The Academy of Management Journal has taken the lead among major journals in spotlighting alternative methods that take advantage of rich empirical data. In a series of “From the Editor” commentaries, scholars cogently have explicated related topics such as qualitative research (Gephart, 2004), grounded theory building (Suddaby, 2006), the value of richness (Weick, 2007) and the persuasive power of the single case (Siggelkow, 2007). In this commentary, we focus on the related research strategy of theory building from cases, particularly multiple cases.

Scholars have used case studies to develop theory about topics as diverse as group process (Edmondson, Bohmer, & Pisano, 2001), internal organization (Galunic & Eisenhardt, 2001; Gilbert, 2005), and strategy (Mintzberg & Waters, 1982). Classic scholars (Chandler, 1962; Whyte, 1941) as well as the authors of highly regarded AMJ papers (Dutton & Dukerich, 1991; Sutton & Raphaeli, 1988) have used the method. Indeed, papers that build theory from cases are often regarded as the “most interesting” research (Bartunek, Rynes, & Ireland, 2006) and are among the most highly cited pieces in AMJ (e.g., Eisenhardt, 1989a; Gersick, 1988), with impact disproportionate to their numbers. Not surprisingly then, the winning authors (Ferlie, Fitzgerald, Wood, & Hawkins, 2005; Gilbert, 2005) of the most recent AMJ Best Article Award relied on this method.

Building theory from case studies is a research strategy that involves using one or more cases to create theoretical constructs, propositions and/or midrange theory from case-based, empirical evidence (Eisenhardt, 1989b). Case studies are rich, empirical descriptions of particular instances of a phenomenon that are typically based on a variety of data sources (Yin, 1994). Cases can be historical accounts, such as Weick’s (1993) study of the Mann Gulch fire, but they are more likely to be contemporary descriptions of recent events, such as Gilbert’s (2005) study of adaptation to discontinuous environmental change by newspaper organizations. The central notion is to use cases as the basis from which to develop theory inductively. The theory is emergent in the sense that it is situated in and developed by recognizing patterns of relationships among constructs within and across cases and their underlying logical arguments.

Central to building theory from case studies is replication logic (Eisenhardt, 1989b). That is, each case serves as a distinct experiment that stands on its own as an analytic unit. Like a series of related laboratory experiments, multiple cases are discrete experiments that serve as replications, contrasts, and extensions to the emerging theory (Yin, 1994). But while laboratory experiments isolate the phenomena from their context, case studies emphasize the rich, real-world context in which the phenomena occur. The theory-building process occurs via recursive cycling among the case data, emerging theory, and later, extant literature. Although sometimes seen as “subjective,” well-done theory building from cases is surprisingly “objective,” because its close adherence to the data keeps researchers “honest.” The data provide the discipline that mathematics does in formal analytic modeling.

A major reason for the popularity and relevance of theory building from case studies is that it is one of the best (if not the best) of the bridges from rich qualitative evidence to mainstream deductive research. Its emphasis on developing constructs, measures, and testable theoretical propositions makes inductive case research consistent with the emphasis on testable theory within mainstream deductive research. In fact, inductive and deductive logics are mirrors of one another, with inductive theory building from cases producing new theory from data and deductive theory testing completing the cycle by using data to test theory. Moreover, since it is a theory-building approach that is deeply embedded in rich empirical data, building theory
from cases is likely to produce theory that is accurate, interesting, and testable. Thus, it is a natural complement to mainstream deductive research.

But while theory building from cases is increasingly prominent, challenges in writing publishable manuscripts using this research strategy exist. Some reviewers who work on large-scale, hypothesis-testing research may misunderstand the method (e.g., expect random sampling), or simply regard their own methods as superior. Some reviewers who work with other research strategies that also use rich empirical data (e.g., naturalistic inquiry) may emphasize thick narrative descriptions but be less interested in generating testable and generalizable theory. Still other reviewers may be sympathetic to research that is based on rich empirical evidence but be confused by the jumble of labels used to describe such research, which include grounded theory building, qualitative research, theory building from cases, and naturalistic inquiry. Having been involved with numerous research projects and written many papers that develop theory from cases, we are particularly sympathetic to authors. So, our purpose is to highlight the opportunities that differentiate building theory from cases from other research strategies, describe some of its common challenges, and suggest possible antidotes.

