RIGOR WITH OR WITHOUT TEMPLATES? THE PURSUIT OF METHODOLOGICAL RIGOR IN QUALITATIVE RESEARCH

Bill Harley
Department of Management and Marketing
The University of Melbourne
bharley@unimelb.edu.au

Joep Cornelissen
Rotterdam School of Management
Erasmus University of Rotterdam
cornelissen@rsm.nl

Biographical Note:

Bill Harley is Professor of Management in the Faculty of Business and Economics at the University of Melbourne, Australia. He is best known for his work on employee experiences of managerial practices with a focus on power and control in organisations.

Joep Cornelissen is Professor of Corporate Communication and Management at Rotterdam School of Management, Erasmus University, The Netherlands. The main focus of his research involves studies of the role of corporate and managerial communication in the context of innovation, entrepreneurship and change, and of social evaluations of the legitimacy and reputation of start-up and established firms.

ABSTRACT

The domain of qualitative research is replete with templates; standard protocols for the analysis of qualitative data. The use of such templates has sometimes been considered as automatically enhancing the rigor of qualitative research. In this paper, we challenge the view that in the context of qualitative research rigor is tied into the application of established protocols. Instead, we argue that rigor emanates from the way in which researchers engage in a deliberate reasoning process of inferring theoretical claims from their data. Such reasoning exists outside of templates, although it may make use of templates. Framing rigor as an emergent quality of reasoning, we distil the main processes through which qualitative researchers derive inferences from data and provide criteria for reflecting on the rigor with which they do so. We then extend these criteria into a set of practical recommendations through the presentation of examples and the framing of questions to focus researchers' thinking on the application of the criteria. In doing so, the paper aims to help qualitative researchers to conduct research which leads to rigorously derived theoretical insights, without having to resort to templates to attempt to do so.

KEYWORDS

Templates, qualitative, rigor, inference, reasoning processes, methodology.

INTRODUCTION

A perennial challenge for scholars who collect and analyse empirical data as a means to generate knowledge is how to demonstrate that the theoretical conclusions they reach are sound. This concern with the soundness of theoretical conclusions derived from data is at the heart of many of the most significant contributions to philosophy of science and methodology (see Lipton 2004). In the field of management and organizational studies, researchers have similarly debated what criteria should be applied within different methodological approaches, to justify theoretical inferences from data. In many instances, such criteria have coalesced into 'templates'; standard protocols by which researchers justify inferences from data and demonstrate 'rigor'. Such templates have become central to both quantitative and qualitative research approaches – with the latter being the key focus in this paper (Gioia, Corley & Hamilton 2013, Gjerde and Alvesson 2019).

Templates provide a means to derive theoretical claims from qualitative data that are – based on codified standards – warranted and grounded in defensible epistemic virtues. For example, Gioia et al. (2013) offer a protocol that describes a series of coding steps that allow a researcher to derive grounded theory from qualitative data. Their protocol, when used, implies a choice in favour of the classic epistemic virtues of parsimony and transferability (in the form of developing transferable constructs or a process model) (Amis & Silk, 2008), rather than providing a contextually detailed, ethnographic interpretation of a setting (Cornelissen, 2017).

As protocols such as this one have become established in the management field, many qualitative researchers have started simply to apply the protocol in question as a template for their work; forcing their data through the protocol, perhaps without being aware of, let alone reasoning through, the methodological assumptions associated with the template. In short,

they have come to substitute their own reasoning with a procedural application of a template as a proxy for rigor in their analysis (Cornelissen, 2017).

Taking this observation as a starting point, in this paper we challenge understandings of rigor that consider it as integral to templates, or even confine rigor to the proper application of a template. We offer instead an alternative perspective; one that conceptualizes rigour as an *outcome* of processes of inferential reasoning. Thus, rigor as we develop it here does not reside in protocolized techniques or methods, but in the very process of researchers reasoning about their use of particular techniques or methods and reporting such reflections in their papers.

To wit, when qualitative researchers engage in reasoning, and consciously deliberate the qualities and rigor of their own analysis, they may incorporate aspects of templates, such as specific methodological practices or coding techniques, or even combine such practices or techniques. But rather than using templates as a substitute or quick fix for rigor, when they deliberate and reflect on their choices, they do not constrict their reasoning to a template, or abandon conscious reasoning altogether. Instead, they make considered and defensible choices in moving from data to their theoretical claims.

In this paper, we challenge existing notions of rigor as strictly tied into templates and present and develop an alternative reasoning-based conceptualisation. We elaborate how rigor can be enhanced in processes of reasoning in qualitative management research. Our central contribution, building on prior work (e.g., Mantere & Ketokivi 2013; Ketokivi et al., 2017), is to offer an understanding of rigor as something that is established in and through a researcher's reasoning processes, as opposed to it being an intrinsic quality of a certain protocolized template. We work towards this contribution by offering a framework featuring a set of criteria that we offer as 'benchmarks' of rigorous reasoning. Having presented our

framework and criteria, we then offer practical guidance to qualitative scholars on how these criteria might be applied in making theoretical inferences from empirical data and in showing rigor in their own work. With these contributions, we aim to provide scholars with a way to strengthen the rigor of their own research, and in ways that are not limited to the rote application of templates.

This paper is structured as follows. In the first part of the paper we briefly consider discussions around templates and rigor in the context of qualitative research. We critique the often-made assumption that rigor inheres in certain templates for qualitative analysis and argue that templates can potentially undermine rigor when applied in a 'proceduralist' (Hammersley 2011) fashion. The second part of the paper proposes an alternative perspective; one which, we argue, has the potential to allow qualitative researchers to demonstrate rigor over and beyond the use of templates. This alternative turn involves a reframing of rigor, focused on processes of reasoning, and we map out a set of criteria for rigorous inferential reasoning from data. In the third part of the paper, we provide practical guidance to qualitative scholars about how processes of reasoning might be made more rigorous. In doing so, we deliberately avoid proposing rules or protocols ourselves and instead present and analyse a series of examples which illustrate each of our criteria for rigor. We then set out a series of questions aimed at stimulating reflexivity on the part of qualitative researchers. The final part discusses the contributions made by the paper and its implications for qualitative research.

TEMPLATES IN QUALITATIVE MANAGEMENT RESEARCH

It is increasingly common for qualitative scholars to analyse and present data in the form of 'templates' (Köhler, Smith and Bhakoo 2019). By templates we do not specifically mean 'template analysis' (King 2004) as a particular step in analysing textual data (although this

would be an example of a template approach). Instead we are here focused more broadly on templates as protocolized ways of conducting and analysing qualitative research, exemplified by approaches such as the 'Gioia Methodology' (Gioa et al 2013) and Eisenhardt's comparative case study method (1989). Templates provide qualitative scholars with general protocols for the analysis and presentation of their data. Within the management field, templates are often also closely linked to notions of rigor.

For example, in a recent chapter reviewing rigor in qualitative research, Reinhardt, Kreiner, Gioia & Corley (2018) survey the work of qualitatively oriented colleagues. Based on interviews with these colleagues the authors argue that rigor is oftentimes closely associated with the use of a particular template or protocol, but that there is much more to rigor than this. One of their interviewees (Davide Ravasi) argues that:

> qualitative research is not rigorous just because it follows a structured, rigid protocol for data collection and analysis; it is rigorous if it is transparent about the unexpected, surprising observations that led you to reorient your focus, the twists and turns your project took as your observations challenged initial research interests and working assumptions (Reinhardt et al., 2018: 519).

Based on this and other comments, Reinhardt et al. (2018) criticize the frequent and implicit bias of equating rigor with the use of a standard protocol and point instead to rigor as something that emerges from the transparent display of a researcher's reasoning from data to theory. In their recommendations, however, Reinhardt et al. (2018) point in particular to presentational issues that, they suggest, support rigor, such as a rich description of the research context (fostering credibility), reporting and revealing the voices of the subjects studied (enhancing verisimilitude), and overall transparency on the analytical steps that were followed (allowing for auditability). Here we complement their view by unpacking characteristics of the actual inferential steps that are involved in moving from data to theory, and how these can be seen to be deserving of the 'rigor' label, or not.

