Critical realism: a suitable vehicle for entrepreneurship research?

How to cite:


For guidance on citations see FAQs
Critical realism: a suitable vehicle for entrepreneurship research?

Richard Blundel

This chapter provides an outline of the origins and principal features of critical realist social theory, followed by a focused review of the methodological implications of this philosophical perspective. The primary purpose of the chapter is to consider why critical realism might offer a suitable ‘vehicle’ for qualitative research in the field of entrepreneurship, and to assess its explanatory potential with reference to recent empirical studies informed by a realist perspective. The concluding section is a reflection on the issues faced by researchers who are considering critical realism against alternative approaches, together with suggestions for further reading.

INTRODUCTION

Chapter two introduced the principal paradigms available to entrepreneurship researchers, and also highlighted some broad ontological and epistemological themes concerning the potential choice of methodological ‘vehicle’ for a particular study. In the present chapter, I consider how one of the more widely-cited social theoretic paradigms, critical realism, might be employed. The opening section comprises a brief account of the critical realism (‘CR’) in the context of social science research, which outlines its principal features, indicates its distinctive ontological and epistemological assumptions, and locates CR in relation to its antecedents and to some competing approaches. The central section includes a more focused appraisal of CR as the basis for research methodology, including its relevance to qualitative research in the entrepreneurship field [1]. The discussion is illustrated with examples of recent empirical work that has been informed by a CR perspective. The chapter concludes with a reflection on the methodological issues facing researchers who may be considering the use of CR against rival approaches and some suggestions for further reading.

RESEARCHING IN A CRITICAL REALIST PERSPECTIVE

Origins and development

The philosophical perspective now widely known as critical realism has gained in prominence over the last thirty years, during which it has made a transition from the natural sciences into social theory, leading to applications in various fields of social science. The core concepts of CR reflect a long tradition of realist philosophy, but its more recent development can be traced to the work of two philosophers of science, Rom
Harré and Roy Bhaskar. Harré’s influential (1972) *The philosophies of science* and Bhaskar’s (1975) *A realist philosophy of science*, established what was termed a ‘transcendental realist’ view of the relationship between the nature of human knowledge and that of objects of investigation in the natural sciences. In his (1979) work, *The possibility of naturalism*, Bhaskar extended these principles to the realm of the social sciences. In doing so, he reworked the term, ‘naturalism’, referring to the claim that there can be a unity of method between the natural and social sciences, into a ‘critical naturalism’, which acknowledges real differences in the nature of the objects investigated. The core ideas of CR flow from this combination of transcendental realism and critical naturalism.

The underlying position is that social scientists are engaged in a similar project to their counterparts in the natural sciences, but that researching social phenomena requires a distinctive set of methodological tools. Empirical researchers have attempted to apply, adapt and refine CR’s philosophical propositions in various fields, including: economic geography (Sayer and Morgan 1986); economics (Lawson 1997, Fleetwood 1999); and organisation studies (Ackroyd and Fleetwood 2000, Fleetwood and Ackroyd 2004), resulting in many different perspectives and emphases (Danermark *et al.* 2002: 1). Underlying this variety is a common concern with a central question in social science, human agency and its relationship with social structure. This concern can be traced back to the rise of CR, which was associated with the rejection of ‘structurist’ grand narratives, and corresponding efforts to recognise the role that knowledge and meaning played among human actors. Interestingly, in relation to qualitative research methods, much of the growing interest in CR appears to have been stimulated by direct experience in the field. For example, like other researchers in urban, regional and industrial studies, Andrew Sayer (2000: 5) found it impossible to reconcile the richness, complexity and sheer variety encountered in concrete social worlds with the tidy abstractions demanded by the ‘all-embracing, all-explaining’ discourses of this period. CR offered a ‘middle way’ for social scientific research, avoiding both reductionist forms of modernism, that took little or no account of interpretive understanding, and the problems of relativism and incommensurability that followed from postmodernism’s discursive ‘turn’ (*ibid*: 67-80) [2].