**Justifying Theory Building**

Sound empirical research begins with strong grounding in related literature, identifies a research gap, and proposes research questions that address the gap. But when using theory building from cases as a research strategy, researchers also must take the added step of justifying why the research question is better addressed by theory-building rather than theory-testing research. The implicit assumption is that theory building from cases is less precise, objective, and rigorous than large-scale hypothesis testing. Moreover, failure to convince readers that a theory-building study is warranted in the first few pages can sink a manuscript before readers ever reach the findings. In other words, readers may ask, *So why is this an inductive study?*

A key response to this challenge is to clarify why the research question is significant, and why there is no existing theory that offers a feasible answer. Conflicting theories are not enough. Rather, it is critical to convince readers that the research question is crucial for organizations and/or theory, and demonstrate that the existing research either does not address the research question at all, or does so in a way that is inadequate or likely to be untrue.

An example is early research on making fast strategic decisions (Eisenhardt, 1989a). The introduction makes a strong case that fast strategic decision making is crucial for firm performance in high-velocity environments, including an example of a firm that died because its executives decided slowly. The introduction then goes on to demonstrate that the research literature has mostly ignored this critical topic. The background section describes several ideas from the literature that address speed (albeit obliquely), but then shows that the logic underlying these ideas is unconvincing. For example, although some of the literature implies that centralized strategic decision making might be fast, centralization could not solve problems of access to relevant information, implementation, and confidence to decide in the face of uncertainty. Thus, it is unlikely that fast decision making is simply a matter of centralization per se. The background section concludes by asking whether a “snap decision” process could actually be realistic.

The challenge of justifying inductive case research partially depends on the nature of the research question. For theory-driven research questions that extend existing theory (Lee, Mitchell, & Sabyinski, 1999), a researcher has to frame the research within the context of this theory and then show how inductive theory building is necessary. Typically, the research question is *tightly* scoped within the context of an existing theory, and the justification rests heavily on the ability of qualitative data to offer insight into complex social processes that quantitative data cannot easily reveal. For example, Greenwood and Suddaby (2006) studied how a known instance of institutional change at the center of a field occurred (i.e., promotion of change by elite firms within the accounting profession). They justified their approach in terms of extending institutional theory and the ability of qualitative data to explicate the complex social processes involved.

In contrast, for phenomenon-driven research questions, a researcher has to frame the research in terms of the importance of the phenomenon and the lack of plausible existing theory. Here the research question is *broadly* scoped to give the researcher more flexibility. The justification rests on the phenomenon’s importance, and the lack of viable theory and empirical evidence. For example, Bingham and Eisenhardt (2006) justified their study of what executives learn when they engage in a repeated organizational process (in their study, internationalization) by observing that learning is a ubiquitous process, and yet the vast empirical literature on learning ignores the content of what is actually learned. More broadly, theory-building research using cases typically answers research questions that address “how” and “why” in unexplored
Comparisons that clarify whether an emergent finding is simply idiosyncratic to a single case or consistently replicated by several cases (Eisenhardt, 1991). Multiple cases also create more robust theory because the propositions are more deeply grounded in varied empirical evidence. Constructs and relationships are more precisely delineated because it is easier to determine accurate definitions and appropriate levels of construct abstraction from multiple cases. For example, Brown and Eisenhardt (1991) found that, although some firms used alliances to experiment with the future, others used futurists and exploratory products. With multiple cases, the authors set an appropriate level of abstraction (i.e., probes) that was more accurate than the individual instantiations (e.g., alliances, exploratory products). Multiple cases also enable broader exploration of research questions and theoretical elaboration. For example, Brown and Eisenhardt (1998) added successful and unsuccessful turnaround cases that enabled them to add further longitudinal elements to their theory. Because case numbers are typically small, a few additional cases can significantly affect the quality of the emergent theory. For example, adding three cases to a single-case study is modest in terms of numbers, but offers four times the analytic power. Thus, theory building from multiple cases typically yields more robust, generalizable, and testable theory than single-case research.

