical claims based on a set of empirical observations are coherent when researchers can link the claims to the evidence in a systematic and clear way. Philosophers of science have argued that the internal coherence of assumptions and explanatory principles is a basic driver of reasoning and, more generally, a foundational prerequisite for science (Kuhn, 1962; Lakatos, 1976). We would argue that there are three crucially important aspects of coherence which contribute to rigorous processes of reasoning. These three categories perhaps overlap to some extent, but we will treat them separately here for the sake of clarity.

First, a prerequisite for coherence is clarity, detail and transparency in the way in which a chain of reasoning is set out and represented in a written research report or article. This may seem an obvious point. Indeed, some might argue that this is such an obvious criterion as to not require explicit mention. However, at the same time many published papers fare rather poorly against this base criterion.

As a second criterion, we propose coherence in terms of consistency between ontology, epistemology, methodology and method within individual studies. This might be thought of as methodological coherence. Perhaps this is an obvious point to make, but without coherence in terms of ontology, epistemology, methodological assumptions and method, it is difficult to achieve coherence in terms of the links between empirical data and the truth claims which underpin theoretical claims. Unless truth claims can be justified with reference to our methodological assumptions, processes of reasoning are likely to be undermined. We must make clear that we are not concerning ourselves with debates about the validity or otherwise of any particular set of ontological, epistemological and methodological assumptions, but rather with questions of internal coherence among the assumptions adopted by researchers.

The requirement for methodological coherence applies to both quantitative and qualitative studies. Regarding the former, if we assume that most quantitative studies in management are informed by a realist ontology – reality exists external to the observer and awaits discovery – and a positivist epistemology – reality can only be known by direct or indirect observation – then a logical corollary is the assumption that such studies allow claims to be made about an objective reality. Having said this, while it would be reasonable within this tradition to make claims about the presence or absence and the frequency of phenomena (or indicators of them) and what phenomena are associated with other phenomena (within the limits of chosen data collection techniques), it would be rather more controversial to make claims about unobservable processes or causal mechanisms which might explain such associations. We do not mean by this that causal claims cannot be made – of course this is what social scientists working in this tradition seek to do – but rather that such claims must be *inferred* from the empirical evidence, informed by theory, rather than being claimed to have been captured empirically.

In terms of methodological coherence, in qualitative research there is likely to be a wider range of ontological and epistemological assumptions, but the same principle applies

that there must be consistency. For example, it simply would not be credible for a researcher working in the social constructionist tradition to make the kinds of claims about objective reality which characterize positivist research. It might, however, be possible for scholars who adopt the assumptions of critical realism to make such claims. Coherence in ontological, epistemological and methodological assumptions is important for rigor in reasoning, but it cannot overcome the problem of induction in empirical research, which brings us to the third form of coherence which we propose.

Third, rigorous reasoning requires consistency in the logical processes through which scholars assemble arguments which lead them from a set of empirical observations to the theoretical inferences which they make based on such observations. Lakatos (1976) described how scientists, as part of larger research programs, accommodate incoherent observations and beliefs through their reasoning to make them coherent again by adding simple exceptions, additional qualifications, sub-categories, and more complex rules. Kuhn (1962) argued that scientists acquire through their education and ongoing research a stock of exemplars – concrete problem solutions in their area – and use them to guide their research (see also Lipton 2004, 2005). They pick new problems that are similar to an exemplar problem, use arguments and explanations that are similar to those that worked in that exemplar, and they assess their success by reference to the solutions and outcomes that the exemplars illustrate. The exemplars, in other words, set up a web of “perceived similarity relations” that coherently define a program of research (or paradigm), including coherent linkages between empirical phenomena and observations and assumptions and explanatory principles. In the management field, the importance of coherence is similarly acknowledged in writings (Locke and Golden-Biddle, 1997) that recognize that a lack of coherence in the form of a fragmented literature, puzzling observations or competing explanations is a key driver for research efforts where through synthesis and further elaboration researchers are then able to restore coherence by offering more encompassing or detailed arguments and explanations.