More generally, rigor is a defining feature of the scientific method and the demonstration of rigor in research is an important element in establishing the legitimacy of a piece of empirical work. Within quantitative research, rigor commonly involves an emphasis on standardised techniques and protocols for the analysis and presentation of data (Harley & Cornelissen 2019, Harley 2015). Protocolized templates in qualitative research are similar in that they also specify techniques and protocols for data analysis and presentation and assume that rigor is tied into the precision and faithfulness with which the protocol is applied to the analysis of data and the reporting of results.

We argue, however, that seeing the use of templates as a way of demonstrating or enhancing rigor may in fact have the effect of *undermining* the rigor of qualitative research. At the heart of this issue is precisely the fact that templates provide standardised protocols for analysing and presenting qualitative data. Hammersley characterises such stock approaches as 'methodology-as-technique' or 'proceduralism'; "the idea that good practice amounts to following a set of rules that can be made explicit as a set of prescriptive dos and don'ts, or even in the form of recipes" (2011: 21). Templates harbour the risk of researchers following a proceduralist approach – following a set of prescribed steps to be taken in data analysis and presentation – when they are applied unreflexively in the belief that they will automatically enhance rigor.

The main problem which arises from proceduralism is the assumption that following a set of rules is *ipso facto* the same as demonstrating rigor. As we will argue in the next section of the paper, however, rigor must be understood as encompassing a much wider range of considerations than a set of technical steps in analysing and presenting data. With such a strict definition, the use of templates can as much mask as reveal the rigor of a study, while at the same time diverting attention from other considerations.

Tying rigor to the application of templates raises a further difficulty, which is that it ignores the fact that the 'quality' of qualitative research varies across traditions and methods (Amis & Silk, 2008). Qualitative research includes a variety of ontologically and epistemologically distinct approaches, each with their own set of quality-related considerations and criteria. Different sets of ontological, epistemological and methodological assumptions prescribe different approaches to analysing and presenting data and a 'one size fits all' approach would therefore be problematic (Pratt, Kaplan and Whittington 2019; Reinhardt et al., 2018). Indeed, casting rigor as a quality of templates that are based on a 'foundationalist' epistemology (Amis & Silk, 2008) (the Eisenhardt template for comparative case analysis is an example) is problematic in that it disregards that rigor, or other conceptions of quality, are different in other non-foundationalist traditions. Further, the specification of approaches to analysis and presentation of data can act as a 'strait jacket', which constrains flexibility and creativity in theorising (Cornelissen 2017).

We must make clear that our argument is not with templates as such. Templates oftentimes codify best practices and conventions for a particular qualitative method. For many forms of qualitative work, informed by specific ontological underpinnings, they may be entirely appropriate. King (2004) for example argues that a form of research where templates are appropriate is:

> …realist qualitative work which accepts much of the conventional positivistic position of mainstream quantitative social science. That is to say, research which is concerned with 'discovering' underlying causes of human action, and which seeks to achieve researcher objectivity and to demonstrate coding reliability (King 2004: 279).

If applied in this kind of research, the use of templates may be a sensible way to conduct and present the analysis in a rigorous manner.

What we argue against, however, is the rote application of templates and the assumption that rigor is an intrinsic quality of the proper application of a template. This is a fallacy, but also has, as mentioned, negative practical consequences. In fact, thinking in this way may undermine the very ideal of rigor, as a quality that arises – and has to be demonstrated – from a reasoned process through inferential links between data and theoretical claims. Earlier we foreshadowed the argument that rigor can be conceptualised in other ways and that there is a much broader range of considerations than simply how data are analysed and presented following a particular protocol or template. In the next part of the paper, we argue that an alternative way of thinking about rigor is to focus more generally on the processes of reasoning that are involved in making inferences from data.

RIGOR AS A QUALITY OF REASONING

In developing this alternative path, we focus as already mentioned on rigor in reasoning; the processes by which theoretical inferences are made based on empirical data. The starting point for our argument is that rigor is not a term with a single fixed meaning, but rather a term which means different things depending on the perspective from which it is assessed (Irwin 2019: 197). In much management research, rigor is conceptualised as involving technical issues around measurement and data analysis, stemming from quantitative analysis (Harley 2015). This is, however, by no means the only way to think about rigor in research.

In empirical research, we generally seek to make theoretical claims based on various kinds of data or evidence, which of course implies that technique matters, but it also suggests that rigor must be focused more fundamentally on the way in which theoretical claims are derived from data. That is, if we are concerned with empirical research, rigor necessarily involves a concern with inferential processes of reasoning.

This recognition means in turn that we need to focus on the logic that links steps in the use of a specific method to specific theoretical claims. This broader concern with rigor in reasoning is in fact the prevalent and original view of rigor within the philosophy of science, combining methods with issues of theory and inferential reasoning (see Lipton, 2004). An obvious illustration of this is Popper's (1959) work on falsification. Popper's argument about what made for scientific knowledge was not concerned with measurement or analytical technique, but with reasoning processes. We similarly argue that if we wish to get to the nub of the question of how we judge conclusions reached on the basis of empirical research, then a focus on reasoning is paramount.

Before we begin to set out our criteria for rigor in reasoning, it is necessary to explain how we conceptualise reasoning. First, we take it as given that any reasoning which involves making theoretical claims based on empirical data is necessarily *inductive*. Only claims which follow necessarily from other claims can be said to be the products of deductive reasoning. Claims based on empirical data are, however, by definition inductive as well as 'underdetermined', in the sense that several claims might be made based on a given empirical finding, and thus the claims one makes involve an inductive conceptual leap (see Lipton 2004: 5-7). How one reasons to underpin, or warrant, this leap is, we argue, key to a demonstration of rigor in a particular study. Following Toulmin (2003), we understand reasoning as comprising three elements: grounds, claims and warrants (see also Mantere & Ketokivi 2013). Grounds are the empirical data, claims are the theoretical conclusions inferred from the data and warrants are the rules of reasoning which are applied to the data to make inferences. When researchers make a claim based on an analysis of the data, they must persuade themselves and others that their claim is warranted. That is, they must demonstrate that the specific inference that they are putting forward is a logical and fairly complete one (over and beyond alternatively inferred interpretations or explanations). In other words,

researchers must, in doing so, persuade others of the *rigor* inherent in their reasoning process (Ketokivi et al., 2017).

Such a process of reasoning is, we argue, a deliberate one; where researchers draw on techniques of data analysis but then consciously draw out their own inferences and display in their papers their reasoned choices and the deliberations that they have made. In comparison, the difficulty with a procedural use of a template is that it oftentimes pushes such deliberate reasoning to the background or nullifies it altogether; by priming researchers to follow the template rather than the other way around – accommodating the template, or elements of it, in one's own considered choices. The result of such 'proceduralism' is that a more encompassing consideration of rigour, as stemming from the nature and outcomes of a researcher's reasoning process (rather than seeing it as a fixed quality of the rule-based application of a protocol or template as a whole) is not being considered.

Explicating Rigor in Reasoning

The key question that this reasoning perspective triggers is, on what basis we might assess whether a reasoning process is rigorous? That is, on what basis might we judge whether the inferences which arise from a process of reasoning should be regarded as justified? To assess the rigor of reasoning, we propose and demonstrate the applicability of two broad criteria: the coherence of the argumentation; and the process through which researchers have arrived at what they regard as the best explanation for their findings (Lipton 2004). There may be other criteria, such as transparency (Aguinis & Solarino, 2019; Reinhardt et al., 2018), which others have related to rigor, but we consider these here as secondary to the strength of the core inferential processes that are involved in reasoning (Lipton, 2004).

We recognise that coherence is a contested concept in philosophy and cognitive science, but that in general it implies a lack of contradiction within the components which support a claim to truth or a belief (see Thagard, 2000). When we use the term here, we are not pretending to use it in the strict formal sense in which a logician might use it, but in the way in which it is commonly used in everyday language and in the applied reasoning practices of management researchers, as referring to consistency or fit between parts of a whole. The extent to which the parts of a process of inferential reasoning can be shown to fit together and not involve any contradictions, suggests that the process is internally consistent and, thus, to that extent rigorous.