But although multiple cases are likely to result in better theory, theoretical sampling is more complicated. The choice is based less on the uniqueness of a given case, and more on the contribution to theory development within the set of cases. That is, multiple cases are chosen for theoretical reasons such as replication, extension of theory, contrary replication, and elimination of alternative explanations, and elaboration of the emergent theory.

Theoretical sampling of single cases is straightforward. They are chosen because they are unusually revelatory, extreme exemplars, or opportunities for unusual research access (Yin, 1994). For example, Weick (1993) used an extreme case of lost sensemaking in the wilderness fire-fighting disaster at Mann Gulch; Galunic and Eisenhardt (1996, 2001) examined organizational adaptation in an exemplar firm that was the highest performing technology-based corporation in the world for several decades; and Dutton and Dukerich (1991) studied the New York Port Authority, where they had unusual access through friends. Thus, single-case research typically exploits opportunities to explore a significant phenomenon under rare or extreme circumstances.

But while single-case studies can richly describe the existence of a phenomenon (Siggelkow, 2007), multiple-case studies typically provide a stronger base for theory building (Yin, 1994). Again, to use the analogy of laboratory experiments, the theory is better grounded, more accurate, and more generalizable (all else being equal) when it is based on multiple case experiments. Multiple cases enable comparisons that clarify whether an emergent finding is simply idiosyncratic to a single case or consistently replicated by several cases (Eisenhardt, 1991). Multiple cases also create more robust theory because the propositions are more deeply grounded in varied empirical evidence. Constructs and relationships are more precisely delineated because it is easier to determine accurate definitions and appropriate levels of construct abstraction from multiple cases. For example, Brown and Eisenhardt (1991) found that, although some firms used alliances to experiment with the future, others used futurists and exploratory products. With multiple cases, the authors set an appropriate level of abstraction (i.e., probes) that was more accurate than the individual instantiations (e.g., alliances, exploratory products). Multiple cases also enable broader exploration of research questions and theoretical elaboration. For example, Brown and Eisenhardt (1998) added successful and unsuccessful turnaround cases that enabled them to add further longitudinal elements to their theory. Because case numbers are typically small, a few additional cases can significantly affect the quality of the emergent theory. For example, adding three cases to a single-case study is modest in terms of numbers, but offers four times the analytic power. Thus, theory building from multiple cases typically yields more robust, generalizable, and testable theory than single-case research.

But although multiple cases are likely to result in better theory, theoretical sampling is more complicated. The choice is based less on the uniqueness of a given case, and more on the contribution to theory development within the set of cases. That is, multiple cases are chosen for theoretical reasons such as replication, extension of theory, contrary replication, and elimination of alternative explanations (Yin, 1994). For example, Graebner and Eisenhardt (2004) studied acquisition from the seller perspective by examining three replicated cases in which the executives sold their companies, a contrary replication in which executives could have sold their companies but did not, and then further cases in different industries that explored industry-level explanations. A particularly important theoretical sampling approach is “polar types,” in which a researcher samples extreme (e.g., very high and very low performing) cases in order to more easily observe contrasting patterns in the data. Although such an approach can surprise reviewers because the resulting theory is so consistently supported by the empirical evidence, this sampling leads to very clear pattern recognition of the central constructs, relationships, and logic of the focal phenomenon.
Dealing with Interview Data

Case studies can accommodate a rich variety of data sources, including interviews, archival data, survey data, ethnographies, and observations. For example, Hargadon and Sutton (1997) combined observations of brainstorming sessions, interviews with corporate actors, and ethnographies of two projects in their case study of routine innovation at Ideo. But as research incorporates more cases and moves away from everyday phenomena such as work practices to intermittent and strategic phenomena such as acquisitions and strategic decision making, interviews often become the primary data source. Interviews are a highly efficient way to gather rich, empirical data, especially when the phenomenon of interest is highly episodic and infrequent. But interviews also often provoke a “knee-jerk” reaction that the data are biased in which impression management and retrospective sensemaking are deemed the prime culprits. The prototypical reader asks, Is the theory just retrospective sensemaking by image-conscious informants?