If we turn to the level of individual research studies, coherence manifests itself in a set of explicitly articulated and logical inferential links between observations and theoretical claims. Such claims are often presented as the theoretical “products” of the research in the form of a set of hypotheses or propositions, an integrative framework or process model, or a narrative theory. Shepherd and Sutcliffe (2011) argue that this coherence is established bi-directionally, whereby the sensory image of the data (observations) informs the set of claims, and the claims in turn inform the formation of the sensory representation. In an iterative manner, a management researcher works through the inferential linkage between the two so that there is less of a gap between them and coherence is established. “The end result of this process is a coherent story. The more coherent a story, the more it is accepted as a plausible explanation of the phenomenon” (Shepherd and Sutcliffe, 2011: 365).

Defined in this manner, coherence consists of an explicitly articulated set of inferential linkages – in the form of basic and auxiliary assumptions, and explanatory principles or rules – that offer a strong “fit” between observations and theoretical claims. It follows that when theoretical claims are simply put forward, but without an explicit account of the theoretical grounds and warrants that coherently link observations and claims that a solid basis for such claims is lacking. In other words, the reasoning that is then offered lacks rigor in the way in which the claims to knowledge are developed and put forward in the paper.

Thus far, we have argued that coherence – in terms of transparency, methodological assumptions and logical processes – is essential to rigor, regardless of the tradition in which the research is conducted and the specific methodology adopted. That is, processes of reasoning must be transparent and coherent if they are to provide a rigorous foundation for reaching theoretical claims based on empirical research. We now turn to a second criterion for rigor in reasoning – inference to the best explanation.

Inference to the Best Explanation

We have argued above that apart from genuinely deductive reasoning, which excludes all reasoning based on empirical research results, theoretical conclusions are underdetermined by data. That is, there is no logical necessity that a particular theoretical explanation follows from a set of results and therefore we have no choice but to work inductively. As empirical researchers who wish to generate sound theoretical explanations, how can we overcome this problem? Besides an explicit articulation of one's reasoning (as opposed to leaving this implicit), the rigor of a process of reasoning also rests on the degree to which the reasoning is seen to offer an inference to the best explanation (Lipton, 2004). Inference to the best explanation is also commonly referred to as abduction (Mantere & Ketokivi, 2013).

To the extent that a management researcher can show as part of her or his reasoning and data analysis that s/he has anticipated or accommodated alternative explanations the resulting theoretical claims are more robust, or rigorous. The contrasting scenario is one in which a researcher simply offers an explanation, instantiates it with reference to empirical observations or data, and in doing so narrowly confirms a particular line of reasoning. We would consider this process as lacking rigor, as it simply explains away possible alternative explanations but without proper justification.

A more rigorous form of reasoning accounts for this and infers one from amongst a set of competing explanations of the same observations. The notion of “best” thus does not suggest a singular or unique explanation that exists in and of itself, but it should be construed as operating over multiple potential explanations. The process involves considering plausible candidate explanations, and then trying to find inferential linkages that discriminate between them. In doing so, it also asks of management researchers that they assess candidate inferences by asking a subjunctive question; they should ask how good the explanation *would* be, if it were true (Tsang and Elsaesser, 2011). Operating the inferential “engine” in this way is generally seen to enhance the rigor of scientific reasoning.

In contrast, a straightforward inductive extrapolation without any subjunctive assessment is seen to fall prey to a fallacy of a hasty or too easy generalization that is too permissive of the data (Lipton, 2004). A comparative, or contrasting, subjunctive assessment is in other words needed so that researchers explore in a measured and considered manner the fit between theory and data (Sutcliffe and Shepherd, 2011) rather than simply assuming such a fit. Researchers may employ different senses when determining what makes for the best amongst a competing set of explanations, depending in part on the research tradition in which they are working. Lipton (2004) differentiates between the “likeliest”, or the most probable explanation, and the “loveliest” explanation; the one which would, if correct, be the most explanatory or provide the most understanding. Conventional positivist researchers may lean because of the hypothetico-deductive model to a focus on the “likeliest” explanation, whereas interpretivist researchers, in their quest for understanding, or *Verstehen*, may veer towards the “loveliest” option.

