Strategic Management Journal
Strat. Mgmt. J., **32**: 139–158 (2010)
Published online EarlyView in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/smj.868
Received 12 February 2009; *Final revision received 5 May 2010*

TESTING MANAGEMENT THEORIES: CRITICAL REALIST PHILOSOPHY AND RESEARCH METHODS

KENT D. MILLER[1]* and ERIC W. K. TSANG[2]
[1] *The Eli Broad Graduate School of Management, Michigan State University, East Lansing, Michigan, U.S.A.*
[2] *School of Management, University of Texas at Dallas, Richardson, Texas, U.S.A.*

This study identifies the practical and philosophical difficulties associated with testing strategic management and organization theories. Working from a critical realist perspective, we affirm the importance of falsification and verification efforts for progress in theory development. We advocate a four-step approach for advancing theory testing that prioritizes identifying and testing for the presence and effects of hypothesized causal mechanisms, rather than solely focusing on correlational methods to jointly test the set of effects composing a theoretical system. Going beyond prior critical realist writings, we provide practical guidance for deploying established research methods to test management theories. Copyright © 2010 John Wiley & Sons, Ltd.

INTRODUCTION

Both practical and philosophical obstacles hinder management researchers' efforts to test theories conclusively. Some challenges to testing management theories are inherent to the complex, open, and changing nature of organizations and their contexts (Astley and Van de Ven, 1983; Fabian, 2000). Other obstacles relate to researchers' beliefs and practices. For example, imprecisely stating theories obstructs empirical testing (Astley and Zammuto, 1992; Bacharach, 1989). Verifying or falsifying evidence remains disputable due to differing beliefs about what constitutes relevant evidence and how such evidence should be collected and analyzed. In response to such obstacles, management researchers have given greater attention to building theories (Lewis and Grimes, 1999; Locke, 2007; Pentland, 1999; Weick, 1989) than to testing them (Davis and Marquis, 2005; Hambrick, 2007).

Most empirical studies in strategic management use correlational methods that do not directly test the explanatory mechanisms proposed by our theories (Bromiley and Johnson, 2005; Tsang, 2006). As a result, we fail to provide direct evidence verifying or refuting our theoretical explanations and do not rule out alternative explanations. In short, researchers often fail to test theories rigorously. As empirical research grows in the management field and methods proliferate, we need to consider how our efforts contribute to cumulative progress toward theory falsification and verification. To support theory testing efforts, the current study provides both philosophical arguments and practical research methods.

We present critical realist philosophy of science as an alternative to strict falsificationism (Popper, 1959) that supports the critical evaluation of theories on the basis of empirical data. By positing a realist ontology and a fallibilist epistemology, critical realism motivates efforts

Keywords: critical realism; mechanisms; philosophy of science; research methods; theory testing
*Correspondence to: Kent D. Miller, The Eli Broad Graduate School of Management Michigan State University, East Lansing, MI 48824-1122, U.S.A. E-mail: millerk@bus.msu.edu

toward falsifying theories in order to advance scientific knowledge. However, falsifications are never certain and definitive; they may be revised (Sayer, 1992: 205–206). Critical realist philosophy acknowledges the potential fallibility of all knowledge claims, and supports modesty regarding verification and falsification alike. The aim of researchers evaluating theories is to offer 'judgements of superiority and inferiority' (Sayer, 1992: 206).

The current study begins by providing background on the challenges—both practical and philosophical—that inhibit the testing of management theories.[1] We then offer an overview of critical realism, highlighting key aspects pertaining to the evaluation of theories. However, the problem remains that prior writings come up short of translating critical realist philosophy of science into applied research methods. As Danermark *et al*. (2002: 150) succinctly summarize, 'Critical realism is not a method.' Acknowledging this claim, the latter portion of our study addresses the gap between critical realist philosophy and practical methods for theory testing within empirical management research. We translate critical realist philosophy into four actionable research steps. Our approach emphasizes identifying causal mechanisms and testing their effects directly, as steps complementary to testing theoretical systems using multivariate correlational methods. Rather than advocating novel methods, we discuss how to deploy various established research methods within a critical realist approach to theory testing.

OBSTACLES TO TESTING MANAGEMENT THEORIES

The major obstacles to testing management theories in a rigorous manner can be categorized into four domains: the nature of the social phenomena that management researchers investigate, imprecise and fragmented theorizing, inadequate research designs, and inevitable reliance upon untested assumptions. This section discusses each of these major obstacles to testing management theories.

Nature of social phenomena

In the social sciences, 'it is unrealistic to assume that all relevant data will be consistent with a theory even if the theory is correct' (Lieberson, 1992: 7). There are several factors that make conclusive theory testing more difficult in the social sciences in general, and management theory in particular, than in the natural sciences.

First, organizations are diverse, complex, and changing social phenomena, with multiple levels of analysis, as well as multiple and contingent causal processes (Astley and Van de Ven, 1983; Fabian, 2000). Owing to our inability to identify all the pertinent variables, we are often unable to state precisely the conditions on which different types of behavior depend or even the contingencies that make particular theories relevant to certain empirical contexts. Where boundary conditions are ambiguous, it may not be possible to determine whether theoretical explanations are direct competitors. From a theory testing view, the availability of competing hypotheses derived from different theories is a fortuitous situation, but it occurs infrequently in management research due to the diverse range of phenomena studied. By focusing on diverse phenomena, researchers tend to position their theories as complementary rather than competing. By contrast, the history of natural science shows that testing competing hypotheses is an effective way to determine the relative merits of alternative theories (Losee, 2005).

A second factor that renders theory testing difficult concerns the element of personal volition in human behavior. Though constrained by habits, rules, routines, and institutions, humans have freedom to choose their actions (Archer, 2000; Downward, Finch, and Ramsay, 2002; Searle, 2001). In strategic management, this belief is reflected in arguments supporting strategic choice as causal, and not merely epiphenomenal (Child, 1997). A similar conviction is found in research on creativity in entrepreneurial venturing (Sarasvathy, 2001). Although contextual factors have important influences on organizations, Child (1972) submits that organizational decision makers are not passive. They exercise choice and take actions that enact their organizations' environments. Not only are strategic behaviors nondeterministic, they can be intentionally unpredictable and deceptive in efforts to gain advantages over rivals. Freedom and

[1] By 'management theories,' we mean strategic management and organization theories.

DOI: 10.1002/smj

creativity curtail the predictive power of theories of human behavior.

Third, researchers' activities may change the beliefs and practices of managers and thereby undermine the continuity of the phenomena investigated (Numagami, 1998). Self-fulfilling and self-defeating prophecies are not causes for concern in the natural sciences. A theory of, say, planetary motion will not change how planets actually move. By contrast, managers may alter their behaviors on the basis of the knowledge created by researchers (Knights, 1992). The extent to which research influences managerial practice likely varies widely depending on the theory and organization of interest.

Moreover, researchers and managers are related to one another such that researchers themselves can, at times, be direct causal agents in organizational actions (Bradbury and Lichtenstein, 2000). For example, making explicit the implicit theory-in-use within organizations can catalyze organizational change (Argyris and Schön, 1978), thereby undermining the behavioral relevance of a previously identified theoretical explanation. Conversely, an invalid theory, if vigorously promoted to managers and given sufficient time, could become a more accurate description of their behavior (Brennan, 1994). Ferraro, Pfeffer, and Sutton (2005, 2009), for example, maintain that economic theories become self-fulfilling as their behavioral assumptions become normative in organizations. A related issue is that unlike the natural sciences, falsification of a theory in the social sciences may not eliminate it permanently because the historical and institutional conditions that originally supported the theory, though subsequently superseded or transformed, may, with further historical change, return again to restore support (Hutchison, 1988).

Theorizing

One of the primary obstacles to testing management theories is their imprecision. Bacharach (1989: 501) comments that management theories 'are often stated in such a vague way that the theorists can rebut any discrediting evidence.' Astley and Zammuto (1992: 446) note: 'while linguistic ambiguity increases the potential number of empirical tests conducted on a theory, it also reduces the chance that those tests can amount to a refutation of the theory.' Imprecise wording of hypotheses can make them logically nonfalsifiable. For example, Perrow's (1994: 216) contention that 'no matter how hard we might try, the characteristics of complexly interactive and tightly coupled systems will cause a major failure, eventually' cannot be falsified due to its open time horizon. Imprecise wording is not unique to strategic management and organization research. Kuhn (1962) and Lakatos (1970) documented scientists' tendency to state and interpret theories in ways that make them immune to falsification.

Donaldson (1995) portrays management theories as oriented toward different levels of analysis and different value assessments, as well as having distinct languages and methodologies. For each theory, there is a dedicated group of scholars working on research to test hypotheses derived from their theory. To establish their distinct niches within a research field, groups of researchers emphasize how their theories differ from one another (Mone and McKinley, 1993). Over time, barriers between groups of theorists grow (Aldrich, 1992). As researchers perceive their theories to be incommensurable and discontinue conversations with advocates of alternative theories (Mahoney, 1993), integration becomes very difficult (Jackson and Carter, 1991). For instance, in a debate with finance researchers Amihud and Lev (1999) and Denis, Denis, and Sarin (1999) concerning the relation between equity ownership structure and corporate diversification strategies, Lane, Cannella, and Lubatkin (1999) claim that agency theory and strategic management present two different worldviews. Such claims of incommensurability imply difficulty in comparing—logically and empirically—the merits of distinct theories.

Not only do we have a plurality of theories within management research, there is also no consensus about the criteria for evaluating theories. Critical realism (Fleetwood, 2005; Tsang and Kwan, 1999), positivism (Donaldson, 1996; Wicks and Freeman, 1998), constructivism (Mir and Watson, 2000), interpretivism (Lee, 1991), and pragmatism (Powell, 2002, 2003; Wicks and Freeman, 1998) are among the most prominent philosophies of science reflected in management studies.