**Principal features**

The aim of this section is to introduce some of the principal features of CR that readers are likely to encounter in the literature, using relatively straightforward language and illustrations. It is clearly impossible to encompass a philosophical position in a few short paragraphs, without omitting or compressing many of the complex arguments upon which it is based. Consequently, I have focused attention on the methodological aspects of CR, taking the viewpoint of a researcher who may be considering this paradigm for a particular empirical study. The discussion is divided into four parts. The first two parts deal with CR’s world-view, introducing the terms: ‘structures’, ‘mechanisms’, ‘causal powers’, ‘stratification’ and ‘emergence’. The remaining parts discuss ‘critical naturalism’, the focal concept that connects CR to its philosophical roots in the natural sciences, and ‘retroduction’, CR’s distinctive mode of scientific inference and explanation. Each part is illustrated with examples from the natural and social worlds, see Figure 3.1.
Structures, mechanisms and causal powers. The term ‘structure’ refers to the way an object is constituted. Hence, the structure of a natural object, such as a water molecule, is based on the fusing of one hydrogen atom with two oxygen atoms. Similarly, a social object, such as an entrepreneurial network, is based on interactions between individual human beings. By virtue of its structure, any object has certain ‘causal powers’. These are the things that an object is able to do, or more broadly, its, ‘potentials, capacities, or abilities to act in certain ways and/or to facilitate various activities and developments.’ (Lawson 1997: 21). Hence, water has the capacity to extinguish a fire and an entrepreneurial network can form the basis for a series of different ventures over time (Johannisson 2000). Critical realists also make use of the term ‘mechanisms’ when referring to the ways that the causal powers of an object are exercised. These mechanisms are sometimes described as ‘generative’, in the sense they can give rise to concrete phenomena, such as an event that we might experience. However, activation of causal powers is not automatic, since it depends on the presence of other conditions. Hence, as Sayer (2000: 58) has noted, ‘a particular mechanism can produce completely different actions at different time, and inversely, the same event can have completely different causes’. To take a highly simplified example, two individuals might have similar capacities to become successful entrepreneurs, yet due to differing conditions (e.g. prevailing socio-economic conditions in their respective home regions), only one of them might realise her potential. Another implication is that similar events can be the product of an entirely different pattern of causes. Distinguishing these ‘contingent’ relationships between mechanisms is central to CR’s view of causation as depicted in Figure 3.2 [3].

Figure 3.2  A critical realist view of causation

Source: Sayer (2000: 15, Figure 1.2).
Stratification and emergence. CR asserts that the social world consists of real objects that exist independently of our knowledge and concepts, and whose structures, mechanisms and powers are often far from transparent. This reflects a well-established realist tenet concerning the independence of the world from our thoughts about it (Sayer 2000: 10). As Danermark et al. (2002: 20) have noted, the CR proposition that reality has hidden depths is hardly remarkable. It is not only a prerequisite for scientific activity, but also part of everyday experience, when people conjecture amongst themselves as to what may be going on ‘behind’ or ‘beneath the surface of’ an observed event (e.g. after witnessing extreme weather conditions, or the decline of an industrial district). However, CR does present researchers with a distinctive view of the world, and of their relationship to both natural and social phenomena. In Bhaskar’s (1975: 56) terms, reality consists of three domains, the empirical, the actual and the real. -The world of human experience and knowledge of events (the ‘empirical’ domain), is seen as ontologically distinct (i.e. separate and different) from the ‘actual’ domain in which events occur, irrespective or whether people have observed them. Thus, while different teams of climate scientists may produce competing theories about extreme weather events, the natural phenomena that they study remain the same (n.b. in the case of the social world this relationship with science is rather more complex; social phenomena are themselves products of human knowledge, so do not enjoy the same independent existence as their natural counterparts - see ‘critical naturalism’, below) [4]. -The further distinction of a ‘real’ domain, comprising structures and associated mechanisms, signals CR’s decisive break with the so-called ‘flat’ ontologies, most commonly associated with empirical realist and interpretivist philosophies of science. Realists argue that these paradigms place inappropriate limits on the scope of scientific exploration of the social world, in the first instance ignoring anything that is unobservable by researchers, and in the second, confining research to the direct experiences or accounts of human actors (Sayer 2000: 11). Hence, from a CR perspective, an entrepreneur’s account of her experience in starting a new venture only provides a provisional starting-point for explanation (Bhaskar 1979: 80, Whittington 1989: 85-86). -One of the primary tasks of science is to probe beneath the ‘empirical’ and ‘actual’ domains in pursuit of generative mechanisms that occupy ontologically distinct strata. For human actors, the potential for agency arises from the resulting interactions between different strata:

‘Just as for society as a whole, none of these strata provide any unique or dominant determination, but each presents a range of courses according to which actors can direct their activities. At the dinner table, guests are torn between the physiological drive of hunger, psychological tendencies towards greed and social pressures for delicate good manners.’