Lipton (2004) argues that whilst both senses, as criteria, may diverge when picking any explanation out in a competition, in most instances they converge. Because of this overlap, Lipton himself favors considering the inference to the best explanation as primarily based on determining the loveliest potential explanation, rather than the likeliest one. His reasoning in this respect is pragmatic, in that this argument serves to further discriminate the inference to the best explanation from classic hypothetico-deductive and inductive models of inference (whereas a focus on the likeliest explanation would draw it closer to the hypothetico-deductive model in particular). But his reasoning is also epistemic, as the likeliest explanation is typically an outcome or property of the loveliest explanation, so that an account of inference based on the loveliest potential explanation gives “the symptoms of likeliness, the features an argument has that led us to say that the premises make the conclusion likely” (Lipton, 2004: 60).

The implication for all researchers is that they engage in abductive and counter-factual reasoning in which they consider and evaluate alternative explanations before settling

on a “best” explanation. This approach importantly goes beyond issues of additional controls and finding the best fitting models in regression analysis. Instead, as we have argued before, it is not merely about incorporating such details – or “rigor” – as part of measurement and analysis, but about a broader view that considers such method issues in the context of broader chains of reasoning in which besides a deductively articulated logic other lines of argument have also been considered. Similarly, for qualitative research, the strengths of abductive reasoning go beyond a notion of rigor as tied to the qualitative protocol that was chosen to code and represent the data and to work towards a theoretical model (Cornelissen, 2017). Instead, it focuses attention on the broader process of contrastive and subjunctive reasoning that qualitative researchers engage in as they entertain various theoretical explanations and *within* which particular coding steps, as indicating the right or best inference, play a part.

In summary, we have elaborated two broad principles that offer a way of qualifying the notion of rigor in methodology, or more specifically scientific reasoning. There may well be other principles, but our attempt here is to highlight two crucial ones within ongoing processes of scientific inference. These principles are also not discrete rules or benchmarks as such; but should rather be considered as broad prescriptive guidelines that researchers can use to assess the rigor of a particular piece of research. These principles also apply to research studies within different traditions and using different methods, and as such offer a more equal footing upon which the qualities of research studies can be evaluated and indeed compared.

We also believe that the principles, whilst separate, contribute to one another and collectively contribute to a more “rigorous” piece of academic research. The principle of coherence relates to the internal consistency of an argument and the way in which it is transparently explicated and drawn out in a study – clarifying in detail the assumptive grounds and warrants upon which a claim, or conclusion, to knowledge rests. The inference to the best explanation principle is in comparison more external in focus in that it focuses on various inferences that work over alternative sets of grounds and warrants, as a basis for enhancing reflexivity (as a process outcome), as well as more instrumentally, as a means ultimately towards inferring the best explanation, *ceteris paribus*. In a sense, the inference to

the best explanation may be seen to assume the former, as incoherent and non-conclusive inferences from the data may not be selected as plausible candidates for explanation.

Similarly, a coherent argument may also presuppose the inference to the best explanation principle; as the latter may help lead to more rounded and well considered – in short, coherent and elegant – explanations.

CONCLUDING COMMENTS

Our starting point for this essay was the argument that the primary reason we conduct empirical research is to produce theoretical knowledge, making rigor a central concern in the conduct of research and the processes through which we draw theoretical inferences. We have critiqued the institutionalization of a narrow and ultimately pernicious belief amongst management researchers that rigor as an important quality of empirical research rests on exact measurement and incorporating specific technical procedures at the level of methods. This belief has taken the status of a taken-for-granted myth in that it is accepted at face value and reaffirmed through specific research practices in our field including methods training, the reviewing of papers and how more generally management researchers tend to judge the quality and contribution of a particular study. However, it is a belief that narrows the much broader notion of rigor that has historically existed within the philosophy of science, leading to a reduction of methodology to technique.

To address this issue, we have drawn out the basis of this root understanding, and the way in which it has infused the thinking of both quantitative and qualitative researchers alike. We offer a reframing of rigor as more rightfully placed at the level of inferential reasoning, which includes methods, but importantly is not confined to them, and places them as part of broader chains of more or less robust – or rigorous – forms of reasoning. This reframing shifts the basis for judging rigor, and we have articulated two principles by which rigor in this broader sense may be judged. In offering these principles and in shifting the basis of rigor away from method, as technique, to reasoning, as methodology, our ultimate aim with this

essay is to take a small step towards broadening the debate on rigor in our field in a way that applies to both qualitative and quantitative research.