We argue that there are two key components of coherence in a process of reasoning. First, there is methodological coherence, which refers to the extent to which ontological, epistemological and methodological assumptions are consistent with each other and the method employed and the extent to which the resulting theoretical claims also line up with these assumptions. The starting point for methodological coherence is one's ontological assumptions, which in turn should inform epistemology, methodology, method and finally the kinds of theoretical claims which arise from a study (for discussion of the importance of aligning the assumptions of a research study with methods and claims see: Bell, Bryman and Harley 2019: 25-34; Crotty 1998: 1-12). We do not regard any particular set of ontological assumptions as harbouring the potential of rigor more than any other – the key here is that the epistemology, methodology and so on which follow them should be consistent with them and with one another. Thus, while as noted above we are not using coherence in the strict sense that a logician might, nonetheless the same principle applies in the sense that there should not be contradictions between the elements which form the chain from ontology to theoretical claims.

Second, logical consistency as part of coherence refers to the extent to which there are explicit logical links provided in a paper which lead from data to theoretical presuppositions, to candidate theoretical frameworks and to actual theoretical inferences. Again, this implies that each part of an argument leading to theoretical inferences should be consistent with each other part, with no apparent contradictions. Moreover, the links between the steps in the argument should be spelt out clearly in a manuscript.

There is of course overlap between these two components of coherence, but nonetheless we treat them separately. This is because one is focused primarily on philosophical assumptions and the way that they inform methods and interpretations of data, while the other is focused on the explication of the steps which take a researcher from empirical data to firm theoretical conclusions. The same fundamental principle applies to both, as we have suggested, in that there should be a clear explication of linkages between components and a lack of contradiction between them.

Perhaps at face value both aspects of coherence would seem obvious to most qualitative researchers. However, when put to the test in a single piece of research, they are harder to practice. There may be slippages in the inferential steps; such that a qualitative researcher ends up, for example, with a variance based theoretical model that is not consistent with the constructivist or processual underpinnings at the ontological level that she or he started from (Amis & Silk, 2008). As such, we propose that the general principle of coherence is still an important general criterion to strive for in the conduct and writing up of qualitative research.

Turning to the second criterion, the notion of ‘inference to the best explanation’ reflects the fact that the theoretical conclusions which a researcher can draw from a given piece of empirical data are, as already mentioned, always ‘underdetermined’ by the data

themselves (Lipton 2004). Except in genuinely deductive reasoning, which excludes the use of empirical data, researchers always reason inductively and thus must justify the explanatory conclusions that they arrive at in relation to their data (Mantere & Ketokivi, 2013).

This general challenge of inductively warranting theoretical conclusions from data is generally known through the question of 'what is this a case of', or as the general conceptual leap from qualitative data (Klag & Langley, 2013). However, besides the general point, there has been very little consideration of the inferential linkages – offering the warrant – between empirical data and theoretical claims, as reasonable conclusions. The term 'inference to the best explanation' is often used interchangeably with abductive reasoning (Behfar & Okhuysen, 2018) and implies a process whereby, confronted with a finding or findings from the analysis of empirical data, the researcher attempts to develop the 'best' explanation which they can offer for the findings from among several possible explanations. This principle of inference to the best explanation has a long history. Thagard (1978) cites Darwin's evolutionary theory, Lavoisier's oxygen theory of combustion and Huygen's wave theory of light as examples of situations where, based on empirical data, scholars sought to develop and justify theoretical explanations.

We suggest here that this second criterion for rigorous reasoning comprises three sub-criteria. Each of these can be understood as principles and practices which can be applied in processes of reasoning, in order to seek to develop a robust and compelling – in short 'rigorous' – explanation for findings. We treat these criteria as distinct from those which comprise the criterion of coherence above, because these latter criteria relate to the specific steps which might be applied in seeking to develop an explanation for findings, over and above the coherence of the resulting argument. That is, they specify practices which are more specific than mere logical coherence. Thus, while logical coherence is vitally important, these

latter criteria specify practices which may be employed to aid a particular form of logical reasoning which we argue enhances rigor.

The first element of inference to the best explanation is the consideration of several competing explanations for empirical findings, rather than simply inferring a single explanation[1]. As noted above, the theoretical inferences drawn from empirical data are underdetermined. This implies that it is not credible, in terms of reaching conclusions, to propose an explanation as if there were no alternatives considered as part of the research process. If one is to make the case that one's explanation of a finding is 'best' (which does not imply that it is the only explanation, but rather that a systematic case can be made that it should be preferred over others), it is surely necessary to demonstrate that this has been determined through a systematic consideration of plausible rival explanations. Such explanations may come from a variety of sources, for example the overall theoretical framing of the research project, from reference to prior work, or intuition, or from 'common sense' (Klag & Langley, 2013). To an extent, we would argue that it does not matter where the explanations emerge from providing the researcher makes the abductive process explicit and can demonstrate that the proposed candidate explanation is warranted as plausible.

The warrant here is crucial. The warrant in qualitative research projects is an inductive generalization; asking of the researcher to inductively enumerate – that is, list – why the proposed candidate explanation has more correspondences with the data than any reasonable alternatives. It is possible that a warrant may also stem from a specific set of assumptions or an existing theoretical tradition, offering a principled argument in a certain

[1] We assume here that all forms of qualitative research and theorizing are geared towards some form of explanation over and beyond just a description of a phenomenon (Cornelissen, 2017). This general notion of explanation includes the traditional qualitative focus on 'interpretation'; as similarly involving a process of inferring specific readings as renderings of a text or phenomenon that make it understood (Mantere & Ketokivi, 2013).

direction (Ketokivi et al., 2017). Yet, in qualitative research projects, such a correspondence needs to be compatible with any induced generalizations from the data.

The second component of inference to the best explanation, contrastive reasoning, provides an important means to focus explanation and thereby to narrow the range of plausible theoretical inferences one might draw. Contrastive reasoning is a process whereby, confronted by a finding, the researcher asks 'Why this finding and not another one?' (Folger and Stein 2017). The logic here is that when we make inferences, we are seeking to explain a finding. Typically we wish to explain something specific, rather than all aspects of a phenomenon. The more specific we can be about what it is that we are seeking to explain, the more likely we can focus on those things which are most relevant in developing an explanation and thus develop an explanation which we can argue is a sound one.

Advocates of contrastive reasoning argue that such reasoning oftentimes entails a fact (the finding) and a foil (a hypothetical finding) and that we should contrast the fact and the foil to gradually focus our theorising on those explanatory factors that are most useful in explaining why the fact and not the foil was found (Lipton 2004). For example, if one asks why a manager fired a certain employee, the explanation which one develops is likely to be different depending on whether the foil one poses is that an employee has been fired or that a manager has made a decision (to fire a person). Explanations which are developed through such forms of contrastive counterfactual reasoning are more likely to be 'best', because they are more likely to be zooming in on precisely the aspect of the phenomenon which requires explanation and thus to have more explanatory power than if this were not the case (Tsang & Elsaesser, 2011). We in turn consider the reasoning demonstrated in a qualitative research paper more rigorous if rather than offering highly stylized or general explanations, the researcher(s) has/have engaged in such contrastive reasoning to zoom in on potent

explanations. In doing so, a researcher has shown that they have gone through a reflective process of entertaining possible emphases as alternatives (as theoretical ‘frames’) and in the process, has made a more persuasive case for the explanation that she or he has formed and puts forward.

The third component also provides another criterion for assessing explanations and working towards the ‘best’ one. An obvious challenge in a situation where there are various plausible rival explanations is how to discriminate between them and decide which one is to be proposed as the preferred explanation. In choosing among competing explanations, it is important not simply to accept the explanation which supports our implicit or explicit theoretical assumptions, but rather to aim for maximum explanatory power (Thagard 1978). Lipton (2004) conceptualises this in terms of the ‘likeliest’ explanation (the explanation which seems most probable based on our assumptions) and the ‘loveliest’ explanation (the explanation which most fully explains the findings) and argues that in making an inference to the best explanation one must weigh up ‘likeliness’ and ‘loveliness’ of possible explanations before reaching a conclusion. Some arguments will be extremely ‘lovely’ in that they elegantly explain all aspects of a set of findings, but so improbable that they must be rejected. Some arguments may be highly probable, but explain so little that they will be rejected. In practice, according to Lipton (2004), often the loveliest and likeliest explanations will converge significantly.