(Whittington 1989: 88 – emphasis added)

In the case of entrepreneurship research, it has long been recognised that investigations restricted to single strata (e.g. explanations based on efforts to isolate the psychological traits of ‘successful’ entrepreneurs), are likely to prove unsatisfactory (cf. Low and Macmillan 1988, Aldrich and Zimmer 1986). However, this begs the question of how the properties of different strata are related to one another. CR’s response is the proposition that both the natural and social worlds are characterised by the concept of ‘emergence’. This suggests that when the properties of different strata combine, they give rise to qualitatively new phenomena, or objects. More precisely, these new objects are emergent in the sense of possessing new properties – structures, causal powers and mechanisms – that depend upon, but cannot be reduced to, those of their constituents (Sayer 2000: 12-13, Danermark et al. 2002: 59-66). Bhaskar (1975: 169) illustrated this point by reconstructing the historical development of chemistry, in which an observable
chemical reaction was explained in terms of the properties of objects in successively ‘deeper’ strata (i.e. electrons, sub-atomic particles). In this example, the structures and associated causal powers (i.e. chemical bonding) of the ‘higher’ strata are emergent, and therefore fundamentally different in nature, from those of the underlying strata. Social structures, their causal powers and mechanisms are seen as being similarly emergent from human interaction. For example, while recognising that entrepreneurial networks are a product of interaction between individuals, CR also directs attention to the new and non-reducible properties of the network itself, including its structural form, causal powers and the mechanisms through which these are exercised. Realists argue that disregard for stratification and emergent powers has undermined social research, contributing to reductionist explanations, the misidentification of causality and the perpetuation of territorial disputes between theories and disciplines (Sayer 1992: 120) [5].

Critical naturalism. This concept derives from Bhaskar’s efforts to work through the implications of transcendental realism for the social sciences. Critical naturalism can be seen, in simple terms, as CR’s strategy for accommodating ‘messy’ and ‘ambiguous’ social phenomena, without abandoning the social scientific task. In common with interpretivists, and those who pursued the postmodern ‘turn’, critical realists have rejected ‘naturalism’, recognising that the social world cannot be understood in the same way as its natural counterpart (see also chapter four). However, in contrast to these paradigms, realists have been unwilling to stop their search at the level of meaning, but prefer to see its interpretation as merely the starting point for the pursuit of deeper causal explanations [6]. The following short extracts from the literature indicate some of the more important differences that realists have attempted to address, as CR philosophy has been translated from the natural world in order to encompass social phenomena. For researchers, it has meant taking due account of distinguishing characteristics of the social world, including: the impact of intentionality on human action (i.e. our purposeful pursuit of perceived goals, such as happiness or profit); the emergent nature of social structures, such as marriage or organisation, which are both relatively autonomous and inherently meaningful; and the complex relationship between agency and structure that this implies:

‘Our pursuit of a separate science in the social sphere, centred upon the intentionality of human agency and involving a recognition of the reality and relative autonomy of action-conditioning social structure, amounts to an acknowledgement of the irreducibility of society to nature.’
(Lawson 1997: 63)

‘What does it mean to write of the social world? The natural world is natural because it does not require action on behalf of human beings for its existence. The social world is social because, by contrast, it does require action on behalf of human beings for its existence.’
(Ackroyd and Fleetwood 2000: 10)

‘Critical realism acknowledges that social phenomena are intrinsically meaningful, and hence that meaning is not only externally descriptive of them but constitutive of them (though of course there are usually material constituents too). Meaning has to be understood, it cannot be measured or counted, and hence there is always an interpretive or hermeneutic element in social science.’
(Sayer 2000: 17)