In addition, with these principles we aim through our own reasoning to have offered a universal set of principled aspirations at the level of reasoning, rather than obsessing about an illogical extension of technical norms of “rigor” from quantitative to qualitative research. These principles may have to be further defined, and other principles may equally apply, or as a matter of fact turn out to be even better in judging the rigor of reasoning. Nonetheless, the underlying premise of having these general principles is that it makes quantitative and qualitative research to some extent “commensurable” when based on a common set of criteria. Such a situation may be beneficial to the field as a whole, in that it may foster more interaction, discussion and even integration between scholars working in different research traditions (Agarwal and Hoetker, 2007; Ketokivi et al., 2017). In addition, these principles ought to aid quantitative and qualitative researchers each in their own ways as well. For quantitative researchers, for example, these principles may help them become more aware of how ultimately the quality or strength of their study is to be judged at the level of their reasoning and how they support their causal inferences. Such awareness may help such researchers work around the inherent limits of the hypothetico-deductive model and to put their inferences on a stronger footing.

Two further points need to be made clear. First, as we have alluded to above, precisely how the criteria of coherence and inference to the best explanation might be applied in conducting or evaluating a piece of research depends crucially on the ontological and epistemological assumptions underpinning that research. In this sense, we recognize and echo Amis and Silk’s (2008) argument that what counts as legitimate research depends on our underlying assumptions. Our recommendations are therefore intended to provide a broad framework for thinking about rigor within different research traditions, rather than providing a single ‘recipe’. Indeed, the motivation for this paper was the apparent dominance of a single narrow conception of rigor and our perceived need to go beyond this and develop principles of rigor which apply across a plurality of approaches. While we have provided some

preliminary discussion above, more detailed work on the application of the criteria, and indeed of additional or alternative criteria, remains to be done. This would require detailed consideration of how our principles could be applied within a range of traditions and methodological frameworks.

Second, in putting forward our argument, we must recognize that we ourselves struggle to be rigorous in our own empirical research and we are quite confident that our work could be criticized when judged against the criteria which we have set out here. Our intention is not to suggest that we hold the secret of rigorous research which others do not, but rather to stimulate debate and constructive engagement. The ultimate aim of this essay is to encourage management scholars to engage critically and reflexively with questions of rigor, to refocus on rigor as a methodological issue and, hopefully then, to conduct more rigorous research.

REFERENCES

- Agarwal, R. and Hoetker, G. (2007), A Faustian bargain? The growth of management and its relationship with related disciplines, *Academy of Management Journal*, 50(6), 1304-1322.
- Alvesson, M., Hardy, C. & Harley, B. (2008) Reflecting on reflexivity: Reappraising practice, *Journal of Management Studies*. 45(3), 480-501.
- Alvesson, M. & Sandberg, J. (2013) Has management studies lost its way?: Ideas for more imaginative and innovative research, *Journal of Management Studies*, 50(1), 128-152.
- Amis, J. and M. Silk (2008) The philosophy and politics of quality in qualitative organizational research, *Organizational Research Methods*, 11(3), 456-80.
- Anderson, N., Herriott, P. & Hodgkinson, G. (2001) The practitioner-researcher divide in Industrial, Work and Organisational (IWO) psychology: Where were we now, and where do we go from here?, *Journal of Occupational and Organisational Psychology*, 74(4), 391-411.
- Bell, E., Khothiyal, N. & Willmott, H. (2016) 'Methodology-as-technique and the meaning of rigour in globalised management research', *British Journal of Management*, 28(3), 534-50.
- Bell, R. & Thorpe, R. (2013) *A very short, fairly interesting and reasonably cheap book about management research*, London, Sage.
- Bluhm, D. J., Harman, W., Lee, T. W. & Mitchell, T. R. (2011), Qualitative research in management: a decade of progress, *Journal of Management Studies*, 48(8), 1866–1891.
- Brandom, R. (1994), *Making it explicit: reasoning, representing, and discursive commitment*, Cambridge: Harvard University Press.
- Corley, K.G. and Gioia, D.A. (2011), Building theory about theory building: What constitutes a theoretical contribution? *Academy of Management Review*, 36(1), 12-32.