In the natural sciences, a common view is that there can be only one true theory explaining any particular phenomenon. Thus, researchers espousing competing theories are keen to provide not only evidence that supports their theories but also evidence that challenges rival theories

DOI: 10.1002/smj

(Chalmers, 1999). By contrast, among management theorists, it is generally accepted that the same phenomenon can be explained by different theories (Allison, 1971; Ghoshal, 2005). If explanations are not mutually exclusive, greater space is opened up for researchers to create original theories that provide novel explanations. Taking the view that various theories can coexist in harmony, some researchers advocate examining organizational phenomena through multiple theoretical lenses (Nambisan, 2002; Rajagopalan and Spreitzer, 1996). For example, in their study of the differential impact of foreign institutional and foreign corporate shareholders on the performance of emerging market firms, Douma, George, and Kabir (2006) use agency, resource-based, and institutional theories, claiming that a multitheoretic approach provides a more holistic perspective. Viewed positively, this approach is compelling given the complexity of organizations and the lack of a general theory to account for the rich variety of organizational phenomena. Viewed negatively, it could evidence a failure to carefully specify and scrutinize the assumptions, explanations, and implications of different theories to determine whether they are compatible or incompatible. Whereas management researchers are not entirely averse to engaging one another in debates (Fabian, 2000), these exchanges tend to remain at a theoretical, and sometimes a philosophical, level (De Cock and Jeanes, 2006). For instance, the debate between Ghoshal and Moran (1996) and Williamson (1996) concerning how and how far organizations are able to reduce opportunism has remained at the theoretical level, stimulating few follow-up empirical studies. By avoiding clearly stating competing hypotheses and digging into the evidence, empirical claims within our theories are never at risk in such debates.

Research design

Because management theories often do not provide guidance that is precise enough to know if our models are properly and fully specified, we cannot assess the nature and extent of misspecification. For example, our theories specify causes and effects, but rarely tell us the precise timing (Mitchell and James, 2001). A theory may alert us to causality running in one direction, but neglect direct or indirect causal relations running the other way. Rarely are our tested models derived from mathematical identities, which provide some assurance against specification errors. Instead, researchers are left to their own discretion in choosing which variables to include and the functional forms of their models. Because our theories are incomplete, we augment our models with control variables to reflect supplemental explanations to the extent that our data provide relevant proxies. Exercising discretion in model specification does not present an inherent bias against falsification, but it can be an occasion for opportunistic behavior by researchers who are intent on verification. Researchers seeking to verify their theories can search for specifications that produce theory-supporting results (Caudill and Holcombe, 1987, 1999; Leamer, 1978).

Most tests of management theories are not conducted under the conditions of a closed system, which is defined as 'one in which a constant conjunction of events obtains; i.e., in which an event of type *a* is invariably accompanied by an event of type *b*' (Bhaskar, 2008: 70). Regularity of events occurs when a phenomenon is buffered from outside influences, but conditions of closure are rarely achievable in the nonexperimental social sciences. Macro-organizational phenomena are often not amenable to laboratory research, although it may be possible to extrapolate, *mutatis mutandis*, from laboratory research on individuals and groups to the organization level. Efforts to isolate social phenomena in laboratory experiments can introduce artificiality, which alters behavioral responses (Harré and Second, 1972). It is impossible to specify all the necessary initial conditions in a test situation, even in controlled experiments (Caldwell, 1984). The openness of the social world entails that a causal mechanism that is under study will not *alone* determine the course of events observed by researchers (Peacock, 2000).

We can learn more from disconfirming cases than from confirming cases, yet our reasoning and sampling tend to have a confirmatory bias (Wason and Johnson-Laird, 1972). The originators of a theory may hold it with such conviction that they—consciously or unconsciously—pay attention primarily to supporting results. Observing this more than a century ago, Chamberlin (1965 [1890]) advocated the method of multiple working hypotheses: instead of basing empirical tests on a single theory, researchers should try to develop tenable hypotheses for the phenomenon of interest

DOI: 10.1002/smj

from as many conceivable perspectives as possible. By doing so, researchers are less likely to be biased by theory when collecting observations and conducting analyses.

Researchers tend to adopt what Klayman and Ha (1987: 211) call a 'positive test strategy,' that is, examining instances in which a theory is expected to hold. 'A theory-confirming researcher perseveres by modifying procedures until prediction-supporting results are obtained' (Greenwald *et al*., 1986: 220). Journal review processes reinforce this confirmatory orientation by favoring theory-supporting results (Feige, 1975; Pfeffer, 2007). In a similar vein, Aldrich (1992: 36) laments that management research 'systematically disdains the two processes central to scientific endeavor: attempted replication of previous findings, and publication of negative findings.' A positive test strategy leads to inflated confidence in a theory's corroborating evidence and generalizability; it also discourages exploration of possible alternative explanations.

Unavoidable assumptions

In science and daily life, we operate on the basis of untested assumptions. No one can be a consistent skeptic; action requires trust in the veracity of a guiding theory (Putnam, 1974). Neither theoretical nor practical knowledge is free of presumptions (Gadamer, 2002; Rescher, 1988, 2000). For example, researchers' interest in empirical evidence presupposes that our perceptions of data give some indication of the way things are. Also, our use of language to communicate our theories and findings assumes that others will interpret our research in ways that approximate our intended meaning (Campbell, 1963). As Polanyi (1962) points out, we cannot possibly doubt all of our beliefs simultaneously; we can only doubt some beliefs on the basis of other beliefs that we hold unquestioningly. Hence, Polanyi (1962) portrays scientists as operating from a 'post-critical' or fiduciary perspective, rather than from universal skepticism. Because of our unexamined assumptions and human fallibility, and our limited access to data, our conclusions are always tentative and our generalizations are risky; nevertheless, researchers venture bold claims about the relevance of their findings in settings beyond those studied.

The Duhem-Quine thesis (Duhem, 1954; Quine, 1961) raises a logical concern that threatens to undermine the prospects for testing theories. According to this thesis, theories are never tested alone. Rather, they are tested together with other untested auxiliary hypotheses, which include background assumptions and rules of inference. A theory and its associated auxiliary hypotheses together form a *test system*. Let H be the principal hypothesis derived from a substantive theory of interest, A the set of auxiliary hypotheses, and O the observational consequence entailed by the conjunction of H and A so that $H \wedge A \Rightarrow O$.[2] An observation 'not O' implies 'not $(H \wedge A)$' rather than 'not H' alone. In other words, the whole test system, instead of the principal hypothesis, is falsified by the contradictory outcome (Grünbaum, 1960), and the existence of auxiliary hypotheses confounds tests of the principal hypothesis. When a finding does not support a theory, one can argue that the problem resides in the failure to control for certain conditions of the theory, rather than in the theory itself (Nooteboom, 1986).

Consider, for example, the above mentioned exchange about the relation between equity ownership structure and corporate diversification strategies. Finance and strategic management researchers arrive at distinct conclusions concerning the relation because they employ different test systems. Researchers in the two fields measure key constructs differently and use disparate reasoning to qualify and interpret empirical evidence (Lane *et al*., 1999). These differing auxiliary assumptions lead to the perception that the two research streams are not mutually informing, and obstruct reaching a shared conclusion based on the evidence from past research.

A critical problem posed by the Duhem-Quine thesis is that when an instance of falsification occurs, logic alone cannot isolate the elements of a test system responsible for the falsification. This problem of underdetermination is particularly serious in the social sciences where there is little agreement among researchers as to how certain auxiliary hypotheses should be tested independently (Meehl, 1978).

Testing management theories under the conditions of an open, rather than closed, system further aggravates the problems associated with isolating the relations of interest from other confounding effects. Specifying the boundary conditions of a theory is critical to advancing theory testing

[2] The logical operator $\wedge$ signifies 'and.'

DOI: 10.1002/smj

(Bacharach, 1989). For instance, the Uppsala internationalization process model proposed by Johanson and Vahlne (1977) maintains that firms exhibit a pattern of incremental commitment to foreign markets as learning occurs over time. Early statements of the model were vague about its boundary condition that the model applies primarily to overseas expansions motivated by the objective of market seeking. If other motives—such as resource seeking—are dominant, the model does not apply (Petersen and Pedersen, 1997). For years, lack of clarity about this boundary condition caused confusion about how far empirical data supported the model.

CRITICAL REALISM AND THEORY EVALUATION

For guidance on how to evaluate theories, we turn to critical realism, which is a growing intellectual movement in the social sciences (Cruickshank, 2003), especially economics (Lawson, 1997), management (Fleetwood and Ackroyd, 2004), marketing (Hunt, 1992), and sociology (Hamlin, 2002).[3] This section briefly describes some key elements of this philosophy of science, including its epistemological stance, reliance upon explanations stated in terms of mechanisms, and view that nature consists of irreducible strata.

Epistemological stance

Critical realism interrelates ontology and epistemology. On the one hand, it posits a realist ontology, that is, the existence of a world independent of researchers' knowledge of it. On the other hand, critical realism holds to a fallibilist epistemology in which researchers' knowledge of the world is socially produced. These two claims jointly motivate the need for and possibility of critically evaluating theories. As such, they make up the core of the critical realist philosophy of science. Because of the existence of an external referent, knowledge claims may be challenged and their merits assessed logically and empirically. Bhaskar (2008 : 43) notes: 'To be a fallibilist about knowledge, it is necessary to be a realist about things.'

Science is a personal and social process and nothing in the methods of science guarantees success in arriving at truth (Miller, 2005). Researchers' fallibility and unavoidable reliance upon presumptions rule out claims to indubitable knowledge, but do not preclude knowledge claims in general (Bernstein, 1983; Polanyi, 1962). Lacking an indubitable basis for science, we can, nevertheless, reasonably assert the veracity or falsity of scientific theories—albeit, not definitively. A critical realist perspective affirms the possibility of truthful knowing but acknowledges that human limitations undermine claims to indubitable or objective knowledge.

Critical realism rejects *judgmental relativism*—the inability to judge the merits of theories—as a general implication of researchers' fallibility and social situatedness (Sayer, 2000). Critical realism takes a balanced and modest stance regarding the prospects for affirming and rejecting theories based on empirical evidence. As Sayer (2000: 70) summarizes, 'Just as assessing truth or practical adequacy is not an all-or-nothing affair, neither is falsification. In practice it is messy, provisional and often partial, and the empirical observations are theory-laden—or at least conceptually mediated—though this does not make it impossible for the theory in question to be contradicted.'