In summary, while the causal powers of natural objects, such as weather systems, are exercised ‘mindlessly’, without any (self-conscious) sense of meaning, interpretation and intent, those of social objects, such as entrepreneurial activities, display these
characteristics in abundance. The implication is that social scientists need to engage in a so-called ‘double hermeneutic’, generating explanatory knowledge about phenomena that are themselves ‘knowing’, in contrast to their natural science counterparts, whose subject-matter is ‘unknowing’. This highlights a central tension arising from CR’s ontology. Given the proposition that science seeks to explain a world consisting of ‘real’ objects, which in CR terms represent the ‘intransitive’ or objective dimension of knowledge, how is it to incorporate this ‘transitive’ or subjective dimension? Bhaskar’s (1975) concept of critical naturalism acts as the conceptual bridge between these competing demands:

‘[C]ritical realism is only partly naturalist, for although social science can use the same methods as natural science regarding causal explanation, it must also diverge from them in using “verstehen” or interpretive understanding. While natural scientists necessarily have to enter the hermeneutic circle of their scientific community, social scientists also have to enter that of those whom they study.’

(Sayer 2000: 17)

It is clear that the concept of critical naturalism represents a far-reaching methodological challenge to empirical researchers (Danermark et al. 2002: 38-39). Consequently, any conclusions that we reach regarding the empirical application of critical naturalism are likely to be central to our assessment of CR as a suitable ‘vehicle’ for entrepreneurship research.

**Retroduction.** CR has adopted a distinctive form of scientific inference, termed ‘retroduction’, which involves the explanation of events in the social world by seeking to discern the structures and mechanisms that are capable of producing them (Sayer 1992: 107). This explanatory task involves quite different methodological operations to those associated with ‘induction’ and ‘deduction’ [7]. Consider, for example, a research project investigating the growth of entrepreneurial firms: *inductive* inference might move from a series of similar observations to an empirical generalisation such as, ‘rapid growth is associated with variables X, Y and Z.’; *deductive* inference might move from a set of premises, such as the existence of certain variables to a conclusion their implications for growth in a particular case; while *retroductive* inference would move from the description and abstract analysis of the growth process as a concrete phenomenon to a reconstruction of the basic conditions (i.e. the structures, causal powers and mechanisms) that make it possible [8]. Retroduction involves a type of scientific generalisation that is concerned with the isolation of fundamental structures whose powers can be said to act ‘transfactually’ (i.e. continuing to exist, even though their operations may not be manifested at the level of events or observations). Its ‘analytical’ approach to generalisation contrasts sharply with the more common type associated with inductive inference, which focuses on the extrapolation of empirical regularities (Danermark et al. 2002: 77) (Figure 3.3). As a consequence, retroduction requires different scientific methods in order to achieve its purposes (Easton 2000: 214).
So what are the implications of retroduction for the working practices of social scientists? Some of the more important issues can be illustrated with reference to an imaginary research study involving case studies of entrepreneurial firms. First, in their effort to reflect the inherent complexity of concrete phenomena, the researchers are likely to draw on multiple sources of data, which may comprise various types of qualitative evidence, derived from ethnography, observation, in-depth interview, historical and archival research, as well as some quantitative evidence, such as industry statistics. Second, in selecting cases, the researchers are guided by the requirements of analytical, rather than empirical, generalisation. This means that they select cases in order to explore and to clarify the necessary and contingent relationships between structures (Danermark et al. 2002: 105). To achieve this, their selection might include some extreme or ‘pathological’ cases, where firms have experienced major transitions or crises (Bhaskar 1979: 48, Collier 1994: 165). In addition, they pay considerable attention to both the spatial and temporal boundaries of case-based research, in an effort to ensure that wider structural conditions are addressed (Whittington 1989: 85). Third, in sifting through their rich idiographic sources, the researchers incorporate the accounts of human actors, not simply in their own terms, but as part of the search for the ‘rules’ that constitute these accounts (Tsoukas 1989: 555). For example, the researchers treat entrepreneurs’ statements about the perceived constraints of the growth of their firms as a starting-point for a retroductive probing of the structural preconditions of these perceptions. Lastly, the study itself proceeds through several iterations, with the researchers moving repeatedly between more concrete and more abstract activities in order to refine their explanation. In Tsoukas’s (1989: 558) terms, the are moving concurrently on two tracks, one of which is ‘up in the clouds’, and concerned with abstraction and theoretical conceptualisation, while the other is ‘down to earth’, engaged in the idiosyncratic details of the case material. The process has been described in a model comprising five distinct but closely-related activities (Danermark et al. 2002: 109-111) (Table 3.1). As the authors have emphasised, the model is not prescriptive, nor does it imply a strictly linear process. The emphasis on different activities is also bound to vary, according to the nature of a particular research project, as are the actual research methods employed (ibid: 109, 73). However, it provides a concise summary of

![Figure 3.3 Two types of generalisation](source: Danermark et al. (2002: 77, Figure 3))
the preceding discussion, illustrating the distinctively retroductive methodological implications that social scientists have derived from the CR paradigm.