Cornelissen, J. (2017) Preserving theoretical divergence in management research: Why the explanatory potential of qualitative research should be harnessed rather than suppressed, *Journal of Management Studies*, 54(3), 368-383.

Delbridge, R. & Fiss, P. (2013) Editors' comments: Styles of theorizing and the social organization of knowledge, *Academy of Management Review*, 38(3), 325-331.

Donaldson, L., Qiu, J. & Luo, B. (2013) For rigor in organizational management theory research, *Journal of Management Studies*, 50(1), 153-172.

Hammersley, M. (2011) *Methodology: who needs it?*, London, Sage.

Harley, B. (2015) The one best way? 'Scientific' research on HRM and the threat to critical scholarship, *Human Resource Management Journal*, 25(4), 399-407.

Ketokivi, M., Mantere, S. and Cornelissen, J. (2017), Reasoning by analogy and the progress of theory, *Academy of Management Review*, 42(4), 637-658.

Ketokivi, M. & Mantere, S. (2010) Two strategies for inductive reasoning in organizational research, *Academy of Management Review*, 33(2), 315-333.

Kieser, M. and Leiner, L. (2009) Why the rigor-relevance gap in management research is unbridgeable, *Journal of Management Studies*, 46(3), 516-533.

Kuhn, T. (1962), *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Lakatos, I. (1976), *Proofs and refutations*, Cambridge: Cambridge University Press.

Lipton, P. (2004) *Inference to the Best Explanation*, 2nd Edn, London: Routledge.

Lipton, P. (2005) 'The Medawar Lecture 2004: The Truth about Science', *Philosophical Transactions of the Royal Society*, 360(1458), 1259-1269.

Locke, K & Golden-Biddle, K. (1997), Constructing opportunities for contribution: Structuring intertextual coherence and “problematizing” in organizational studies, *The Academy of Management Journal*, 40(5), 1023-1062.

Longino, H. (2002), *The fate of knowledge*, Princeton: Princeton University Press

Mercier, H. and Sperber, D. (2011), Why do humans reason? Arguments for an argumentative theory, *Behavioral Brain Sciences*, 34(2), 57-74

Merriam, S.B. (1995), What can you tell from an N of 1? Issues of validity and reliability in qualitative research, *PAACE Journal of Lifelong Learning*, 4, 51-60

Mantere, S. & Ketokivi, M. (2013) Reasoning in organization science, *Academy of Management Review*, 38(1), 70-89.

Phillips, N. & Hardy, C. (2002) *Discourse Analysis: Investigating Processes of Social Construction*, Thousand Oaks: Sage.

Popper, K. (1934) ‘The problem of induction’, in *The logic of scientific discovery*, Abingdon: Routledge Classics.

Shepherd, D.A. and Sutcliffe, K.M. (2011), Inductive top-down theorizing: A source of new theories in organization, *Academy of Management Review*, 36(2), 361-380.

Snyder, L. (2006) *Reforming philosophy: A Victorian debate on science and society*, Chicago: University of Chicago Press.

Sonenshein, S. (2007), The role of construction, intuition, and justification in responding to ethical issues at work: The sensemaking-intuition model, *Academy of Management Review*, 32(4), 1022-1040

Swedberg, R. (2016), Before theory comes theorizing or how to make social science more interesting, *The British Journal of Sociology*, 67(1), 5-22.

Toulmin, S. (2003) *The uses of argument*, Cambridge: Cambridge University Press.

Toulmin, S. Rieke., R.D, & Janik, A. (1979) *An introduction to reasoning*,
Basingstoke: Macmillan Publishing.

Tsang, E.W.K & Ellsaesser, F. (2011). How contrastive explanation facilitates theory building. *Academy of Management Review*, 36(2), 404–419.

Van Maanen, J., Sørensen, J.B., and Mitchell, T.R. (2007), The interplay between theory and method, *Academy of Management Review*, 32(4), 1145–54.

Weick, K.E (1989), Theory construction as disciplined imagination, *The Academy of Management Review*, 14(4), 516-531.

Wright, P. (2016) Making great theories, *Journal of Management Studies*, 54(3), 384-390.