The challenge of interview data is best mitigated by data collection approaches that limit bias. A key approach is using numerous and highly knowledgeable informants who view the focal phenomena from diverse perspectives. These informants can include organizational actors from different hierarchical levels, functional areas, groups, andographies, as well as actors from other relevant organizations and outside observers such as market analysts. It is unlikely that these varied informants will engage in convergent retrospective sensemaking and/or impression management. For example, in our study of acquisitions from the seller perspective, Graebner and Eisenhardt (2004) we relied on interviews with executives from two hierarchical levels at the selling firms, executives from two hierarchical levels at the buying firms, board members from both the buying and selling firms, and investment bankers who provided background information about M&A.

Another key approach to mitigating bias is to combine retrospective and real-time cases (Leonard-Barton, 1990). Retrospective cases rely on interviews (and archival data) that build up the number and depth of cases efficiently and so enable a researcher to cover more informants and include more cases. Such interviews are particularly accurate when the focal events are recent. In contrast, real-time cases employ longitudinal data collection of interviews and, often, observations, both of which help to mitigate retrospective sensemaking and impression management.

A more subtle challenge arises from the confusion between qualitative data and qualitative research. Theory-building cases usually rely extensively on qualitative data from interviews and other sources, such as observations, historical books, archives, and so forth. This research is often termed “qualitative” simply because it relies significantly on qualitative data. But qualitative research can also refer to the use of qualitative data in research strategies other than organizing data into cases and using replication logic to build theory. For example, Elsbach and Kramer (2003) accumulated qualitative data on individual “pitches” in their study of face-to-face interviews in Hollywood, but they pooled their data rather than organize it into cases. Adding to the confusion, some scholars have a very specific definition of what constitutes “qualitative research” that goes well beyond the type of data. For example, Gephart (2004) described qualitative research as “multimethod research that uses an interpretive, naturalistic approach to its subject matter (Denzin & Lincoln, 1994)” and “addresses questions about how social experience is created and given meaning” (Gephart, 2004: 454–455). According to this view, qualitative research is highly descriptive, emphasizes the social construction of reality, and focuses on revealing how extant theory operates in particular examples. This view is different in terms of research activities, goals, and epistemology from the more objective and positivist stance of theory building from cases as well as from other research strategies also termed “qualitative.” The key implication is that some readers will confuse different kinds of research that seem similar because they use qualitative data, and these readers may be disappointed if the research does not then match their understanding of “qualitative research.”

A straightforward approach for coping with the varied meanings of “qualitative research” is to avoid the term. Rather, clarify the research strategy being used, and contrast it with other “qualitative” approaches with differing epistemological assumptions. Specifically, when inducting theory from cases, be explicit about the theory-building goal and to liberally use footnotes that sharpen the distinctions among the multiple meanings of qualitative research. The key here is to convey the theory-building strategy clearly while avoiding confusion, philosophical pitfalls, and unrealistic reader expectations.

Presenting Empirical Evidence

A critical aspect of empirical research is presenting the evidence from which the theory of interest
was inducted. In large-scale deductive studies, there is a widespread norm of presenting theory and then empirical evidence in compact numerical tables that summarize statistical analyses of large amounts of data. But case data cannot be so tightly summarized, because much of it consists of rich qualitative detail.