Table 3.1  An explanatory research process involving retroduction

<table>
<thead>
<tr>
<th>Activity</th>
<th>Nature of activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Description</td>
<td>Prepare a description of the phenomenon, making use of actors’ accounts and a variety of other sources.</td>
</tr>
<tr>
<td>2: Analytical resolution</td>
<td>Distinguish various components, aspects or dimensions of the phenomenon and establish (tentative) boundaries to the components studied.</td>
</tr>
<tr>
<td>3: Theoretical redescription</td>
<td>Interpret and redescribe the different components, applying contrasting theoretical frameworks and interpretations in order to provide new insights (n.b. this activity is sometimes referred to as ‘abduction’).</td>
</tr>
<tr>
<td>4: Retroduction</td>
<td>For each component, seek to identify basic, or ‘transfactual’ conditions, including structures, causal powers and mechanisms, that make the phenomenon possible.</td>
</tr>
<tr>
<td>5: Abstract comparison</td>
<td>Elaborate and estimate the explanatory power of the structures, causal powers and mechanisms that have been identified during activities 3 and 4.</td>
</tr>
<tr>
<td>6: Concretization and contextualization</td>
<td>Examine how different structures, causal powers and mechanisms manifest themselves in concrete situations.</td>
</tr>
</tbody>
</table>

Source: Danermark et al. (2002: 109-111, Table 4 – modified). Note: the term ‘activities’ has been substituted for the original ‘stages’ in order to emphasise the non-linear nature of the process.

IS CRITICAL REALISM RELEVANT TO ENTREPRENEURSHIP RESEARCH?

In this section, I discuss several reasons why the critical realist paradigm might provide a suitable vehicle for entrepreneurship research, with specific reference to qualitative approaches. As the editors have indicated, the methodological debate in our field is at best at a highly provisional stage. With this in mind, I have presented the material in the form of a rhetorical case for CR-inspired research, intended to stimulate discussion (n.b. critics and alternative approaches are addressed in a later section). The argument builds on five themes: first, that CR that can help to revive a longstanding realist tradition in entrepreneurship research; second, that CR can promote the much-needed contextualisation of entrepreneurial phenomena in research studies; third, that CR can facilitate greater theoretical integration between disciplines and across multiple levels of analysis; fourth, that CR can enhance the explanatory potential of existing qualitative research techniques, including the case study approaches; and fifth, that as a
consequence, CR has the potential to contribute more ‘useful’ knowledge than rival paradigms.

**Reviving a realist tradition**

Realism has long intellectual roots in entrepreneurship research, and its contributory disciplines (Ackroyd and Fleetwood 2000: 9, Swedberg 2000: 12-18). For example, it is possible to detect a common thread of ideas in economics, emerging out of its polarisation in the *Methodenstreit* (i.e. battle over methods) at the end of the 19th century. The pioneering sociologist, Max Weber, proposed a new approach to overcome the divide between an overly-abstract, non-historical version of economics, and an overly-historical, non-theoretical one. Weber’s *Sozialökonomik*, an attempt to synthesise history with theory, had a great impact on Joseph Schumpeter’s thinking (Swedberg 1991: 83-89), including his approach to entrepreneurship:

> ‘[The] sociology of enterprise reaches much further than is implied in questions concerning the conditions that produce and shape, favour or inhibit entrepreneurial activity. It extends to the structure and the very foundations of capitalist society.’