Mechanisms as explanations

Bhaskar (2008) distinguishes between the domains of the real, actual, and empirical. The *real* domain consists of generative *mechanisms*, which refer to 'the ways of acting of things' (Bhaskar, 2008: 14).[4] Through enabling or preventing change, mechanisms give rise to events in the *actual* domain. Some events are experienced, and some are not. The *empirical* domain is made up of events experienced through direct or indirect observation, such

[3] It is beyond the scope of this article to compare critical realism with other philosophical perspectives, but we note a few references for the interested reader. Sayer (2000: Part II) compared critical realism and various postmodernist perspectives. Bhaskar (1998: Chapter 4), a key founder of critical realism, wrote a critique of positivism and hermeneutics. Manicas (1987: Chapter 12) critiqued positivist empiricism. Danermark *et al*. (2002: Chapter 6) argued the merits of critical realism, as a methodological guide, over positivism (which is associated with quantitative methods), as well as hermeneutics and phenomenology (which are associated with qualitative methods). In strategic management, Mir and Watson (2000, 2001), on one side, and Kwan and Tsang (2001), on the other, debated the merits of constructivism and critical realism.

[4] For further background on mechanisms see: Bunge (1997), Hedström (2005), Hedström and Swedberg (1998), Machamer, Darden, and Craver (2000), and Pajunen (2008).

DOI: 10.1002/smj

as detection through instruments. The reality of the world does not assure our access.[5] Our situation, including our sensory capacities and instruments, can render aspects of the world imperceptible. Mechanisms can be—but need not be—stable, whereas events and empirical outcomes are always contingent. Only to the extent that mechanisms are intransient can they serve as explanations across cases and time.

The implications of mechanisms need not be actualized or empirically evident. Realized events are *conjunctures* of all of the mechanisms operating in a situation (Bhaskar, 1998). Lawson offers a similar contention: 'A demi-regularity, or *demireg* for short, is precisely a partial event regularity which *prima facie* indicates the occasional, but less than universal, actualization of a mechanism or tendency, over a definite region of time-space. The patterning observed will not be strict if countervailing factors sometimes dominate or frequently co-determine the outcomes in a variable manner' (Lawson, 1997: 204, emphasis in original). This observation places great importance on establishing a theory's boundary conditions, and accounting for countervailing mechanisms in order to test a theory. For example, if an observed choice of governance structure contradicts the prediction of transaction cost economics, it is not necessarily an incident of falsification. Rather, it may indicate the presence of countervailing mechanisms, such as tax incentives or government restrictions. Therefore, testing a theory's explanatory power involves going beyond deducing a theory's implications and seeking empirical data regarding those implications; it requires checking whether the proposed theoretical mechanisms account for the data.

Bhaskar (2008) notes that the openness of social systems has both extrinsic and intrinsic sources. *Extrinsic closure* refers to the standard notion of controlling for extraneous variables. *Intrinsic closure* inhibits individuals from altering their behavioral responses to circumstances. Recognizing system openness, critical realists (Archer, 1998; Sayer, 1992) generally have dismissed conclusive falsification as unattainable within social science research. Furthermore, Bhaskar (2008: 160–162) cautions against 'pseudo-falsification,' which can occur if a theory is not sufficiently developed to deal with anomalous findings if the data are wrong, or if countervailing mechanisms cause researchers to conclude prematurely that a theory is false.

Rather than focusing solely on empirical outcomes, critical realists seek explanations for contingent relations, understood in terms of causal mechanisms. As such, critical realism rejects Hume's (2000 [1739]) conception of causality in terms of universal empirical regularities. A science concerned with potentialities rather than actualities, proposes and examines real causal mechanisms. Hence, critical realists seek to test explanations, not just correlations between observable antecedents and consequences (Carter and New, 2004). The search for explanations also goes beyond the pragmatic concern for whether acting on a theory produces the expected and desired effect.

Sayer (1992) emphasizes the priority of explanation over prediction in theory testing. The validity of predictions depends upon *ceteris paribus* conditions, which do not apply in open systems. By contrast, explanations based on causal mechanisms may function despite their effects remaining unrealized due to other countervailing mechanisms. In light of the openness of social systems, Bhaskar concludes, '. . .the rational development and replacement of theories in social science must be *explanatory and non-predictive*' (Bhaskar, 1998: 45–46. emphasis in original). Testing explanations requires examining not only the causal mechanism of interest, but also the empirical evidence regarding intervening and countervailing mechanisms. In this way, claims regarding intervening and countervailing mechanisms should not be invoked simply as a *post hoc* means to bolster a theory in the face of contrary evidence; they should be tested.

Strata within reality

Critical realism portrays reality as stratified. The mechanisms operating in each stratum are unique and, despite being made up of components (Pajunen, 2008), may exhibit properties that are not explicable in terms of mechanisms operating at other levels. It is a conceptual error to conflate distinct strata (Archer, 1995). Instead, emergent properties must be treated as dependent upon other levels, but not reducible to the mechanisms operating at those levels. Social phenomena emerge from biological and psychological

[5] The 'epistemic fallacy' associates reality only with the experiential, thereby limiting ontology to epistemology.

DOI: 10.1002/smj

strata, but explanations drawn from biology or psychology may be too distant to account for them. Following such reasoning, Nickel and Rodriguez (2002) question management researchers' extrapolations from individual-level theories, such as Kahneman and Tversky's (1979) prospect theory, to formulate hypotheses about organizational risk taking. The appropriate tests for theories involve examining mechanisms at the level of analysis at which they occur, thereby reflecting the relevant emergent properties that would be lost by examining mechanisms only at lower strata. Machamer *et al*. (2000: 13–14) refer to this level of analysis as where a mechanism 'bottoms out.' Mechanisms should not be reduced to their components if doing so strips them of essential properties. In adopting a stratified view of nature, Bhaskar (1998) rejects reductionist approaches to social science such as methodological individualist accounts of social phenomena. Critical realists reject conflation of levels and reductionism.

Stratification of the world justifies a parallel stratification of science. Indeed, a stratified view of reality simplifies our research efforts by allowing us to focus on those mechanisms germane to the phenomenon of interest, without needing to examine mechanisms associated with constituent levels. Such focus should not, however, be taken to mean that only one level of analysis is required to explain any given social phenomenon. For example, social phenomena often involve both agency, at the individual level, and structure, at the social level (Archer, 1995). This multilevel perspective is evident in research on technological change (Orlikowski, 1992), strategy (Jarzabkowski, 2008; Pozzebon, 2004), and entrepreneurship (Sarason, Dean, and Dillard, 2006) as structuration processes. Mechanisms are not always uncovered by dropping to a lower level of analysis; mechanisms can operate downward from a higher level or laterally within the same level, as well as upward from a lower level (Anderson *et al*., 2006).

CRITICAL REALIST METHODS

The issues to which critical realism alerts us tend to go unacknowledged in the design and implementation of management research. Most researchers work from an implicit *empiricist ontology*, which conflates the empirical and the real, rather than a *realist ontology* informed by Bhaskar's (2008) distinction between the real, the actual, and the empirical. Using variance designs (Mohr, 1982), researchers often study empirical regularities expressed as correlations, rather than examining causal explanations directly. Although research methods training sensitizes us to the distinction between closed and open systems, our theories and empirical tests reflect to only a limited extent the critical realist emphasis on outcomes as resulting from contingent conjunctions of mechanisms.

The methodological implications of critical realism for management researchers remain sketchy. Ackroyd (2004) offers general implications of critical realism for management research. Tsang and Kwan (1999) address the role of replicating prior research in theory development. Tsoukas (1989), Easton (2000), and Harrison and Easton (2004) propose critical realist approaches to case study research. Mingers (2004) critiques researchers' use of statistical modeling. Fairclough (2005) explains a critical realist approach to organizational discourse analysis.

Some critical realists (Wuisman, 2005; Yeung, 1997) acknowledge the gap between critical realism's contribution to philosophy of science and the dearth of methodological guidance in this literature. Those who have begun to address this gap have done so primarily for the purposes of theory *generation* and *verification*, rather than theory *falsification*. We see a need to augment critical realist writings by specifying methods for testing theories. To the extent possible, proposed methods should address the obstacles to theory testing identified earlier.

This section advances practical guidance for evaluating theories—particularly management theories—that would facilitate theory testing from a critical realist perspective. We first propose a four-step approach to theory testing, and then advocate research programs that combine both extensive and intensive designs.

Testing mechanisms

Within critical realism, to theorize is to propose mechanisms that explain events. To identify mechanisms, researchers rely upon a mode of reasoning beyond the two forms traditionally emphasized in philosophy of science—deduction and induction (Wuisman, 2005). *Retroduction* is the mode of inferential reasoning that reconstructs the conditions for the occurrence of an empirical

DOI: 10.1002/smj

phenomenon (Bhaskar, 2008; Danermark *et al*., 2002; Sayer, 1992). Retroduction seeks to identify the generative mechanisms that generalize beyond the immediate instance of the phenomenon and are critical to its occurrence. To abstract from specific empirical instances to mechanisms, we attempt to distinguish general and essential conditions that underpin the phenomenon from incidental and nonessential conditions (i.e., spurious effects).

Less emphasized in the coverage of critical realist methods is the needed complementary process of *testing* postulated mechanisms. Sayer (1992) advocates testing mechanisms directly instead of testing their observable implications. This view reflects an analytical shift from empirical correlations to real mechanisms. Whereas correlation analyses can evaluate the predicted implications of a theory, they fail to substantiate or refute a theory's causal explanation directly or to rule out alternative explanations. If mechanisms are to be truly explanatory, they must be amenable to testing directly (Bromiley and Johnson, 2005; Tsang, 2006). However, critical realist discussions identify two characteristics that present formidable obstacles to testing explanations: (1) mechanisms may not be directly observable and (2) their effects are contingent. If critical realists are to engage in both theory generating and theory testing research, they must overcome these two obstacles by providing methods to (1) identify and measure the relevant mechanisms in play in a situation and (2) test for conjunctions of mechanisms as explanations for empirical outcomes.