This consideration of the virtues of various explanations within the philosophy of science resembles discussions about the qualities of social science research. Abbott (2004) distinguishes between elegant stylized explanations that put form over substance, based on their ‘parsimony’ or ‘generality’ versus explanations that are ‘simple’ yet effective in explaining a portion of social reality. Weick (1999) and Cornelissen (2017) have similarly in

the context of qualitative research referred to the trade-offs between probable explanations that provide a simple cut on a particular reality (likeliness) versus ones that are accurate enough and describe a setting in rich detail (loveliest). In most instances of qualitative research, we expect the reasoned choice of an explanation to involve a well-articulated balance between theorized abstractions that, whilst simple, still adequately reflect the setting studied (Cornelissen, 2017). In relation to this third sub-criterion, we propose that qualitative researchers specify the virtues, or qualities, of the explanations that they put forward, explicating themselves the grounds upon which they selected their theorized explanation (rather than implicitly abiding by the virtues that are tied into a particular template). The more they do so explicitly in a manuscript, the more we consider the resulting reasoning about their inferred explanations rigorous.

Now that the criteria have been mapped out, an obvious question is what this means for qualitative research. Addressing this question will be the task of the next section of the paper, in which we will explore the application of these criteria by examining papers which in our view demonstrate rigor and providing some practical guidance for qualitative researchers.

RIGOR IN PRACTICE: PRACTICAL RECOMMENDATIONS FOR QUALITATIVE RESEARCHERS

In this section, we take the general criteria which we have set out above and consider how qualitative researchers might put them into practice and thereby increase the rigor of their reasoning processes. Before doing so, we need to be very clear that our aim in developing criteria which might be applied to different kinds of qualitative research is not to propose templates. Rather, our aim is to demonstrate how the criteria we have proposed might form a starting point for management researchers working in different qualitative traditions to think reflexively about how to increase the rigor of their reasoning processes. In

a sense, we are proposing, in the spirit of reasoning, to advocate a focus on rigor in our reasoning as a set of epistemic virtues that we would encourage researchers to pursue. In addition, we are also not proposing specific normative or prescriptive criteria for reasoning, which would assume perfect rationality on the part of researchers and the ability to apply formulaic or 'computational' reasoning processes. Rather, we accept Mantere and Ketokivi's (2013) argument that reasoning processes involve active and imperfect forms of sensemaking, whereby the theory which is created through reasoning reflects in part the idiosyncrasies of researchers. Because of these factors and the absence of perfect rationality (Mantere & Ketokivi, 2013), what counts as the best explanation for any finding will always be specific to a case and the criteria that we propose recognise this fact.

It is also important at this point to be clear that ultimately the rigor of a research project and its outputs will be assessed based on written outputs in the form of articles or reports. To show others that our research is rigorous, our focus must be on showing and telling readers what we have done in relation to the suggested criteria. In this sense, the practice of writing and the rhetorical devices authors use to persuade readers that their conclusions are justified (see Locke and Golden-Biddle, 1997; Langley and Abdallah 2011) cannot be decoupled from considerations of rigor.

Our approach to providing recommendations to qualitative researchers reflects our concerns about the potential problems arising from the use of templates, our desire to avoid overly prescriptive criteria and our recognition that rigor is ultimately demonstrated through the written presentation of our research. In the discussion below, we use examples from published work which we regard as demonstrating rigor regarding each criterion.

We do not advocate that others slavishly follow the same procedures that are used in the examples we discuss. Rather, we hope that the examples we discuss and the advice we

provide will stimulate reflexive thinking on the part of researchers. The example papers do not represent the output of a systematic review. Rather, we chose them purposively on the basis that we regard them as providing very good illustrations of the criteria for rigour in reasoning. In selecting examples, we chose papers which were qualitative; were quite recently published; which did not adopt standard templates; and which varied in their methodological approaches.

We make no claims about whether the papers we draw on are *overall* good examples across all criteria, but only that they provide useful examples of the specific criteria which they are used to illustrate. Table 1 provides a summary of the criteria we framed in the previous section, lists useful example articles and sets out questions we suggest that researchers ask themselves as a means to focus their thinking with reference to each criterion. In the discussion below, we focus in detail on only one (and in one case two) example paper for each criterion, but briefly discuss and list in the table others which we believe researchers may find useful to consult to see different ways of addressing the criterion.

- TABLE 1 ABOUT HERE -

The first criterion, coherence, comprises two parts. The first of these, methodological coherence, entails seeking to ensure that ontological, epistemological and methodological assumptions are made explicit and are consistent with the method employed and the claims about reality which are made. Whilst this may seem an obvious point, our observation is that it is remarkably common for authors to omit to discuss such issues in their papers.

Methodological Coherence

A recent paper which provides a useful example of rigor regarding this criterion is Einola and Alvesson's (2019) process study of teams. Early in their paper, the authors

consider how others have researched team processes, then make a very clear statement of their ontological position: "…we take a radically different approach [from prior work on teams] and work with another ontological view on process. Instead of studying team processes, we study *teams as a process*…we are more interested in the becoming of teams, or *teaming*, than teams as static or dynamic entities" (p. 2). They then go on to elucidate this further: "Process then refers to the ontology of the team, viewed as something people do and a matter of interactions based on and further fuelling subtle and shifting experiences and meanings" (p. 5).

Having specified their ontological position, they clarify their epistemological position, arguing that from a process perspective the way to engage with the reality of teams as processes is to understand the sensemaking of participants: "In this study, we consider sensemaking as an organizational process enabling (or disenabling, as we shall see later) the accomplishment of other key organisational processes, in this case the becoming of a team. Sensemaking is thus a core aspect and a way to understand how teams are processed" (p. 7). In effect, they are suggesting that if a team exists as a process rather than an entity (ontology), then to know the reality of team processes it is necessary to understand how participants understand and experience that process (epistemology).

The methodological approach is also clearly specified and follows logically from the ontological and epistemological assumptions underlying the project. The authors describe their methodological approach as falling within the "interpretive sensemaking tradition to case studies" (p. 7). They argue that they adopt a methodology which reflects the assumption that knowledge is generated through understanding the understandings of participants, and in which the researcher is recognised as an active participant. Following from this latter point,

they also reference ‘self ethnography’, an approach in which the researcher is familiar with the setting and has direct access to it.

In this study, the setting is provided by a set of 22 teams of graduate business students, dispersed geographically and diverse in terms of nationality, working virtually on a series of assignments, which were set by the authors. The authors characterise these teams as “global virtual teams” (p. 8). Reflecting the ontological, epistemological and methodological underpinnings, data collection was longitudinal (over a three-month period) and drew on a range of sources, with a very clear focus on the experiences and understandings of participants.

These sources comprised explanations of the teams’ ways of working written collectively by team members, multiple essays and videos produced individually by team members and reflecting on the functioning of the team, and finally interviews with members of those teams which appeared to be functioning effectively as assessed by the quality of their output of team assignments and lack of reported open conflict. In parallel, one author kept a personal diary, in which she recorded informal observational data, communication with participants and course instructors, and discussion of observations with both students and staff. The method is made very explicit in the paper, and it appears entirely consistent with the underlying assumptions in the sense that the aim is to explore participants’ understandings of processes, without denying the active role of the researchers in co-creating knowledge.

Einola and Alvesson (2019) then explicitly position their analytical strategy with reference to their underlying ontology and epistemology, in the process explaining why they eschew templated coding approaches which they associate with “inductive positivist research” (p. 10). They explain their approach as follows:

> …we applied the principles of live coding, in which the systematic, procedural nature of the inert coding is substituted by a more open-ended and imaginative way of working with research materials, joining validation and discovery that aims at new interpretations and theoretical sightlines, instead of building conceptual walls between validation and discovery…a careful hermeneutic reading and re-reading of our material as it accumulated, gradually conveying a feeling for the parts (team member narratives as they made sense of their team experience over time) and the whole (combining the individual narratives to team narratives, then contrasting the teams)…constant efforts to make sense of specific examples of material that seemed to be representative for significant parts of the whole, while conveying some important clues of something specific and 'deeper' that needed to be discovered.(pp. 10-11).

What strikes us about this approach is that there is no pretence of an analysis done by purely disinterested observers, but rather a recognition that the knowledge is produced out of an interpretation of a variety of data, notably data which aims to reflect respondents' interpretations. This, of course, reflects the explicit assumptions of the researchers.