(Schumpeter 1951: 224-225)

Schumpeter’s ideas influenced Edith Penrose, whose seminal (1959) study, *The Theory of the Growth of the Firm*, reflects a similar realist concern with uncovering structures and mechanisms, and specifically those ignored by mainstream economics in its ‘black box’ treatment of the firm. Penrose’s interest was sparked by involvement in a substantial piece of qualitative research, examining the growth of a former subsidiary of Du Pont (Penrose 1960). Her eclectic theory incorporates a subtle treatment of meaning and intentionality in human actors (i.e. the dynamics of entrepreneurial judgement at the level of the managerial team, encapsulated in her concept of ‘productive opportunity’), but also acknowledges the relative autonomy of environmental selection mechanisms [9]:

> “Expectations” and not “objective facts” are the immediate determinants of a firm’s behaviour, although there may be a relationship between expectations and “facts” - indeed there must be if action is to be successful... In the last analysis the “environment” rejects or confirms the soundness of the judgements about it, but the relevant environment is not an objective fact discoverable before the event’.

(Penrose 1959: 41)

Penrose’s emphasis on the subjective element, whereby firm behaviour is, in the first instance, the product of an ‘image’ of the environment in the mind of the entrepreneur (Boulding 1956), contrasts with much of the later resource-based literature. However, by elaborating her theory, she helped to perpetuate a strand of research that retains a strong realist flavour (e.g. Lawson and Lorenz 1999, Best 2001). Investigations may start at the level of entrepreneurial perceptions, but their scope should be much broader; researchers are challenged to discover how the phenomenon that Penrose conceptualised as ‘productive opportunity’ articulates with other structures and mechanisms.

**Contextualising entrepreneurship**

Critical realism raises questions about the pre-conditions for social phenomena. It is therefore well-placed to frame an investigation into contextual and process issues. In
considering the context in which entrepreneurship occurs, we begin to raise important questions about the boundaries, both temporal and spatial, of our research:

‘We need to know not only what the main strategies were of actors, but what it was about the context which enabled them to be successful or otherwise. This is consistent with the realist concept of causation and requires us […] to decide what it was about a certain context which allowed a certain action to be successful. Often the success or failure of agents’ strategies may have little or nothing to do with their own reasons and intentions’.

(Sayer 2000: 26)

Many contributors have called for greater attention to be paid to the context in which entrepreneurial activity takes place (Low and Macmillan 1988, Zafirovski 1999, Ucbasaran et al. 2001). For example, entrepreneurial networks have been identified as important contextual phenomena that display degrees of social embeddedness (Granovetter 1985, Johannisson and Monsted 1997) and latency (Ramachandran and Ramnarayan 1993). Network-based case studies have also been used to deconstruct the (culturally-conditioned) myth of entrepreneurs as ‘heroic’ individuals (Jones and Conway 2000). However, leading figures continue to argue that interaction between entrepreneurial activity and the broader context is a relatively underdeveloped research area (Acs and Audretsch 2003: 329, Davidsson and Wiklund 2001: 81-12). The potential contribution of CR is to facilitate a more nuanced understanding of the context in which entrepreneurs exercise strategic choice; CR’s mechanisms-based paradigm is seen as a moderating influence on excessively voluntaristic (and deterministic) accounts of entrepreneurial agency (Whittington 1989: 75). Building on CR’s methodological precepts, entrepreneurial research should be capable of better spatial and temporal explanations, tracing the changing ‘zones of manoeuvre’ of entrepreneurial firms as they interact with the competitive capacities of their contexts (Clark 2000: 303-313).

**Integrating different levels of analysis**

Entrepreneurship research has blossomed in many academic disciplines, including psychology, anthropology, organisation studies, geography, economic history and economics. These activities have generated a rich and diverse harvest of empirical and conceptual material. However, this variety masks the fact that the field is fragmented, with specialists making little use of one another’s work (Ucbasaran et al. 2001: 57). Furthermore, in pursuing the methodologies traditionally associated with these disciplines, entrepreneurship researchers have tended to focus their attention on particular levels of analysis. In their comprehensive review of ‘past research and future challenges’, Low and MacMillan (1988: 151-152) suggested that entrepreneurship researchers may choose among five levels of analysis in pursuit of relevant phenomena: the individual, group, organisational, industrial and societal. They noted a tendency for most previous research to be conducted at a single level of analysis, but argued that a few recent examples of multi-level research (e.g. Aldrich and Auster 1986), demonstrated the potential for achieving a richer understanding of entrepreneurship processes. This led them to conclude that both entrepreneurship research designs would be enriched if they were able to incorporate multiple levels of analysis:

‘The relationships between phenomena that can be observed at different levels of analysis are important not just for academics, but for both practitioners and public policy makers as well. From the entrepreneur’s perspective, the success of the individual enterprise will be affected by factors that can only be observed at different levels of analysis. *To miss any one of these perspectives increases the probability that key factors will be overlooked and that unanticipated*
events will take the entrepreneur by surprise. From the public policy-maker’s perspective, the insights generated by multi-level studies have the potential to improve targeting of government efforts to encourage successful entrepreneurship.’