The concern about unobservable mechanisms has been addressed elsewhere in strategic management research. Godfrey and Hill (1995) distinguish between constructs that are *measurement unobservable* (i.e., there is no instrument appropriately calibrated) and *state unobservable* (i.e., observation perturbs the state). In the former case, the key emphasis needs to be on advances in measurement. For the latter case, effort should be directed toward techniques for unobtrusive data collection (Webb *et al*., 2000). The two forms of unobservability are not unique to critical realist research, and the methods for addressing both challenges are well established, although success in overcoming them is by no means guaranteed. Mechanisms that are unobservable from the researcher's perspective may be apparent, nevertheless, to participants within the research setting. Questioning participants (using interviews or questionnaires) can elicit evidence regarding mechanisms, although self-report measures are susceptible to various kinds of biases and must be treated with caution (Schwarz, Groves, and Schuman, 1998).

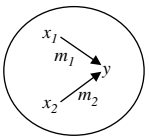
Strategic management research tends to move from theoretical arguments motivating sets of hypotheses directly to theory testing using multivariate models that incorporate proxies for multiple theoretical and control variables. Such research seeks to test full theoretical systems using variance (i.e., correlation) methods. Bolstering this approach is the concern that regression models that are less than fully specified produce biased coefficients. We support the use of multivariate correlational methods, but believe that they do not, by themselves, fulfill the exigencies of critical realism. In particular, they neglect the requirement to attend directly to testing causal mechanisms as part of the process of verifying or falsifying a theory. Examining mechanisms directly calls for added research steps that specify the hypothesized mechanisms, test for the presence of these mechanisms, and determine whether they function as hypothesized. These three steps, together with testing the full theoretical system, make up our four-step critical realist approach to theory testing.

Step 1. The initial step involves identifying the causal mechanisms believed to account for the hypothesized relations in a study. This step requires interpretive work at the interface of theory and the empirical context of interest. Researchers must resolve ambiguities in their theories and derive their implications for particular settings. The goal is a contextualized specification of the explanatory properties and processes that underlie hypothesized causal relations. Because a theory may propose different mechanisms to explain different phenomena, researchers need to select those mechanisms that they believe operate in their particular research setting. More than one mechanism may be relevant to a given causal relation; furthermore, mechanisms may have complementary or conflicting implications for a hypothesized relation. To illustrate the importance of identifying all relevant mechanisms, Bromiley and Johnson (2005) point out the varied and conflicting ways that top management team diversity affects firm performance.

To detail the steps in our critical realist approach to testing theories, consider a simple theory contending that variables x_1 and x_2 jointly determine

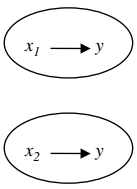
DOI: 10.1002/smj

Step 1: Identify the hypothesized mechanisms.



Step 2: Test for the presence of the mechanisms (m_1 and m_2) in the empirical setting.

Step 3: Test isolated causal relations.



Step 4: Test the theoretical system.

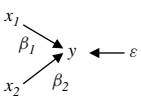


Figure 1. Steps in a critical realist approach to testing theories

y. The top portion of Figure 1 shows the path diagram representing this theoretical system. The hypothesized mechanisms, m_1 and m_2 link x_1 and x_2, respectively, to y. We maintain the important distinction between the parameters relating x_1 and x_2 to y (designated β_1 and β_2 in step 4), and the mechanisms themselves, m_1 and m_2. The circle encompassing the path diagram reflects the explicit or implicit *ceteris paribus* condition, which treats the theoretical system as closed. Specifying a model according to step 1 of Figure 1 challenges theorists to clarify theoretical mechanisms (Anderson *et al*., 2006) and boundary conditions (Bacharach, 1989).

Step 2. The next step in testing this theory is to consider whether the proposed mechanisms are indeed present in the empirical setting. Step 2 of Figure 1 conveys the emphasis on substantiating the existence of the causal mechanisms. Failure to validate the presence or the nature of postulated mechanisms provides compelling evidence to reject a theory's arguments. Prioritizing testing for mechanisms, rather than jumping directly to testing their hypothesized relations to dependent variables, runs contrary to the pattern in strategic management research (Bromiley and Johnson, 2005).

Tests for the presence of a mechanism can involve looking for collateral implications of the mechanism beyond those predicted by the theory of interest (Goldthorpe, 2001). Even if a particular mechanism is not directly observable, the more observable effects that logically are attributable to the mechanism, the more compelling is the case for its presence. For example, consider an argument claiming that experience with an alliance partner leads to greater willingness to make asset-specific investments, with growth in trust as the explanation. Trust between alliance partners may not be directly observable, but it should produce a variety of manifest effects on negotiations, contracts, and coordination processes. In this case, multiple observable indicators provide indirect support for the presence of an unobserved mechanism. The reasoning here parallels the use of multiple indicators for latent constructs in structural equation modeling but, in this case, the construct of interest is a causal mechanism.

Bhaskar (1998) maintains that people's reasons can operate as causes, which in turn are responsible for producing or inhibiting change. Such intentional human behavior provides an opportunity for testing the mechanism of a theory. The retroduction of intentions relies upon either truthful disclosure by the individuals involved, generalizations about similarly-situated others, or projections of the researcher's own intentions onto the studied subjects. In order to infer reasons and evaluate the subjects' espoused reasons, we draw heuristically upon our own self-understanding for insights into others' unobservable intentions. Our common humanity and experiences, rather than a posture of objectivity, inform the hermeneutical act of inferring intentions from others' actions (Gadamer, 2002; Ricoeur, 1981).

Bromiley and Johnson (2005) and Tsang (2006) contend that behavioral assumptions are fundamental to the explanations of management theories, yet they are often overlooked in empirical evaluations. As such, it remains unclear whether the mechanisms really exist and account for the phenomena as purported. For example, transaction cost economics assumes that opportunism is a key factor affecting governance costs, and economizing on transaction costs in turn determines the governance choice (Wathne and Heide, 2000). This explanation of managerial choice relies on an implicit view that perceived, rather than objective, transaction costs are what managers

DOI: 10.1002/smj

take into account when making their decisions (Chiles and McMackin, 1996). To test the transaction cost economizing assumption, researchers need to examine how managers actually perceive and evaluate transaction costs. In other words, investigating managers' perceptions and intentions would provide a direct test of the alleged mechanism. However, in a survey of the research, Tsang (2006) found that researchers rarely subject the theory's core assumptions to empirical tests. This assumption-omitting way of testing theories fails to provide direct tests of their explanations. Two studies by Buckley and Chapman (1997, 1998) asked managers directly about the internalization and externalization of corporate activities and found no evidence of transaction cost economizing, indicating the absence of the hypothesized mechanism.

Nevertheless, does Buckley and Chapman's (1997, 1998) finding really matter as long as governance structures observed in the empirical domain are consistent with the assumption of transaction cost economizing? Yes, it matters if we care to understand the process giving rise to the outcome. There is more than one reason for observed governance structures to be consistent with expectations based on transaction cost economizing: both natural selection and managerial choice are plausible causal mechanisms. An explanation based on natural selection 'adopts the population of organizations as the level of analysis, the environment as the primary selection mechanism that utilizes some selection criteria (e.g., transaction cost economizing), the long run as the appropriate time frame, and an ex post objective view of costs' (Chiles and McMackin, 1996: 76). In this framing, firms engage in a random series of configurational changes, some of which are by accident transaction cost reducing while others are not. Firms that happen to arrive at low transaction cost configurations will succeed relative to those that do not (Buckley and Chapman, 1997). The surviving governance structures are generally consistent with the logic of transaction cost economics regardless of whether managers' choices were based on transaction cost economizing (Robins, 1987). The issue here is that the natural selection approach invokes a mechanism very different from that of the managerial choice approach (Hodgson, 1993). As such, identifying the mechanisms operating in the empirical setting is an effective way of assessing competing theories of the same phenomenon.

Step 3. If the available evidence affirms the presence of the theorized mechanisms, we move to testing their causal effects. Prior to attempting to verify or falsify an entire theoretical system in an open context, we advocate testing binary or more complex subsets of relations under controlled circumstances. For the hypothetical example shown in Figure 1, we seek tests of the component x_1-y and x_2-y relations in contexts that isolate each relation from other effects. Step 3 of Figure 1 depicts such tests. To do so, we must move from the open system of the empirical context of ultimate interest to experimental or quasi-experimental settings.

In spite of the threats to external validity (Guala and Mittone, 2005), *laboratory experiments* allow for testing the effects of mechanisms under conditions similar to those of a closed system and, due to random assignment of treatments, generally provide stronger evidence that supports or rejects hypotheses than do nonexperimental designs. Many theories invoke mechanisms that consist of chains of causal links in which one event leads to another; experiments help to isolate and test these links. Moreover, experiments enable researchers to assess the merits of competing theories by, for example, testing competing mediating relations proposed by those theories (Shadish, Cook, and Campbell, 2002). Although experiments have been a popular research method in psychology and organizational behavior, they are seldom used in strategic management research to facilitate tests of mechanisms (Croson, Anand, and Agarwal, 2007). A notable exception is the experiment conducted by Sutcliffe and Zaheer (1998) to study the effects of different kinds of uncertainty on the likelihood of vertical integration. Another is Schweiger, Sandberg, and Rechner's (1989) experiment involving managers in dialectical inquiry, devil's advocacy, and consensus approaches to group strategic decision making.

A *behavioral simulation* involves individuals in a setting constructed to mimic key aspects of a naturally occurring organizational situation. The realism of the constructed setting is the key feature distinguishing behavioral simulations from laboratory experiments. In an interactive behavioral simulation, researchers control the simulated context and observe the behaviors and decisions coming from participants' interactions with one another. The method is particularly suited for strategy process research (Dutton and Stumpf, 1991). For

DOI: 10.1002/smj

example, Hough and White (2003) used an interactive behavioral simulation to study the moderating effect of environmental dynamism on the relation of decision making comprehensiveness to decision quality. Gist, Hopper, and Daniels (1998) report that interactive behavioral simulations remain quite rare in management research—even among studies published in organizational behavior journals.