Finally, they present their theoretical claims about the processes they studied, in a way which is consistent with assumptions, methodology and method. These take the form of providing accounts of three teams, organised around consistent themes which were developed in the course of analysis and which focus on processes and sensemaking. They cover the dynamics of team processes, the apparent beliefs of members about how the team should function, the apparent meanings of team membership for different members, and so on. In each case, they draw conclusions about how sensemaking appears to contribute to the becoming (or 'unbecoming') of teams. For example, summarising what appeared to have happened in one of the teams they write: "Diverging individual sensemaking fuels distance, conflict and the corrosion of the team. Unmaking thus prevails" (p. 15). On the basis of their comparisons of the processes in the three teams, the authors ultimately reached a series of conclusions about the ways that sensemaking played a role in shaping the dynamics of team processes. The authors explicitly state that they do not seek to generalise their findings, but rather to enrich the field of studies of teamwork by showing the insights which can be gained

by studying “the becoming of teams” (p. 25) as an alternative to positivist studies which ignore process and instead focus on generalizable, universal models.

We regard Einola and Alvesson (2019) as a useful illustration of rigor in reasoning, when judged against the criterion of methodological coherence. To be clear, we make no judgement of the overall rigour of the article, but use it here only to illustrate this criterion (although we return to it later in this section when we address other criteria). Nor do we make any judgement of the virtue of the ontological assumptions from which the other pieces of the methodological chain follow. Rather, we would suggest that given those assumptions, what follows is characterised by a high level of methodological consistency. Space constraints do not allow such detailed discussion of additional examples, but we list two in Table 1, which start with quite different ontological assumptions. These are Kothiyal, Bell and Clark’s (2018) study of identity work in Indian business schools and Ozanne and Appau’s (2019) paper on the consumption of religion and spirituality in Ghana.

We offer as a practical recommendation that qualitative researchers state their assumptions in their paper and work coherently through them so as to offer a clear and logical basis for the theoretical claims that they are putting forward. We recognize here that authors may have various options to demonstrate such methodological consistency, and may (compared to the example set out above) wish to offer fewer references or signposts to their methodological suppositions. The discussed example is based on a self-proclaimed ‘radical’ counter-factual or antidote to existing research on teams, which then asks of the authors to provide sufficient detail on their own methodological assumptions and to offer these as a coherent alternative. In other instances, a piece of qualitative research may be part of an established theoretical tradition or stream of research, and may thus require less of a detailed and deliberative approach towards seeding novel assumptions (Alvesson & Sandberg, 2011).

In such instances, authors may instead weave their ontological and epistemological position into the theoretical framing and review sections of their paper, as well as in their method section; and then ensuring that there is a logically coherent thread running through the paper, from its premise to its logical conclusions in the form of the nature of the theoretical claims that are put forward. The questions which we set out in the final column of Table 1 are intended to encourage a focus on attending to these recommendations.

Logical Consistency

Our focal example to illustrate logical consistency in practice is Bucher, Chreim, Langley and Reay's (2016) study of professional associations' responses to attempts by government to change boundaries between professions in the healthcare sector. The context for this study was a government initiative in Ontario, Canada to encourage inter-professional collaboration (IPC) in the health sector. The focus of the research was on how professions struggled over professional boundaries. The specific research questions which the authors sought to answer were: "How do professions engage in discursive boundary work in response to initiatives aimed at reshaping professional practice? And How do the field positions of professions influence the types of discursive boundary work that they engage in?" (p. 498).

In our estimation, Bucher et al (2016) present an extremely logical, clear and coherent account of the steps they went through in analysing data, reporting findings and then moving to a theoretical explanation for the findings. The primary data for the study comprised a report from an advisory council to the Ontario health department proposing regulatory changes, and submissions in response from five professional bodies (representing physicians, registered nurses, psychologists, registered practical nurses and psychological associates). A minor and secondary source was Ontario's legislation for governing health professions, which was used to cross-check prior published research in the process of classifying

professions in terms of field position (defined here in terms of the status of the professional group and its centrality to the health field). The authors specify the documents comprising the primary data set very clearly in the text and then summarise them in a table (p. 503) which specifies author of the text, components of the document and its length. Crucially, they provide a link to the submissions on the relevant government websites (p. 502), thereby facilitating easy access for readers who wish to assess findings against the dataset.

The authors then specify clearly the analytical approach: "Data analysis included two stages. First, we examined the professions' field positions in terms of status and centrality. Second, we inductively analyzed the discursive boundary work strategies used by the professions, and related these to field positions." (p. 502). The analysis of field positions was fairly straightforward and brief, based on published research and government sources. The analysis of the submission to explore discursive strategies, which necessarily draws on the analysis of field position, provides a much richer set of examples of logical consistency and we will focus on it here for this reason. As the passage below shows, the authors set out the analytical steps very clearly and logically.

> First, we identified *themes* that were indicative of boundary work by looking for explicit or implicit aspects related to professional roles and practices, capabilities, relationships, and hierarchy employed by each profession in relation to itself and other groups with whom it interacted. By moving back and forth between data and the literature, we gradually condensed the initially large number of themes into four more abstract 'foci for framing,' common across the five texts. […]
>
> In a second step, we focused on each of the professions. We explored their particular use of the four foci for framing in relationship to their field positions, extracting the most salient similarities and differences between professions' strategies. We iteratively considered extant literature to better ground our distinctions. […].
>
> In a third step, we looked across the particular uses of framing for the five professions, for patterns in the way field positions manifested themselves in framing strategies. We identified distinctive framing-field position relationships.
>
> In a final step, we returned to the literature and to our data in an effort to better understand the theoretical mechanisms underlying the relationships discovered in the previous steps.

The authors set out logically and clearly how they moved from the first step of thematic analysis around the construction of professional boundaries, to the relationships between field positions and strategies and finally to theoretical inferences about how the relationships can be explained.

As well as telling what was done, the authors subsequently show how each phase of the analysis was conducted by presenting a series of tables which summarise the findings, surrounded by explanatory text. For example, in the first phase, a table is presented for each of the professional groups in turn setting out the specific discursive framing strategies of each within each of the 'foci for framing' and presenting concrete examples from the submissions to illustrate (pp. 504-510). Each of the tables is surrounded by text which explains how the strategies were inferred from the submission, including excerpts from the text. At each stage, there is rich presentation of evidence and explanation, which readers can cross-check against the submissions should they wish.

This analysis is then condensed into a single table which maps discursive strategies against field positions (degree of centrality and status) systematically to highlight similarities and differences (p. 515). Again, this is buttressed with very detailed explanation of the associations and which then leads into the development of theoretical inferences about how associations might be explained.

Through the process of reasoning which proceeds from empirical data to theoretical inference, Bucher et al (2016) are able to construct an explanation of how professions struggle over boundaries, in which each step of the process is linked to each other step in a logical and clear way. Moreover, they devote a good deal of their paper to demonstrating the inferential process so that readers can follow closely how data analysis led to inferences.

They also do this in such a way that readers can easily follow the steps, and in ways that logically and tightly connect with one another.

In our estimation, Bucher et al (2016) present an extremely logical, clear and coherent account of the steps they went through in analysing data, reporting findings and then moving to theoretical explanations for the findings. Through the process of reasoning which proceeds from empirical data to theoretical inference, they are able to construct an explanation of discursive struggles over professional boundaries, in which each step of the process is linked to each other step in a logical and clear way. Moreover, they devote a good deal of their paper to demonstrating the inferential process so that readers can follow closely how data analysis led to inferences. They also do this in such a way that readers can easily follow the steps, and in ways that logically and tightly connect with one another.

Again, we list two additional example papers in Table 1. The first is Bolander and Sandberg's (2013) study of selection decisions made by interview panels and the second is Pouthier's (2017) study of 'griping' and joking at work. We also set out a series of questions which we hope will help scholars focus their thinking on the implementation of this criterion. We now turn to the criterion of 'inference to the best explanation' and its sub-criteria.