(Low and Macmillan 1988: 152 - emphasis added)

However, Davidsson and Wiklund’s (2001) review of current research practice, based on a content analysis of articles published in leading US and European entrepreneurship journals, revealed that research was dominated by micro-level analysis, with integrated ‘micro/aggregate mix’ approaches continuing to represent a small proportion of published work. While our diverse and primarily single-level research programmes have given rise to recurrent debates over the relative importance of, for example, psychological, organisational and socio-cultural dimensions of entrepreneurship, they have achieved little empirical or conceptual integration (Frank and Landström 1997; Davidsson et al. 2001). For example, entrepreneurship researchers employ a variety of strategies to build or refine process theories. Each seeks to understand ‘patterns in events’, but methodologies differ in the extent to which they probe beyond observed events (i.e. surface-level effects) in order to understand underlying causal sequences or generating mechanisms (Pentland 1999). This is not to deny the many insights into entrepreneurial processes that have already been achieved. For example, population ecologists have made productive use of a single-level methodology, exploring macro-level processes with data that is primarily aggregated and quantitative (i.e. official statistical data sets recording firm entries and exits) (Aldrich and Zimmer 1986, Staber 1997). Similarly, ethnographic researchers, who also tend to apply a single-level methodology, have revealed richly-detailed micro-level processes through direct exposure to localised fieldwork sites, making imaginative use of qualitative research methods (i.e. observing entrepreneurs and recording their perceptions and behaviours) (Ram 1999). Rather, as proponents would argue, a CR-inspired methodology is capable of taking entrepreneurship research a step further, supporting new research strategies better geared to achieve integration across its traditional divides (cf. Layder 1993, Danermark et al. 2002) [10].

Enhancing qualitative research

CR is compatible with a range of qualitative research methods. Its potential role in relation to qualitative evidence can be illustrated with reference to one of the leading texts in this field (Miles and Huberman 1994). As the authors suggest, the decision to adopt a realist perspective may have little impact on data collection. However, research strategies will be affected by the imperatives of critical naturalism and retroductive analysis:

‘Human relationships and societies have peculiarities that make a realist approach to understanding them more complex - but not impossible. Unlike researchers in physics, we must contend with institutions, structures, practices and conventions that people reproduce and transform [...] Things that are believed become real and can be inquired into.’

(Miles and Huberman 1994: 4)

Though it has few references to Bhaskar and Harré, this widely-adopted sourcebook has added considerable substance to CR’s earlier methodological reflections. For example, its approach to ‘within case displays’ illustrates some of the challenges in causal explanation, contrasting investigations that are limited to a single level of analysis to more complex, multi-level approaches. The authors argue that qualitative research methods are particularly amenable to this type of causal analysis:
'Qualitative analysis, with its close-up look, can identify mechanisms, going beyond sheer association. It is unrelentingly local, and deals well with the complex network of events and processes in a situation. It can sort out the temporal dimension, showing clearly what preceded what, either through direct observation or retrospection. It is well-equipped to cycle back and forth between variables and processes - showing that “stories” are not capricious, but include underlying variables, and that variables are not disembodied, but have connections over time.'

(Miles and Huberman 1994: 147)

These techniques are broadly consistent with a CR position, and suggest that researchers should proceed through a combination of what they term a ‘variable-oriented’ conceptual approach (i.e. looking for patterns, or configurations in the data), and a ‘process-oriented’ approach (i.e. assembling chronologies, or stories). The overall emphasis is towards retroductive inference:

'[We are] proposing that answering good “why” and “how” questions requires us to go beyond sheer association to seeing the actual mechanisms of influence in a bounded local setting, which are always multifold, operating over time.'