Quasi-experiments share some characteristics of laboratory experiments, but they occur in field settings and researchers have less than full control over the assignment and scheduling of effects (Shadish *et al*., 2002). Thus, quasi-experiments have some merit by offering conditions that approximate randomly assigned treatments. In contrast with laboratory experiments, quasi-experiments trade off some control for external validity (Grant and Wall, 2009). Again, with few exceptions, management researchers outside the domain of organizational behavior rarely use quasi-experimental designs. An example of such exceptions is Szulanski and Jensen's (2004) investigation of the role of templates in the replication of organizational routines, which was based on an eight-year longitudinal case study that took the form of a naturally occurring, repeated-treatment quasi-experiment.

We recommend that strategic management and organization researchers further explore the potential of laboratory experiments, behavioral simulations, and quasi-experiments for providing data to rigorously test the relations invoked in their theories. Because such research designs fall outside the training typically acquired in strategic management doctoral programs, we see opportunities for collaborating with scholars in fields such as organizational behavior and psychology where such designs are in the mainstream.

Step 4. If empirical data corroborate a theory's mechanisms and their effects, then the next step to take in evaluating the theory is to examine the implications of its mechanisms jointly. This step moves the analysis from isolated mechanisms to the entire theoretical system, thereby adding complexity to the evaluative procedure. At this stage, we are interested in whether all of the theory's mechanisms are necessary and whether they are jointly sufficient to explain the outcome. For a complicated management theory, it may not be feasible to test all of its mechanisms in a single study. In such cases, a second-best alternative is to design tests of subsets of mechanisms that are unbiased by omitted theoretical variables through choosing research settings where the unmeasured variables can be assumed to be either invariant or irrelevant.

The openness of social systems complicates testing jointly a theory's hypotheses. Although the theoretical system is closed (see step 1), the empirical contexts of organizations and industries are not. Step 4 of Figure 1 depicts the open theoretical system in which outside influences give rise to unexplained variance that is relegated to the error term, ε. Here, the error term results from omitting variables relevant to the empirical context but outside the scope of the theory itself. Failure to account for some of the relevant mechanisms diminishes the proportion of variance explained by a model and potentially biases the estimated effects of the theoretical variables. Step 4, in contrast with the original theoretical model (step 1), focuses on estimable partial correlations (β_1 and β_2) and omits the generative mechanisms (m_1 and m_2). This switch from mechanisms to correlations accommodates established multivariate methods. We acknowledge that fitting empirical data to a multivariate model treats the process generating the dependent variable as a black box and, as such, correlational analysis can only complement other research that directly tests mechanisms, not substitute for such tests.

Acknowledging potentially relevant outside contingencies, in large sample studies researchers attempt to include control variables to account for effects beyond those in their theory and interactions to reflect conjunctions of mechanisms. However, critical realist reasoning can lead to an 'interactionist regress' (Sayer, 1992: 189) in which explaining outcomes as contingent conjunctions of mechanisms results in complex interactions that are difficult to understand conceptually and test empirically. Large-sample analytical methods require simplifying generalizations regarding the contingencies affecting observations, whereas case research can identify and take into account idiosyncratic contingencies (Nash, 1999). We discuss next how different research designs contribute to theory testing.

Research designs

Critical realism is highly pluralist in terms of empirical research methods. Because different methods focus on different aspects of reality,

DOI: 10.1002/smj

combining several methods together in a research program yields a rich understanding of the phenomenon of interest (Mingers, 2001). Critical realists distinguish between *extensive* and *intensive designs*, and maintain that both approaches are needed and complementary (Danermark *et al*., 2002; Sayer, 1992, 2000). Extensive research tests empirical generalities across cases and typically employs quantitative methods. Intensive research tries to uncover the explanatory mechanisms relevant in particular cases, and is often associated with qualitative modes of analysis. *Mixed designs* combine both extensive and intensive designs, thereby utilizing their complementary strengths and weaknesses. Critical realism encourages mixed designs in efforts to cope with the complexity of studied phenomena, aspects of which may go undetected by any single research approach (Mingers, 2006).

Extensive designs. Critical realist concerns about relying upon extensive designs revolve around (1) their focus on empirical correlations rather than specifying explanatory mechanisms, (2) assumptions of model stability and closure, and (3) reliance upon statistical tests of significance (Mingers, 2004). Conveying similar concerns, Sayer observes: 'Given the disjunction between mechanisms and events, a strong correlation (or some other quantitative association) need not imply causation, nor a weak one absence of a causal or structural relation. If a theory is to help solve this type of problem it must postulate causal mechanisms and not merely specify how total variation in the dependent variable might relate quantitatively to variation in the independent variables' (Sayer, 1992: 194). Even Starbuck's (2006) prioritization of effect size over statistical significance omits the goal of explanation through identifying mechanisms and their surrounding contingencies. Estimated effect sizes capture causal importance only if models are specified in terms of contingent relations reflecting complex conjunctures of mechanisms.

However, Ron (2002) offers some counterarguments to address such concerns. He points out that open systems do not *necessarily* undermine empirical regularities and the occurrence of such regularities can provide insights into the operative mechanisms. Working from a critical realist perspective, a researcher uses regression analysis to demonstrate the effects of theorized causal mechanisms after controlling for other mechanisms that also could affect the outcome. Ron submits that researchers' repeated respecification of regression models exhibits an exploratory search for mechanisms affecting an empirical outcome. Such exploratory use of regression analysis seeks to explain empirical patterns retrospectively, and, for open systems the result should not be interpreted as a predictive model, nor do conventional criteria for evaluating the statistical significance of coefficients for theory-determined models apply. If our intent is to falsify theories, researchers ought to pursue model respecifications that *undermine* hypothesized relations, not just *post hoc* respecifications that favor the theory (Caudill and Holcombe, 1987, 1999; Kennedy, 2005; Leamer, 1978) or adjustments in theoretical arguments to fit findings (Lipton, 1991).

Manicas (1987, 2006) raises the concern that the causes postulated in critical realist research are not linearly additive and, thus, violate a key assumption of regression modeling. Such a critique misses some possibilities for using multivariate modeling in critical realist research. For example, researchers can use regression models with interaction terms to test contingent effects. If a hypothesized effect varies over the range of values for a regressor, then a spline function can be used. Using regression analysis, relations for which empirical tests are feasible consist primarily of direct effects and lower-level (two-way or three-way) curvilinear effects, rather than more complex contingent relations. Hierarchical linear modeling allows coefficients to vary contingent upon moderating and mediating effects at other levels of analysis (Hofmann, Griffin, and Gavin, 2000; Zhang, Zyphur, and Preacher, 2009) and, as such, is another tool suitable to critical realists' interest in testing the effects of intervening mechanisms on theoretical relations.

For models with complex and dynamic conjunctions, computer simulation modeling offers a means to transition from theoretical arguments to empirical testing. Experimentation with a simulation model generates data for a *response surface* reflecting the effects of multiple variables on a particular outcome of interest. Such data permit estimation of an algebraic function, often called a *meta-model*, which captures the effects of a set of independent variables on the dependent variable. *Response surface methods* estimate a multivariate equation that approximates the functional relation

DOI: 10.1002/smj

of explanatory variables to an outcome variable as generated by a complex simulation model (Law and Kelton, 2000). Provided relevant empirical data are available, an estimated meta-model can be tested. This combination of simulation modeling of complex systems resulting in an empirically testable regression model provides a way to move from complex (nonlinear) dynamic systems to empirical tests using established methods. Such methods can advance empirical testing in research areas such as organizational structure and learning, where simulation research has enhanced our understanding of the relevant mechanisms (Carley, 1992; Lin and Hui, 1999; Rahmandad, 2008).

Intensive designs. Intensive designs complement extensive designs by addressing the differences across cases that one would expect if empirical outcomes result from conjunctions of multiple mechanisms in open systems. The purpose behind intensive designs is to identify and describe the generative mechanisms operating in particular cases, which is often not feasible for extensive designs. An intensive design emphasizes the collection of detailed data within one or more cases. The data are often qualitative, as are the analytic methods. Established methods for case study, ethnographic, grounded theory, and action research guide intensive designs.

A distinctive feature of critical realist deployment of these methods is the emphasis on retroduction. Intensive investigation of a case affords a unique opportunity to evaluate whether the mechanisms proposed by a theory jointly affect empirical outcomes as expected, while taking into consideration intervening and countervailing contextual mechanisms. Easton (2000: 212) concludes: 'Case research which would wish to lay claim to a realist philosophy should be carried out in a different way: to be inquisitive, to look for the roots of things, to disentangle complexities and to conceptualise and re-conceptualise, test and retest, to be both rigorous and creative and above all to seek for the underlying reality through the thick veil which hides it.' For instance, Porter's (1993) ethnographic study of how racism affected the occupational relationships between nurses and doctors in a hospital adopted a critical realist perspective for identifying mechanisms that generated the racist events and found that the universalist-achievement ethos of professionalism was a mechanism countering the ascriptive nature of racist tendencies.

Intentional sampling is an important step if intensive designs are to yield evidence verifying or falsifying theoretical claims. Contrasting cases can provide evidence from natural experiments regarding how mechanisms operate under different conditions. Intertemporal comparisons through repeated studies of the same case can reveal how continuities and discontinuities in mechanisms and contexts affect outcomes (Harrison and Easton, 2004). Pathological or extreme cases can reveal conditions where mechanisms are undermined or unimpeded (Danermark *et al.*, 2002). As Collier (1994: 165) explains, '[B]y seeing how something goes wrong we find out more about the conditions of its working properly than we ever would by observing it working properly.'