Considering Competing Explanations and Contrastive Reasoning

We distinguished three sub-criteria above, and discuss the first two together in sequence, before turning to the third. The reason that we are considering the first two together is that both aspects often combine as the processes driving a reasoning process (and it would thus be hard to separate them for specific examples); whereas the third sub-criterion involves a separate assessment of the warrant for an inferred theoretical explanation. In discussing all three sub-criteria, we refer to Gjerde and Alvesson's (2019) study of middle managers (with

additional very brief discussion of Cornelissen, Mantere and Vaara's (2014) paper on the so-called 'Stockwell shooting' for one of the sub-criteria). The reason for using Gjerde and Alvesson across all three sub-criteria is that it is by far the best example we are familiar with and illustrating the three sub-criteria is much easier within a single example. We do, however, list additional examples in Table 1, which we would encourage scholars to consult.

Quite early in their discussion of methodology, Gjerde and Alvesson write that "it is not possible to work exclusively inductively. Data seldom point in a clear direction" (p. 7). They explicitly reject the use of templated approaches to coding, which they suggest can lead to privileging obvious features of the data rather than digging more deeply for less obvious features which may have more value as a basis for understanding the phenomenon being researched. Our reading of this statement as a starting point for their theorising is that the authors are explicitly stating that the same empirical data can be understood in different ways depending on how one approaches it and that this necessitates 'digging deep' in considering alternative interpretations and in pursuit of the best explanation for what is observed.

The overall aim of this study is to understand how middle managers (in this case heads of departments in universities) experience and make sense of their role, 'sandwiched' between senior management and subordinates. In their interviews with middle managers Gjerde and Alvesson report that a frequently invoked (indeed *the* most frequently invoked) metaphor for their identity was that of "the umbrella carrier" (p. 9), whose role was to shelter subordinates from negative effects of directives from more senior managers. The prevalence of this metaphor in the interviews led the authors to focus on it in follow-up interviews and in their analysis, yet they did not simply accept the obvious inference that the best way to explain how middle managers experienced their roles was as protectors of subordinates. In reaching their explanation, they considered alternatives in number of ways.

First, in making sense of respondents accounts, they considered the possibility that in characterising their work in this way, respondents' responses might be self-serving, rather than simply reflecting how they experienced their roles: "it may also be, as is common in human life, an element of self-serving bias…it may be tempting for heads [of departments] to rationalise their work and importance by portraying others as in need of their…support, facilitating the head as mature, powerful and good…" (p. 7). Second, they also considered other metaphors that came out of the interviews, namely "performance driver" (p. 17), which involved pushing subordinates to improve productivity, and "impotent" (p. 17) which expressed a sense of having little real power within the hierarchy. The authors take these seriously as ways of understanding the experience of middle managers, but argue that they were much less commonly used and much less elucidatory than the metaphor of umbrella carrier, thus rejecting them (in this study, but not in possible future work) as the best way to make sense of the phenomenon in question.

Finally, in the discussion section of the paper, they propose the umbrella carrier metaphor as a new and useful way to understand the experience of middle managers and try to develop an explanation for *why* this metaphor might be invoked by respondents. That is, they move beyond the argument that this metaphor represents a useful theoretical tool for making sense of the experience of middle managers and ask why it seems to be adopted. In doing so, again a range of explanations is considered. Do they project this 'protector' role because they identify with subordinates? (p. 20). Is it, as alluded to previously, a self-serving understanding? (p. 21). Are they responding to explicit calls from others to protect the capacity of subordinates to undertake their academic work? (p. 21). Having considered this range of possible explanations, they then make an argument that the umbrella metaphor can best be explained as a way of managing the complexities of the middle-management role and making sense of 'middleness', which positions the respondents between, but attentive to, both

superiors and subordinates. In doing so, they provide a detailed explanation of their reasons for favouring this explanation, rather than simply asserting it.

The presentation of reasoning in this paper illustrates one way in which authors have sought to demonstrate this key facet of inference to the best explanation. They explicitly reject the idea that one can simply infer an explanation in a direct linear fashion, then having developed an explanation, they subject it to challenge by other competing explanations to demonstrate its explanatory power. Even then, they do not propose that it is the only explanation, but rather a new one which can be added to our ways of understanding middle management.

A particularly strong feat of this paper is that the authors entertain multiple metaphors, as theoretical imagery, for making sense of the role of middle managers and their day-to-day experiences. They contrast alternative metaphors, ranging from ones that simply position a middle manager's role (container, conduit) to ones that describe a particular activity (such as "carrying" in the "umbrella carrier" image) and ascribe varying degrees of agency to these subject positions. In doing so, they iterate between the data and what interviewees have told them and broader theoretical imagery, comparing and contrasting alternative images that best typify and explain the middle manager's situation. As part of their paper, the authors not only clearly stipulate the alternatives they have considered but have also contrasted alternatives as a way of shoring up confidence in their candidate explanation of the middle manager as "umbrella carrier". Building on this example, the practical recommendation for qualitative researchers is to spell out the inductive or abductive process that they have followed; articulating the alternative explanations they have considered and engaging in a contrastive reasoning process to draw out the warrant for their proposed explanation. The more authors do this as part of their paper, and show reflexivity in

their reasoning process from data to warranted theoretical inferences, the more robust and thus rigorous we consider their reasoning to be.

Explicating the Virtues of Explanations

The closely related step in the inferential process is to consider the virtues of the explanations that are compared and contrasted, and the degree to which a researcher demonstrates awareness and reflexivity of the strengths and limitations of the claimed explanation that they put forward. We distinguished here between explanations that are chosen because they are 'likely'; that is, simple and probable, and ones that are 'lovely' in that they explain a phenomenon in patently rich detail. In relation to this third sub-criterion, we propose that qualitative researchers specify the virtues, or qualities, of the explanations that they put forward, explicating the grounds upon which they selected their theorized explanation. The more they do so explicitly in a manuscript, the more we consider the resulting reasoning about their inferred explanations rigorous.

Returning to the Gjerde and Alvesson (2019) paper, whilst they do not use our terminology, they nicely illustrate this sub-criterion. First, they actively question for themselves, as mentioned, whether the 'umbrella carrier' metaphor is the right image to depict the situation and experiences of middle managers. They thus reflect on whether this likely explanation, which has been handed to them by their interviewees, is also at a second order, theoretical level the preferred explanation of the predicament of middle managers. They conclude in the end that it is, but they do so in ways that reflect a deep consideration of the loveliness of this explanation as well. To illustrate this, what Gjerde and Alvesson (2019) do particularly well is to reflect on whether this explanation accurately and richly captures the day-to-day experiences of middle managers. They contrast the image with more stale spatial imagery of middle managers that is prevalent in the academic literature, as well as notions of

upward voice and representation and downward managerial sensegiving and instruction. They thus conclude that it captures more than previous imagery did. They also consciously go for a metaphor, to depict middle managers' experiences, rather than other forms of theorized abstractions (e.g., propositions, process models). Metaphors typify experiences in the form or regular scenarios which as such, and compared to alternative theoretical 'products', offer rich and elaborate detail on the middle manager's role and sensemaking in practice. In other words, Gjerde and Alvesson not only reflect on the warrant for their best explanation, but also do so by considering the virtues of the explanations they are putting forward.

In addition to the Gjerde and Alvesson (2019) paper that we have already discussed, another paper that nicely illustrates this sub-criterion is the study by Cornelissen, Mantere and Vaara (2014) of the so-called 'Stockwell shooting', in which a Brazilian electrician was mistakenly identified as a potential terrorist and killed by armed police. Their study starts with the open question of why an innocent civilian was shot and what had led to this outcome. They state; "In our case, the mystery was tangible: What would explain the tragic outcome of the Stockwell shooting? How could the sensemaking of officers on the job that day, lacking sufficient evidence to support drastic action (i.e., no positive identification), result in such an outcome?" (p. 706). Taking this outcome as a starting point they follow an abductive and counter-factual strategy of interrogating various theoretical lenses, as possible candidate explanations, against the data. They mobilize lenses on emotions, communication and materiality to figure out what lens, and what associated candidate explanations, best explains the data. They ultimately conclude that no singular lens explains the case, but that the interactions across processes led to the tragic outcome. A particularly important feature of their paper is that they provide a rich narrative description of the case, where readers can trace the steps in the unfolding narrative. They then explain for each episode in the narrative

how a particular lens applies and what candidate inferences are warranted. They then go furthermore through a structured process of bringing the lenses and candidate explanations together, thinking through how explanations add up to explain the case, and how a consideration of counterfactual alternatives (that is, considering what might have been, or how the outcome could have been different depending on the interplay of events) shores up confidence in the overall model and ultimate explanation that they put forward.