In a single-case study, the challenge of presenting rich qualitative data is readily addressed by simply presenting a relatively complete rendering of the story within the text. The story typically consists of narrative that is interspersed with quotations from key informants and other supporting evidence. The story is then intertwined with the theory to demonstrate the close connection between empirical evidence and emergent theory. This intertwining keeps both theory and evidence at the forefront of the paper. Gersick (1994), Hargadon and Douglas (2001), and Mintzberg and Waters (1982) are exemplars of this approach.\(^1\)

But presenting a relatively complete and unbroken narrative of each case is infeasible for multiple-case research, particularly as the number of cases increases. If the researcher relates the narrative of each case, then the theory is lost and the text balloons. So the challenge in multiple-case research is to stay within spatial constraints while also conveying both the emergent theory that is the research objective and the rich empirical evidence that supports the theory. Coping with the trade-off between rich story and well-grounded theory is easier to do in a multicase book or a single-case paper. But in journal articles, multicase researchers face a particularly difficult trade-off between theory and empirical richness. It can be especially challenging to satisfy readers who expect the extensive narratives of single-case research. They ask, Where’s the rich story?

The best way to address this challenge of “better stories vs. better theories” is to develop a theory in sections or by distinct propositions in such a way that each is supported by empirical evidence. Thus, the overarching organizing frame of the paper is the theory, and each part of the theory is demonstrated by evidence from at least some of the cases. But since it is generally not realistic to support every theoretical proposition with every case within a text itself, the use of extensive tables and other visual devices that summarize the related case evidence are central to signaling the depth and detail of empirical grounding. In other words, the use of summary tables and aids that summarize the case evidence complements the selective story descriptions of the text and further emphasizes the rigor and depth of the empirical grounding of the theory. A separate table that summarizes the evidence for each theoretical construct is a particularly effective way to present the case evidence. These “construct tables” summarize the case evidence and indicate how the focal construct is “measured,” thus increasing the “testability” of the theory and creating a particularly strong bridge from the qualitative evidence to theory-testing research. Graebner (2004), Gilbert (2005), and Zott and Huy (2007) are excellent examples of blending construct tables with selected text descriptions.

Summarizing case evidence within tables and organizing the text around the theory can be, however, disappointing to readers who are expecting the “richness” of detailed narratives from the empirical data. This is particularly likely among readers whose research predilections favor description over theory. So, although it may seem trivial, it is usually helpful to remind reviewers that the objective is theory development. More significantly, it is critical to invest in developing well-crafted tables, appendixes, and visual aids to demonstrate the theory’s underlying empirical support and the anticipated richness of the case data, and to tie those tables clearly to the text.

**Writing the Emergent Theory**

The objective of building theory from cases is theory. But unlike in large-scale hypothesis-testing research, there is no “sure-to-please” standard template for writing emergent theory in theory-building research. Since different readers have their own preferences, they often ask, Why did you format the theory this way?

A useful way to cope with this challenge is to write the theory in multiple ways. First, sketch the emergent theory in the introduction. Then, in the body of paper, write each proposition (implicitly or explicitly stated), and link it to the supporting empirical evidence for each construct and for the proposed relationship between the constructs. When the research is well done, the propositions will be consistent with most (or even all) of the cases because the researcher has effectively “pattern-matched” between theory and data. It is also crucial to write the underlying theoretical arguments that provide the logical link between the constructs within a proposition. These arguments can be

---

\(^1\) An alternative approach is to present the story and then the theory. But this approach moves the theory off center stage and makes the empirical grounding of the theory less apparent. Nonetheless, it is a reasonable and common approach.
drawn from case evidence (e.g., an informant explaining the logic) and/or from more detached logic. Finally, provide a visual theory summary such as a “boxes and arrows” diagram or summary table. Eisenhardt (1989a), Gilbert (2005), and Maurer and Ibers (2006) offer exemplars of the multiple ways of writing theory within a single paper. Using these multiple ways to present the theory is often a safe starting point for initial manuscript submissions.