Choosing cases that are very different from one another enables tests that allow researchers to judge the relative merits of competing theories (Stinchcombe, 1968). This approach involves finding theories within a particular research area that make arguments that lead to conflicting empirical predictions. Setting up direct empirical tests of conflicting theoretical predictions provides more compelling evidence regarding their relative merits than does testing the implications of a single theory and then invoking the alternative as a *post hoc* explanation for unanticipated findings. Sagan's (1993) case study of the Cuban missile crisis is an exemplar of testing competing mechanisms and implications proposed by the two major theories that explain accidents in hazardous high-technology organizations—normal accident theory (Perrow, 1984) and high reliability theory (La Porte and Consolini, 1991; Roberts, 1989).

Auxiliary hypotheses. The aforementioned Duhem-Quine thesis informs how we test mechanisms and theoretical systems. Because the presence of auxiliary hypotheses (A) confounds the testing of any theoretical hypothesis (H), it is possible to accept a given theoretical hypothesis under one set of auxiliary hypotheses and reject the hypothesis under another set of auxiliary hypotheses. Although the confounding effects implied by the Duhem-Quine thesis cannot be eliminated, Sawyer, Beed, and Sankey (1997) suggest two ways of assessing the effects of auxiliary hypotheses. One method is to conduct sensitivity analyses where the auxiliary hypotheses are altered. For

DOI: 10.1002/smj

example, researchers can draw different samples, use alternative measures, vary model specifications, or employ dissimilar statistical methods for fitting and evaluating models. Consistent falsifying or supporting evidence using alternative auxiliary hypotheses is more conclusive than if alternative auxiliary hypotheses produce conflicting findings.

A second method is to test the auxiliary hypotheses directly. Mäki (2000) distinguishes *negligibility assumptions* (that effects are small enough to neglect), *joint negligibility assumptions* (that a set of effects is collectively small enough to neglect), *applicability assumptions* (identifying contingencies that make effects non-negligible), *early-step assumptions* (restrictions that are later relaxed), and *as-if assumptions* (involving counterfactual claims). At the broadest level, Mäki (2000) distinguishes between the core and peripheral assumptions of a theory: core assumptions indicate the major causes postulated by a theory while peripheral assumptions refer to the minor causes. Although tests of theories center on core assumptions, direct tests of peripheral assumptions can provide further evidence qualifying theories.

Replications that extend prior studies can help researchers examine the implications of alternative auxiliary hypotheses by focusing their attention on one part of the test system at a time. Suppose the findings of a study did not support the test system, $H \wedge A$. Suspecting that a certain auxiliary hypothesis caused the result, a researcher replicates the study with an alternative auxiliary hypothesis (A') while keeping the rest of the test system intact. Finding support for the new test system, $H \wedge A'$, confirms the researcher's conjecture.[6] If the system is still not supported, the researcher needs to consider whether H should be rejected. In a similar vein, Søberg (2005) submits that a series of experiments can be used to locate the source of disconfirming evidence.

CONCLUSION

To successfully test theories, management researchers must overcome a formidable set of practical and philosophical obstacles, including the complexity and contingency of social phenomena, imprecisely specified theories, the openness of social systems, and the unavoidability of untested assumptions. Critical realist arguments regarding the conditions that make science possible also are relevant to making theory testing possible in the face of such challenges (Bhaskar, 2008; Van de Ven, 2007). Critical realism confronts the complexity of social phenomena by espousing explanations stated in terms of mechanisms that generalize, with empirical effects that are contingent. Critical realist philosophy acknowledges that theories are social products, but ontological realism provides some hope of achieving greater precision over time through testing our theories. To deal with the openness of social phenomena, management researchers can employ experimental and quasi-experimental designs to complement nonexperimental designs. Critical realism also supports efforts to replicate prior research and directly test auxiliary hypotheses (Tsang and Kwan, 1999).

We sought to make critical realist philosophy of science practical by addressing its implications for testing management theories. We proposed a series of four complementary steps for testing strategic management and organization theories: (1) identify the hypothesized mechanisms, (2) test for the presence of the mechanisms in the empirical setting, (3) test isolated causal relations using experimental or quasi-experimental designs, and (4) test the theoretical system using correlational methods. Though there have been scattered discussions of these steps and the methods that they entail in the literature, we know of no other attempt to organize them into an integrative research process that facilitates theory testing from a critical realist perspective.

Most empirical studies in strategic management, especially those based on extensive research designs, seek to identify empirical regularities expressed as correlations, and thus are located at step 4 of our recommended four-step approach. Although many research studies also touch upon step 1 during the development of hypotheses, steps 2 and 3 have been neglected for the most part. Without examining causal mechanisms directly, it is difficult to adjudicate conflicting explanations for empirical findings. Even in the face of cumulative, mostly confirming empirical findings, omitting tests for the presence and causal effects of mechanisms leaves theoretical explanations in

[6] An assumption here is that the study has to be repeated in order to examine the auxiliary hypothesis in question; that is, a simple reanalysis of the prior data will not do. This is the case when, for example, the auxiliary hypothesis is about the measurement of a certain variable in a questionnaire survey. In such a case, a new survey with a different set of measurement items for that variable is needed.

DOI: 10.1002/smj

doubt. Both intensive and extensive designs can provide such tests.

Identifying and testing causal mechanisms calls for increased use of designs and methods found most frequently elsewhere in the social sciences—such as generating data through laboratory experiments, behavioral simulations, quasi-experiments, case studies, ethnography, or grounded theory building in field settings. Cross-disciplinary collaborations and drawing guidance from the methods literature in other academic fields can facilitate such research designs. Because each of the four steps requires specific research techniques and no researcher is well versed in all these techniques, there is a need for specialization, with different researchers focusing on different steps, and the potential for collaboration as researchers coordinate sequential steps or work together on more than one step.

ACKNOWLEDGEMENTS

We thank Gregory Dess, Joseph Mahoney, Donald Palmer, Arkadiy Sakhartov, Roy Suddaby, and the anonymous reviewers for their helpful comments on earlier versions of this article.

REFERENCES

Ackroyd S. 2004. Methodology for management and organisation studies: some implications of critical realism. In *Critical Realist Applications in Organisation and Management Studies*, Fleetwood S, Ackroyd S (eds). Routledge: London, U.K.; 137–163.

Aldrich HE. 1992. Incommensurable paradigms? Vital signs from three perspectives. In *Rethinking Organization: New Directions in Organization Theory and Analysis*, Reed M, Hughes M (eds). Sage: London, U.K.; 17–45.

Allison GT. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis*. Little, Brown and Company: Boston, MA.

Amihud Y, Lev B. 1999. Does corporate ownership structure affect its strategy towards diversification? *Strategic Management Journal* **20**(11): 1063–1069.

Anderson PJJ, Blatt R, Christianson MK, Grant AM, Marquis C, Neuman EJ, Sonenshein S, Sutcliffe KM. 2006. Understanding mechanisms in organizational research: reflections from a collective journey. *Journal of Management Inquiry* **15**: 102–113.

Archer MS. 1995. *Realist Social Theory: The Morphogenetic Approach*. Cambridge University Press: Cambridge, U.K.

Archer M. 1998. Introduction: realism in the social sciences. In *Critical Realism: Essential Readings*, Archer M, Bhaskar R, Collier A, Lawson T, Norrie A (eds). Routledge: London, U.K.; 189–205.

Archer MS. 2000. *Being Human: The Problem of Agency*. Cambridge University Press: Cambridge, U.K.

Argyris C, Schön DA. 1978. *Organizational Learning: A Theory of Action Perspective*. Addison-Wesley: Reading, MA.

Astley WG, Van de Ven AH. 1983. Central perspectives and debates in organization theory. *Administrative Science Quarterly* **28**: 245–273.

Astley WG, Zammuto RF. 1992. Organization science, managers, and language games. *Organization Science* **3**: 443–460.

Bacharach SB. 1989. Organizational theories: some criteria for evaluation. *Academy of Management Review* **14**: 496–513.

Bernstein RJ. 1983. *Beyond Objectivism and Relativism: Science, Hermeneutics, and Praxis.* University of Pennsylvania Press: Philadelphia, PA.

Bhaskar R. 1998. *The Possibility of Naturalism: A Philosophical Critique of the Contemporary Human Sciences* (3rd edn). Routledge: London, U.K.

Bhaskar R. 2008. *A Realist Theory of Science* (2nd edn). Verso: London, U.K.

Bradbury H, Lichtenstein BMB. 2000. Relationality in organizational research: exploring *the space between*. *Organization Science* **11**: 551–564.

Brennan MJ. 1994. Incentives, rationality, and society. *Journal of Applied Corporate Finance* **7**(2): 31–39.

Bromiley P, Johnson S. 2005. Mechanisms and empirical research. In *Research Methodology in Strategy and Management* (Vol. 2), Ketchen DJ Jr, Bergh DD (eds). Elsevier Science: Amsterdam, The Netherlands; 15–29.

Buckley PJ, Chapman M. 1997. The perception and measurement of transaction costs. *Cambridge Journal of Economics* **21**: 127–145.

Buckley PJ, Chapman M. 1998. The management of cooperative strategies in R&D and innovation programmes. *International Journal of the Economics of Business* **5**: 369–381.

Bunge M. 1997. Mechanism and explanation. *Philosophy of the Social Sciences* **27**: 410–465.

Caldwell BJ. 1984. Some problems with falsificationism in economics. *Philosophy of the Social Sciences* **14**: 489–495.

Campbell K. 1963. One form of scepticism about induction. *Analysis* **23**(4): 80–83.

Carley K. 1992. Organizational learning and personnel turnover. *Organization Science* **3**: 20–46.

Carter B, New C. 2004. Introduction: realist social theory and empirical research. In *Making Realism Work: Realist Social Theory and Empirical Research*, Carter B, New C (eds). Routledge: London, U.K.; 1–20.

Caudill SB, Holcombe RG. 1987. Coefficient bias due to specification search in econometric models. *Atlantic Economic Journal* **15**(3): 30–34.

Caudill SB, Holcombe RG. 1999. Specification search and levels of significance in econometric models. *Eastern Economic Journal* **25**: 289–300.

Chalmers AF. 1999. *What Is This Thing Called Science?* (3rd edn). Open University Press: Maidenhead, U.K.