We surmise that most qualitative researchers, working with rich and complex data, face this challenge of settling on an explanation that is still representative of the phenomenon that they have studied as a whole (Cornelissen, 2017). In other words, they struggle to find a justifiable balance between simplicity and richness in their explanations. In other instances, researchers may preference simplicity over richness; for example by inferring a more abstract and transferable model from a case (Gioia et al., 2013). The key here from our perspective is not to adjudicate, and formally suggest that the one virtue is preferred over the other. There is no basis for such a claim (Lipton, 2004). Instead, we offer, as already mentioned, the more practical recommendation that qualitative researchers consider these virtues and actively entertain them as the basis for the candidate explanation that they put forward. The more they convey these considerations in a paper, the more rigorous we judge their reasoning to be. The key here is that researchers reflexively show their conscious choices for the "best" explanation, and at the expense of alternatives. For example, a choice for a transferable set of propositions or process model (Gioia et al., 2013) at the expense of a thick, patterned description (Cornelissen, 2017) is a defensible choice, but one that needs to be clearly articulated and justified. The more researchers can do this in an explicit manner, and as part of their own reasoning – as opposed to referring to existing templates or protocols as legitimacy arguments – the more rigorous their research is.

In Table 1, in addition to the two papers we have discussed above, we list three example papers. These are Sasaki, Ravasi and Micelotta's (2019) study of family-owned businesses in Japan (to illustrate consideration of competing explanations), Rennstam and Karreman's (2019) study of managerial control in a high-tech firm (contrastive reasoning) and Pratt, Lepisto and Dane's (2019) study of trust among firefighters (explicating the virtues of explanations). Having discussed and illustrated the criteria of rigorous reasoning in qualitative research and enunciated practical lessons, we now turn to discuss the implications of our reframing of rigor. We elaborate the contributions which we have sought to make and why we think they are important.

DISCUSSION AND CONCLUSION

In this paper, we explore how qualitative researchers can be rigorous in how they go about reaching theoretical conclusions on the basis of empirical data. In developing a case for how this could be done, we have argued that while existing templates for qualitative analysis might be adopted by some scholars as a means to enhance rigor, and thus buttress the legitimacy of their work, in situations where such templates are applied in a proceduralist fashion or are considered wholly as a substitute for rigour they may do considerable harm.

Instead of such proceduralism, we have advocated an alternative approach, which restores the idea of rigour as being first and foremost an outcome of inferential reasoning processes linking data to theoretical conclusions. We have unfolded this perspective and have proposed a series of specific criteria which can be applied in the pursuit of reasoning that can be seen to be of a more rigorous kind – where researchers actively reason through alternative theoretical inferences and are reflective in the process whilst doing so. Based on these criteria we have also provided practical guidance to qualitative scholars who themselves wish to

pursue rigor in the conduct and writing up of their research, without necessarily having to resort to templates to do so.

Our central contribution, is to build on an emerging stream of work in management and organizational research (see Ketokivi & Mantere, 2010; Ketokivi et al., 2017; Mantere & Ketokivi, 2013) that describes how researchers reason when they make theoretical claims and offers similarly practical guidance to researchers. To this body of work, we contribute a distillation of specific reasoning practices in qualitative research and the ways in which these may amount to rigour, as an outcome of such overt reasoning. A further significant contribution is that we provide specific practical guidance to qualitative scholars on how they might increase the rigor of their own reasoning processes in carrying out and writing up their research.

In providing such guidance, we have tried to walk a fine line between specifying prescriptive rules, which would be antithetical to the version of rigor which we espouse, and providing overly general criteria which fail to provide a basis for action. In doing so, we have sought to provide guidance which can be flexibly applied in context and in ways that match the particularities of the setting and the methods used. At the heart of our approach lies a core belief in qualitative researchers focusing on methodology over technique (Hammersley, 2011); knowingly making using of techniques as part of their broader approach to the data analysis and putting forth a chain of reasoning that is transparent and defensible. Indeed, when researchers follow the tenets of the approach that we propose here, they do so in a way that recognizes the particularities of their case and engage in active and deliberate forms reasoning with a healthy dose of reflexivity.

We do not think that the approach itself can be seen as a template, nor will it likely evolve into a protocolized way of doing things. Even if specific criteria or steps (such as around inference to the best explanation) solidify into a set of commonly applied rules, these

will still only gain meaning and appreciation in context – calling on the active reasoning of both researchers as well as reviewers to make sense of the inferential connections made in each context (rather than abiding by the proxy of a template or even the mere application of a specific rule as indicating rigor). That said, we would argue that the suggestions that we make here can be adopted by those who use templates in their work and may in such instances turn out to be beneficial in the process. These suggestions may strengthen their overall reflexivity and will urge them to make considered choices and reason through how they use or adapt a certain template in context.

Furthermore, by discussing examples of published work, we have sought to go beyond *explaining* reasoning-related criteria to *illustrate* how others have applied reasoning practices consistent with our criteria. We certainly are not encouraging scholars to slavishly follow the examples that we have provided. We merely offer these examples to illustrate and demonstrate the kind of rigor that we see as associated with reasoning, in the hope that this will help scholars think about how they might similarly enhance rigor in their own reasoning. As well as using examples, we have presented a series of questions which we hope will provide a means for qualitative researchers to assess the extent to which they are applying the criteria in their own research. These questions can be used as a heuristic or thinking tool by qualitative researchers to orient them to the qualities of their own reasoning, and to improve it – make it more rigorous – in the process. Thus, in a very practical sense, we hope that together the examples and questions will stimulate reflexive thinking on the part of others as they engage in processes of reasoning.

While we hope that we have made a useful contribution to thinking about rigor in reasoning and particularly to providing practical guidance to researchers, of course we do not claim that our contribution here represents the ‘last word’ on how scholars might strengthen the rigor of their reasoning when undertaking qualitative research. Clearly there is a need for

further work in this area. We have argued that our criteria represent a useful set of principles that can be applied in reasoning, but there may be additional, or better, criteria that can be developed. Further, while we have demonstrated how the criteria might be used in some specific examples of published research, more work is required to explore the applicability and limits of the criteria across different forms of qualitative research. We would encourage qualitative scholars to take the criteria we have enunciated here and seek to develop and extend them.

In closing, we hope that this paper contributes to helping qualitative researchers find ways to strengthen the processes through which they derive theoretical conclusions from empirical data. For researchers who take seriously the need to enact and demonstrate rigor in their research, but whose work may not be suited to templates (Amis & Silk, 2008), we have tried to provide guidance about how they might strengthen the processes through which they draw inferences from qualitative data. But even more generally, we have suggested that rigour is something that is established in and through the reasoning that qualitative researchers offer, rather than seeing it as something that is inherent to the application of a standard protocol, as a template. The reasoned and reflective application of a template may establish rigor, but the rote application of templates – which we have labelled as 'proceduralism' – may push such reasoning to the background with the result that form starts to rule over function.

References

Abbott, A. (2004) *Methods of discovery: Heuristics for the social sciences*, London: Norton.

Aguinis, H. & Solarino, M. (2019) Transparency and replicability in qualitative research: The case of interviews with elite informants, *Strategic Management Journal*, in press https://doi.org/10.1002/smj.3015

Alvesson, M. & Sandberg, J. (2011) Generating research questions through problematization, *Academy of Management Review*, 36, 246-71.

Amis, J. & Silk, M. (2008) The philosophy and politics of quality in qualitative organizational research, *Organizational Research Methods*, 11, 456-80.

Behfar, K., & Okhuysen, G. A. (2018) Perspective— Discovery within validation logic: Deliberately surfacing, complementing, and substituting abductive reasoning in hypothetico-deductive inquiry. *Organization Science*, 29, 323–340.

Bell, E., Bryman, A. & Harley, B. (2019) *Business research methods* (5th Edn). Oxford: Oxford University Press.

Bolander, P. & Sandberg, J. (2013) How employee selection decisions are made in practice, *Organization Studies*, 34, 285-311.

Bucher, S., Chreim, S., Langley, A. & Reay, T. (2016) Contestation about collaboration: Discursive boundary work among professions, *Organization Studies*, 37, 497-522.