A more subtle challenge arises from confusion about the meaning of “grounded theory building.” For some scholars, grounded theory building simply means creating theory by observing patterns within systematically collected empirical data. This view often includes some notion of recursively iterating between (and thus constantly comparing) theory and data during analysis, and theoretically sampling cases (as described earlier). As Langley (1999) noted, this is a widely held view of grounded theory building. In this view, the quality of the theory and the strength of its empirical grounding are more central to research quality than the specifics of the theory-building process.

But for other scholars, grounded theory building has a more precise meaning that stems from the original focus of Glaser and Strauss (1967) on the interpretation of meaning by social actors. For example, Suddaby described grounded theory building as “most suited to efforts to understand the process by which actors construct meaning out of intersubjective experience” (Suddaby, 2006: 634). Others go further to emphasize elaborate processes (and terminology) for how researchers should gather field data and discover theory using a hierarchical structure of categories (Corbin & Strauss, 1990). Constant comparison and theoretical sampling take on precise meanings: “constant comparison” means simultaneous collection and analysis of data, and “theoretical sampling” means that decisions about which data to collect next are determined by the theory in progress (Suddaby, 2006). In this view, adherence to specific grounded theory building processes is important in judging research quality. But strict adherence can also result in theory with limited generalizability (Langley, 1999) and idiosyncratic path dependence on the particular empirical starting point.

As when coping with the multiple meanings of “qualitative research,” it is often helpful to deal with the multiple meanings of “grounded theory building” by avoiding the term unless one is actually using the Glaser and Strauss (1967) approach. It is also helpful to preempt misunderstanding by engaging in systematic data collection and theory development processes that are reported with transparent description, particularly regarding how the theory was inducted from the data (e.g., description of cross-case comparison techniques). The key here is to convey the rigor, creativity, and open-mindedness of the research processes while sidestepping confusion and philosophical pitfalls.

Finally, a surprising challenge can arise from readers who are disappointed by parsimonious theory. Particularly when readers are more familiar with the idiosyncratic detail of some single-case research, they may expect the complicated theory that can arise from such cases. Somewhat surprisingly, single cases can enable the creation of more complicated theories than multiple cases, because single-case researchers can fit their theory exactly to the many details of a particular case. In contrast, multiple-case researchers retain only the relationships that are replicated across most or all of the cases. Since there are typically fewer of these relationships than there are details in a richly observed single case, the resulting theory is often more parsimonious (and also more robust and generalizable). A key approach to dealing with this challenge is to ensure that the theory fully exploits the available evidence in terms of possible nuances and alternative interpretations. It also helps to remind readers that parsimony, robustness, and generalizability characterize superior theory.

Conclusion

Theory building from case studies is an increasingly popular and relevant research strategy that forms the basis of a disproportionately large number of influential studies. But like the adherents of any research method, its adherents face some predictable challenges, some of which have, ironically, emerged precisely because research relying on rich qualitative data is becoming more common. The good news is that these often very legitimate challenges can be mitigated through precise language and thoughtful research design: careful justification of theory building, theoretical sampling of cases, interviews that limit informant bias, rich presentation of evidence in tables and appendixes, and clear statement of theoretical arguments. The result is fresh theory that bridges well from rich qualitative evidence to mainstream deductive research. This is the hallmark of building from case studies.
REFERENCES


---

**Kathleen M. Eisenhardt** (kme@stanford.edu) is the Stanford W. Ascherman M.D. Professor of Strategy and Organizations in the Department of Management Science and Engineering and the codirector of the Stanford Technology Ventures Program, Stanford University. She received her Ph.D. from Stanford’s Graduate School of Business. Her research interests include alliance and network processes, strategy as simple rules, and competitive power dynamics in technology-based entrepreneurial and established companies. She studies these issues using both theory building from cases and simulation methods.

**Melissa E. Graebner** (melissa.graebner@mccombs.utexas.edu) is an assistant professor of management at the McCombs School of Business in the University of Texas at Austin. She received her Ph.D. in management science and engineering from Stanford University. Her research interests include corporate governance, trust, and strategic decision making. She examines these issues in the contexts of mergers and acquisitions and entrepreneurial firms.