Chamberlin TC. 1965 [1890]. The method of multiple working hypotheses. *Science* **148**(3671): 754–759.

Child J. 1972. Organizational structure, environment and performance: the role of strategic choice. *Sociology* **6**: 1–22.

Child J. 1997. Strategic choice in the analysis of action, structure, organizations and environment: retrospect and prospect. *Organization Studies* **18**: 43–76.

Chiles TH, McMackin JF. 1996. Integrating variable risk preferences, trust, and transaction cost economics. *Academy of Management Review* **21**: 73–99.

Collier A. 1994. *Critical Realism: An Introduction to Roy Bhaskar's Philosophy*. Verso: London, U.K.

Croson R, Anand J, Agarwal R. 2007. Using experiments in corporate strategy research. *European Management Review* **4**: 173–181.

Cruickshank J (ed). 2003. *Critical Realism: The Difference It Makes.* Routledge: London, U.K.

Danermark B, Ekström M, Jakobsen L, Karlsson JC. 2002. *Explaining Society: Critical Realism in the Social Sciences*. Routledge: London, U.K.

Davis GF, Marquis C. 2005. Prospects for organization theory in the early twenty-first century: institutional fields and mechanisms. *Organization Science* **16**: 332–343.

De Cock C, Jeanes EL. 2006. Questioning consensus, cultivating conflict. *Journal of Management Inquiry* **15**: 18–30.

Denis DJ, Denis DK, Sarin A. 1999. Agency theory and the influence of equity ownership structure on corporate diversification strategies. *Strategic Management Journal* **20**(11): 1071–1076.

Donaldson L. 1995. *American Anti-management Theories of Organization: A Critique of Paradigm Proliferation*. Cambridge University Press: Cambridge, U.K.

Donaldson L. 1996. *For Positivist Organization Theory: Proving the Hard Core.* Sage: London, U.K.

Douma S, George R, Kabir R. 2006. Foreign and domestic ownership, business groups, and firm performance: evidence from a large emerging market. *Strategic Management Journal* **27**(7): 637–657.

Downward P, Finch JH, Ramsay J. 2002. Critical realism, empirical methods and inference: a critical discussion. *Cambridge Journal of Economics* **26**: 481–500.

Duhem P. 1954. *The Aim and Structure of Physical Theory.* Princeton University Press: Princeton, NJ.

Dutton JE, Stumpf SA. 1991. Using behavioral simulations to study strategic processes. *Simulation and Games* **18**: 207–224.

Easton G. 2000. Case research as a method for industrial networks: a realist apologia. In *Realist Perspectives on Management and Organisations*, Ackroyd S, Fleetwood S (eds). Routledge: London, U.K.; 205–219.

Fabian FH. 2000. Keeping the tension: pressures to keep the controversy in the management discipline. *Academy of Management Review* **25**: 350–371.

Fairclough N. 2005. Discourse analysis in organization studies. *Organization Studies* **26**: 915–939.

Feige EL. 1975. The consequences of journal editorial policies and a suggestion for revision. *Journal of Political Economy* **83**: 1291–1296.

Ferraro F, Pfeffer J, Sutton RI. 2005. Economics language and assumptions: how theories can become self-fulfilling. *Academy of Management Review* **30**: 8–24.

Ferraro F, Pfeffer J, Sutton RI. 2009. How and why theories matter: a comment on Felin and Foss (2009). *Organization Science* **20**: 669–675.

Fleetwood S. 2005. Ontology in organization and management studies: a critical realist perspective. *Organization* **12**: 197–222.

Fleetwood S, Ackroyd S (eds). 2004. *Critical Realist Applications in Organisation and Management Studies.* Routledge: London, U.K.

Gadamer H. 2002. *Truth and Method* (2nd revised edn). Continuum: New York.

Ghoshal S. 2005. Bad management theories are destroying good management practices. *Academy of Management Learning and Education* **4**: 75–91.

Ghoshal S, Moran P. 1996. Bad for practice: a critique of the transaction cost theory. *Academy of Management Review* **21**: 13–47.

Gist ME, Hopper H, Daniels D. 1998. Behavioral simulation: application and potential in management research. *Organizational Research Methods* **1**: 251–295.

Godfrey PC, Hill CWL. 1995. The problem of unobservables in strategic management research. *Strategic Management Journal* **16**(7): 519–533.

Goldthorpe JH. 2001. Causation, statistics, and sociology. *European Sociological Review* **17**: 1–20.

Grant AM, Wall TD. 2009. The neglected science and art of quasi-experimentation: why-to, when-to, and how-to advice for organizational researchers. *Organizational Research Methods* **12**: 653–686.

Greenwald AG, Pratkanis AR, Leippe MR, Baumgardner MH. 1986. Under what conditions does theory obstruct research progress? *Psychological Review* **93**: 216–229.

Grünbaum A. 1960. The Duhemian argument. *Philosophy of Science* **27**: 75–87.

Guala F, Mittone L. 2005. Experiments in economics: external validity and the robustness of phenomena. *Journal of Economic Methodology* **12**: 495–515.

Hambrick DC. 2007. The field of management's devotion to theory: too much of a good thing? *Academy of Management Journal* **50**: 1346–1352.

Hamlin CL. 2002. *Beyond Relativism: Raymond Boudon, Cognitive Rationality and Critical Realism.* Routlege: London, U.K.

Harré R, Second PF. 1972. *The Explanation of Social Behavior.* Basil Blackwell: Oxford, U.K.

Harrison D, Easton G. 2004. Temporally embedded case comparison in industrial marketing research. In *Critical Realist Applications in Organisation and Management Studies*, Fleetwood S, Ackroyd S (eds). Routledge: London, U.K.; 194–210.

Hedström P. 2005. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge University Press: Cambridge, U.K.

DOI: 10.1002/smj

Hedström P, Swedberg R (eds). 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge University Press: Cambridge, U.K.

Hodgson GM. 1993. *Economics and Evolution: Bringing Life Back into Economics.* University of Michigan Press: Ann Arbor, MI.

Hofmann DA, Griffin MA, Gavin MB. 2000. The application of hierarchical linear modeling to organizational research. In *Multilevel Theory, Research, and Methods in Organizations: Foundations, Extensions, and New Directions*, Klein KJ, Kozlowski SWJ (eds). Jossey-Bass: San Francisco, CA; 467–511.

Hough JR, White MA. 2003. Environmental dynamism and strategic decision-making rationality: an examination at the decision level. *Strategic Management Journal* **24**(5): 481–489.

Hume D. 2000 [1739]. *A Treatise of Human Nature.* Oxford University Press: Oxford, U.K.

Hunt SD. 1992. For reason and realism in marketing. *Journal of Marketing* **56**(2): 89–102.

Hutchison TW. 1988. The case for falsification. In *The Popperian Legacy in Economics*, de Marchi N (ed). Cambridge University Press: Cambridge, U.K.; 169–181.

Jackson N, Carter P. 1991. In defense of paradigm incommensurability. *Organization Studies* **12**: 109–127.

Jarzabkowski P. 2008. Shaping strategy as a structuration process. *Academy of Management Journal* **51**: 621–650.

Johanson J, Vahlne J-E. 1977. The internationalization process of the firm—a model of knowledge development and increasing foreign market commitments. *Journal of International Business Studies* **8**: 23–32.

Kahneman D, Tversky A. 1979. Prospect theory: an analysis of decision under risk. *Econometrica* **47**: 263–291.

Kennedy PE. 2005. Oh no! I got the wrong sign! What should I do? *Journal of Economic Education* **36**: 77–92.

Klayman J, Ha YW. 1987. Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review* **94**: 211–228.

Knights D. 1992. Changing space: the disruptive impact of a new epistemological location for the study of management. *Academy of Management Review* **17**: 514–536.

Kuhn TS. 1962. *The Structure of Scientific Revolutions*. University of Chicago Press: Chicago, IL.

Kwan K-M, Tsang EWK. 2001. Realism and constructivism in strategy research: a critical realist response to Mir and Watson. *Strategic Management Journal* **22**(12): 1163–1168.

Lakatos I. 1970. Falsification and the methodology of scientific research programmes. In *Criticism and the Growth of Knowledge*, Lakatos I, Musgrave A (eds). Cambridge University Press: London, U.K.; 91–196.

Lane PJ, Cannella AA, Lubatkin MH. 1999. Ownership structure and corporate strategy: one question viewed from two different worlds. *Strategic Management Journal* **20**(11): 1077–1086.

La Porte TR, Consolini PM. 1991. Working in practice but not in theory: theoretical challenges of 'high reliability organizations.' *Journal of Public Administration Research and Theory* **1**: 19–47.

Law AM, Kelton WD. 2000. *Simulation Modeling and Analysis* (3rd edn). McGraw-Hill: Boston, MA.

Lawson T. 1997. *Economics and Reality*. Routledge: London, U.K.

Leamer E. 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. Wiley: New York.

Lee AS. 1991. Integrating positivist and interpretive approaches to organizational research. *Organization Science* **2**: 342–365.

Lewis MW, Grimes AJ. 1999. Metatriangulation: building theory from multiple paradigms. *Academy of Management Review* **24**: 672–690.

Lieberson S. 1992. Einstein, Renoir, and Greeley: some thoughts about evidence in sociology. *American Sociological Review* **57**: 1–15.

Lin Z, Hui C. 1999. Should lean replace mass organization systems? A comparative examination from a management coordination perspective. *Journal of International Business Studies* **30**: 45–80.

Lipton P. 1991. *Inference to the Best Explanation.* Routledge: London, U.K.

Locke EA. 2007. The case for inductive theory building. *Journal of Management* **33**: 867–890.

Losee J. 2005. *Theories on the Scrap Heap: Scientists and Philosophers on the Falsification, Rejection, and Replacement of Theories.* University of Pittsburgh Press: Pittsburgh, PA.

Machamer P, Darden L, Craver CF. 2000. Thinking about mechanisms. *Philosophy of Science* **67**: 1–25.