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, 368-383.

Cornelissen, J., Mantere, S. & Vaara, E. (2014) The contraction of meaning: The combined effect of communication, emotions and materiality on sensemaking in the Stockwell shooting, *Journal of Management Studies*, 51, 699-736.

Crotty, M. (1998) *The foundations of social research: Meaning and perspective in the research process*, London: Sage.

Einola, K. & Alvesson, M. (2019), The making and unmaking of teams. *Human Relations*, in press, DOI: 10.1177/0018726718812130.

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

Folger, R. & Stein, C. (2017) Abduction 101: reasoning processes to aid discovery, *Human Resource Management Review*, 27, 306-15.

Gioia, D., K. Corley and A. Hamilton (2013). Seeking qualitative rigor in inductive research: A note on the Gioia methodology, *Organizational Research Methods*, 16, 15-31.

Gjerde, S. and M. Alvesson (2019) Sandwiched: Exploring role and identity of middle managers in the genuine middle, *Human Relations*, in press, DOI: 10.1177/0018726718823243

Hammersley,M. (2011) *Methodology: Who Needs It?* London, UK: Sage.

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

Harley, B. & Cornelissen, J. (2019) Reframing Rigor as Reasoning: Challenging Technocratic Conceptions of Rigor in Management Research in Zilber, T., J. Amis and J.

Mair (eds) *The Production of Managerial Knowledge and Organizational Theory: New Approaches to Writing, Producing and Consuming Theory, Research in the Sociology of Organizations*, 59, 59–76.

Irwin, A. (2019) Remaking 'quality' within the social sciences: The debate over rigour and relevance in the modern business school, *The Sociological Review*, 67, 194-201.

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

Ketokivi, M., Mantere, S. & Cornelissen, J. (2017) Reasoning by analogy and the progress of theory. *Academy of Management Review*, 42, 637-58.

Klag, M. & Langley, A. (2013) Approaching the Conceptual Leap. *International Journal of Management Reviews*, 15: 149-166.

King, N. (2004) Using templates in the thematic analysis of text, in Cassell, C. & Symon, G. (eds) *Essential Guide to Qualitative Methods in Organizational Research*, Sage, 279-292.

Köhler, T., Smith, A. & Bhakoo, V. (2019) Feature topic for ORM: "Templates in qualitative research", *Organizational Research Methods*, 22, 3-5.

Kothiyal, N., Bell, E. & Clarke, C. (2018) Moving beyond mimicry: Developing hybrid spaces in Indian business schools, *Academy of Management Learning and Education*, 17, 137–154.

Langley, A. & Abdallah, C. (2011) Templates and turns in qualitative studies of strategy and management, in D. Bergh & D. Ketchen (eds) *Building Methodological Bridges*, Emerald Group Publishing Limited, 201 – 235.

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

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

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

Ozanne, J. & Appau, S. (2018) Spirits in the marketplace, Journal of Marketing Management, 35, 451-466.

Popper, K. (1959) *The Logic of Scientific Discovery*, London: Hutchinson.

Pouthier, V. (2017) Griping and joking as identification rituals and tools for engagement in cross-boundary team meetings, *Organization Studies*, 38, 753-74.

Pratt, M., Kaplan, S. & Whittington, R. (2019) Editorial essay: The tumult over transparency: Decoupling transparency from replication in establishing trustworthy qualitative research, *Administrative Science Quarterly*, in press, DOI: 10.1177/0001839219887663.

Pratt, M., Lepisto, D. & Dane, E. (2019) The hidden side of trust: Supporting and sustaining leaps of faith among firefighters, *Administrative Science Quarterly*, 64, 398-434.

Reinhardt, A., Kreiner, G., Gioia, D. & Corley, K. (2018) Conducting and publishing rigorous qualitative research, in C. Cassell, A. Cunliffe & G. Grandy (eds) *The SAGE Handbook of Qualitative Business and Management Research Methods*, Vol. 1, London: SAGE, 515-32.

Renstamm, J & Karreman, D. (2019) Understanding control in communities of practice: Constructive disobedience in a high-tech firm, *Human Relations,* in press, DOI: 10.1177/0018726719843588

Sasaki, I., Ravasi, D. & Micelotta, E. (2019) Family firms as institutions: Cultural reproduction and status maintenance among multi-centenary *shinise* in Kyoto, *Organization Studies*, 406, 793-831.

Thagard, P. (1978) The best explanation: Criteria for theory choice, *The Journal of Philosophy*, 75, 76-92.

Thagard, P. (2000) *Coherence in thought and action*, Cambridge, Mass: MIT Press.

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

Tsang, E. & Ellsaesser, F. (2011) How contrastive explanation facilitates theory building, Academy of Management Review, 36, 404-19.

Weick, K.E. (1999) Theory construction as disciplined reflexivity: Tradeoffs in the 90s. *Academy of Management Review* 24, 797-806.

TABLE 1: A FRAMEWORK FOR APPLYING THE CRITERIA

	CRITERIA	EXAMPLE PAPERS	QUESTIONS
COHERENCE			
Methodological Coherence	The ontology, epistemology, methodology, method/data and theoretical claims in a study are consistent with each other.	*Einola and Alvesson (2019) Kothiyal, Bell and Clarke (2018) Ozanne and Appau (2019)	• What are the ontological assumptions underpinning this piece of research? • What are the epistemological assumptions underpinning this research and are they consistent with my ontological assumptions? • What is my methodology and is it consistent with my ontological and epistemological assumptions? • Are ontology, epistemology and methodology all made explicit if this is necessary? • Is my data collection and analysis consistent with my methodology and underlying assumptions? • Are the claims I make about reality, based on my data, consistent with my methodology and underlying assumptions?
Logical Consistency	There are explicitly articulated and logical links between each part of the explanation of why theoretical claims are inferred from empirical data.	* Bucher, Chreim, Langley and Reay (2016) Bolander and Sandberg (2013) Pouthier (2017)	• Have I clearly and precisely specified the data that form the empirical basis of the research? • Have I clearly and precisely specified the focus of my analysis, ie. what I am looking for in the analysis? • Have I made explicit any theoretical resources which inform my analysis? • Have I explained the steps of the analytical and inferential processes sequentially and explained how each links to the next? • Have I provided sufficient data and explanation to demonstrate the basis on which inferences were made and allow readers to make sense of it and check it? • Are there clear and explicit links from the inferences made in each step of the analysis to the overarching theoretical explanation which is ultimately inferred?
INFERENCE TO THE BEST EXPLANATION			
Consideration of Competing Explanations	In inferring an explanation for findings, several possible explanations are considered.	* Gjerde and Alvesson (2019) Cornelissen, Mantere and Vaara (2014)	• In reaching a provisional inference from the data, have I interrogated them thoroughly? • Have I justified the provisional inference?

		Sasaki, Ravasi and Micelotta (2019)	• In deciding which is the preferred explanation, what other explanations have I considered? • Have I given them serious consideration? • Have I justified adequately why I have chosen my preferred explanation ahead of the rival ones?
Contrastive Reasoning	In reaching theoretical explanations based on empirical findings, asking 'Why this finding and not a different one?'.	* Gjerde and Alvesson (2019) Cornelissen, Mantere and Vaara (2014) Renstamm and Karreman (2019)	• In explaining a finding what elements of it seem salient and important? • In comparing alternative explanations, what is the crucial difference between them? • In selecting a preferred explanation, what others are rival candidate explanations? Can these be ruled out or accounted for? • Have I justified adequately why I have chosen my preferred explanation ahead of the rival ones?
Explicating the Virtues of Explanations	In choosing a theoretical explanation, not immediately accepting the 'likeliest', but seeking the one with the most explanatory power (the 'loveliest').	* Gjerde and Alvesson (2019) * Cornelissen, Mantere and Vaara (2014) Pratt, Lepisto and Dane (2019)	• In considering and comparing alternative explanations, which ones seem simple and sound? • In considering and comparing alternative explanations, which ones seem elegant, yet rich and detailed? • How do I balance between the virtues of simplicity and richness in my choice of a preferred explanation?

* indicates the 'focal' pieces which we discuss in detail in this paper. The remainder of the citations are articles which we believe provide useful examples.