Mahoney JT. 1993. Strategic management and determinism: sustaining the conversation. *Journal of Management Studies* **30**: 173–191.

Mäki U. 2000. Kinds of assumptions and their truth: shaking an untwisted F-twist. *Kyklos* **53**: 317–336.

Manicas PT. 1987. *A History and Philosophy of the Social Sciences*. Basil Blackwell: Oxford, U.K.

Manicas PT. 2006. *A Realist Philosophy of Social Science: Explanation and Understanding*. Cambridge University Press: Cambridge, U.K.

Meehl PE. 1978. Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology* **46**: 806–834.

Miller KD. 2005. The problem of method and the practice of management research. In *Research Methodology in Strategy and Management* (Vol. 2), Ketchen DJ Jr, Bergh DD (eds). Elsevier Science: Amsterdam, The Netherlands; 143–177.

Mingers J. 2001. Combining IS research methods: towards a pluralist methodology. *Information Systems Research* **12**: 240–259.

Mingers J. 2004. Future directions in management science modeling. In *Critical Realist Applications in Organisation and Management Studies*, Fleetwood S, Ackroyd S (eds). Routledge: London, U.K.; 164–193.

Mingers J. 2006. A critique of statistical modelling in management science from a critical realist perspective:

its role within multimethodology. *Journal of the Operational Research Society* **57**: 202–219.
Mir R, Watson A. 2000. Strategic management and the philosophy of science: the case for a constructivist methodology. *Strategic Management Journal* **21**(9): 941–953.
Mir R, Watson A. 2001. Critical realism and constructivism in strategy research: toward a synthesis. *Strategic Management Journal* **22**(12): 1169–1173.
Mitchell TR, James LR. 2001. Building better theory: time and the specification of when things happen. *Academy of Management Review* **26**: 530–547.
Mohr LB. 1982. *Explaining Organizational Behavior*. Jossey-Bass: San Francisco, CA.
Mone MA, McKinley W 1993. The uniqueness value and its consequences for organization studies. *Journal of Management Inquiry* **2**: 284–296.
Nambisan S. 2002. Designing virtual customer environments for new product development: toward a theory. *Academy of Management Review* **27**: 392–413.
Nash R. 1999. What is real and what is realism in sociology? *Journal for the Theory of Social Behaviour* **29**: 445–466.
Nickel MN, Rodriguez MC. 2002. A review of research on the negative accounting relationship between risk and return: Bowman's paradox. *Omega* **30**: 1–18.
Nooteboom B. 1986. Plausibility in economics. *Economics and Philosophy* **2**: 197–224.
Numagami T. 1998. The infeasibility of invariant laws in management studies: a reflective dialogue in defense of case studies. *Organization Science* **9**: 2–15.
Orlikowski WJ. 1992. The duality of technology: rethinking the concept of technology in organizations. *Organization Science* **3**: 398–427.
Pajunen K. 2008. The nature of organizational mechanisms. *Organization Studies* **29**: 1449–1468.
Peacock MS. 2000. Explaining theory choice: an assessment of the critical realist contribution to explanation in science. *Journal for the Theory of Social Behaviour* **30**: 319–339.
Pentland BT. 1999. Building process theory with narrative: from description to explanation. *Academy of Management Review* **24**: 711–724.
Perrow C. 1984. *Normal Accidents: Living with High-risk Technologies*. Basic Books: New York.
Perrow C. 1994. The limits of safety: the enhancement of a theory of accidents. *Journal of Contingencies and Crisis Management* **2**: 212–220.
Petersen B, Pedersen T. 1997. Twenty years after—support and critique of the Uppsala internationalization model. In *The Nature of the International Firm*, Björkman I, Forsgren M (eds). Copenhagen Business School Press: Copenhagen, Denmark; 117–134.
Pfeffer J. 2007. A modest proposal: how we might change the process and product of managerial research. *Academy of Management Journal* **50**: 1334–1345.
Polanyi M. 1962. *Personal Knowledge: Toward a Post-critical Philosophy*. University of Chicago Press: Chicago, IL.
Popper KR. 1959. *The Logic of Scientific Discovery.* Basic Books: New York.
Porter S. 1993. Critical realist ethnography: the case of racism and professionalism in a medical setting. *Sociology* **27**: 591–609.
Powell TC. 2002. The philosophy of strategy. *Strategic Management Journal* **23**(9): 873–880.
Powell TC. 2003. Strategy without ontology. *Strategic Management Journal* **24**(3): 285–291.
Pozzebon M. 2004. The influence of a structurationist view on strategic management research. *Journal of Management Studies* **41**: 247–272.
Putnam H. 1974. The 'corroboration' of theories. In *The Philosophy of Karl Popper*, Schilpp PA (ed). Open Court Publishing: LaSalle, IL; 221–240.
Quine WVO. 1961. *From a Logical Point of View*. Harper and Row: New York.
Rahmandad H. 2008. Effect of delays on complexity of organizational learning. *Management Science* **54**: 1297–1312.
Rajagopalan N, Spreitzer GM. 1996. Toward a theory of strategic change: a multi-lens perspective and integrative framework. *Academy of Management Review* **22**: 48–79.
Rescher N. 1988. *Rationality: A Philosophical Inquiry into the Nature and the Rationale of Reason*. Clarendon Press: Oxford, U.K.
Rescher N. 2000. *Realistic Pragmatism: An Introduction to Pragmatic Philosophy.* State University of New York Press: Albany, NY.
Ricoeur P. 1981. *Hermeneutics and the Human Sciences: Essays on Language, Action and Interpretation*. Cambridge University Press: New York.
Roberts KH. 1989. New challenges in organization research: high reliability organizations. *Industrial Crisis Quarterly* **3**: 111–125.
Robins JA. 1987. Organizational economics: notes on the use of transaction-cost theory in the study of organizations. *Administrative Science Quarterly* **32**: 68–86.
Ron A. 2002. Regression analysis and the philosophy of social science: a critical realist view. *Journal of Critical Realism* **1**: 119–142.
Sagan SD. 1993. *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons.* Princeton University Press: Princeton, NJ.
Sarason Y, Dean T, Dillard JF. 2006. Entrepreneurship as the nexus of individual and opportunity: a structuration view. *Journal of Business Venturing* **21**: 286–305.
Sarasvathy SD. 2001. Causation and effectuation: toward a theoretical shift from economic inevitability to entrepreneurial contingency. *Academy of Management Review* **26**: 243–263.
Sawyer KR, Beed C, Sankey H. 1997. Underdetermination in economics. The Duhem-Quine thesis. *Economics and Philosophy* **13**: 1–23.
Sayer A. 1992. *Method in Social Science: A Realist Approach* (2nd edn). Routledge: London, U.K.
Sayer A. 2000. *Realism and Social Science*. Sage: London, U.K.
Schwarz N, Groves RM, Schuman H. 1998. Survey methods. In *The Handbook of Social Psychology* (4th edn, Vol. 1), Gilbert DT, Fiske ST, Lindzey G (eds). McGraw-Hill: Boston, MA; 143–179.

DOI: 10.1002/smj

Schweiger DM, Sandberg WR, Rechner PL. 1989. Experiential effects of dialectical inquiry, devil's advocacy, and consensus approaches to strategic decision making. *Academy of Management Journal* **32**: 745–772.

Searle JR. 2001. *Rationality in Action*. MIT Press: Cambridge, MA.

Shadish WR, Cook TD, Campbell DT. 2002. *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Houghton Mifflin: Boston, MA.

Søberg M. 2005. The Duhem-Quine thesis and experimental economics: a reinterpretation. *Journal of Economic Methodology* **12**: 581–597.

Starbuck WH. 2006. *The Production of Knowledge: The Challenge of Social Science Research*. Oxford University Press: New York.

Stinchcombe AL. 1968. *Constructing Social Theories.* Harcourt, Brace & World: New York.

Sutcliffe KM, Zaheer A. 1998. Uncertainty in the transaction environment: an empirical test. *Strategic Management Journal* **19**(1): 1–23.

Szulanski G, Jensen RJ. 2004. Overcoming stickiness: an empirical investigation of the role of the template in the replication of organizational routines. *Managerial and Decision Economics* **25**: 347–363.

Tsang EWK. 2006. Behavioral assumptions and theory development: the case of transaction cost economics. *Strategic Management Journal* **27**(11): 999–1011.

Tsang EWK, Kwan K-M. 1999. Replication and theory development in organizational science: a critical realist perspective. *Academy of Management Review* **24**: 759–780.

Tsoukas H. 1989. The validity of idiographic research explanations. *Academy of Management Review* **14**: 551–561.

Van de Ven AH. 2007. *Engaged Scholarship: A Guide for Organizational and Social Research.* Oxford University Press: New York.

Wason PC, Johnson-Laird PN. 1972. *Psychology of Reasoning: Structure and Content.* Harvard University Press: Cambridge, MA.

Wathne KH, Heide JB. 2000. Opportunism in interfirm relationships: forms, outcomes, and solutions. *Journal of Marketing* **64**(4): 36–51.

Webb EJ, Campbell DT, Schwartz RD, Sechrest L. 2000. *Unobtrusive Measures* (revised edn). Sage: Thousand Oaks, CA.

Weick KE. 1989. Theory construction as disciplined imagination. *Academy of Management Review* **14**: 516–531.

Wicks AC, Freeman RE. 1998. Organization studies and the new pragmatism: positivism, anti-positivism, and the search for ethics. *Organization Science* **9**: 123–140.

Williamson OE. 1996. Economic organization: the case for candor. *Academy of Management Review* **21**: 48–57.

Wuisman JJJM. 2005. The logic of scientific discovery in critical realist social scientific research. *Journal of Critical Realism* **4**: 366–394.

Yeung HW. 1997. Critical realism and realist research in human geography: a method or a philosophy in search of a method? *Progress in Human Geography* **21**: 51–74.

Zhang Z, Zyphur MJ, Preacher KJ. 2009. Testing multilevel mediation using hierarchical linear models: problems and solutions. *Organizational Research Methods* **12**: 695–719.

DOI: 10.1002/smj