(Miles and Huberman 1994: 170)

The implication, which echoes the previous argument concerning multi-level analysis, is that a CR-inspired methodology can contribute to better outcomes when researchers are employing qualitative research methods. More specifically, by highlighting the role of unobserved social structures, causal powers and mechanisms, the CR ontology can act as a counterbalance to the ‘micro-sociological’ tendencies of context-specific qualitative approaches such as ethnography (Porter 2002: 142, 157). Relatedly, CR’s fundamental concern with explaining why things occur, and with analysis through a process of retroductive inference, can challenge researchers to move beyond the description of social situations to a more critical assessment of the relationship between structural factors and human agency (ibid: 156-157).

**Generating more ‘useful’ knowledge**

In order to intervene successfully in the world, it is useful to obtain a working knowledge of the relevant structures and generative mechanisms. Or, to paraphrase Kant’s widely-cited aphorism, ‘There is nothing so practical as a good theory.’ The principal advantage of CR’s retroductive methodology, from the perspective of the policy-maker or practitioner, is that its purpose is to develop a theoretical understanding of real mechanisms, and the contingent ways in which they combine to generate effects (e.g. Subramaniyam 2000). While isolated, subjective accounts of entrepreneurial agency may be engaging, they have no referent and therefore lack cumulative explanatory power. With its concern for underlying structure rather than surface-level correlations, its opposition to excessive voluntarism and determinism, and its critique of reductionist explanations, CR seems well-placed to deliver a more informed – though, it has to be conceded, not always ‘actionable’ – understanding of concrete situations. At present, it is difficult to substantiate this argument, given the limited number of published studies that combine a CR methodology with an explicit policy orientation. However, some provisional conclusions may be drawn from three cases presented in the next section, which illustrate contrasting empirical applications of a broadly realist perspective.
The empirical challenge

This section provides examples to illustrate the proposition that research drawing on a CR perspective is capable of delivering more informed explanations of entrepreneurial activity. It reflects repeated calls to move beyond conceptual integration and attempt to replicate it in concrete, empirical research (Aldrich and Martinez 2001: 51). I will focus on three studies, each reflecting different aspects of the entrepreneurial networks agenda: Best’s (2001) analysis of the dynamics of entrepreneurial firms and regional clusters is not explicitly critical realist in approach, yet displays realism’s capacity for integration across multiple levels of analysis; Jones’s (2001) examination of divergent strategies of technology- and content-driven entrepreneurs in the early years of the US film industry combines realism with a narrative approach; and Bowey and Easton’s (2003) study adopts a CR methodology to explain changes in social capital in relationships between entrepreneurs and other actors. The aim is to connect the methodological debate to concrete research practices, noting both the limitations and potential of the paradigm [11].

The dynamics of entrepreneurial clusters

Michael Best’s recent work addresses ‘cluster dynamics’, defined as, ‘interactive processes of capability development and specialization within and amongst firms within a region.’ (Best 2001: ix). It forms part of a research tradition concerned with processes of entrepreneurship, learning and adaptation both within and beyond the boundaries of the firm (Penrose 1959, Richardson 1972, Lawson and Lorenz 1999). Best’s systems integration model extends the spatial and temporal scope of the (neo-Penrosian) ‘technology capability and market opportunity’ mechanism and suggests how it might articulate with other mechanisms operating at several distinct levels of analysis (Figure 3.4). The resulting analysis of capability development in industrial districts may be interpreted, in CR terms, as highlighting the role of pre-existing structures and their associated latent causal powers, while also isolating the contingent relationships that can to these powers being exercised:

This model has been applied empirically to explain the changing fortunes of regional clusters, including the resurgence of high technology manufacturing in eastern Massachusetts and emerging cluster dynamics in the Malaysian electronics sector [12].

‘An industrial district, unlike any single firm, offers the potential for new and unplanned technology combinations that tap a variety and range of production-related activities. This protean character of technological capability, particularly evident in the high tech sectors, is a feature of industrial change even in the oldest sectors. […] Thus, a region’s technological capabilities are an outcome of a cumulative and collective history of technological advances embedded in entrepreneurial firms.’

(Best 2001: 81 – emphasis in original)
Figure 3.4  A cumulative model of cluster dynamics

Source: Best (2001: 70, Figure 3.1 - adapted, bracketed annotation added)

Entrepreneurial trajectories in Hollywood

Candace Jones (2001) has conducted a fascinating historical analysis of the interaction between entrepreneurial careers, institutional rules and competitive dynamics in the early American film industry. Jones’s methodology combines realist and narrative approaches, while her conceptual framework draws on insights from co-evolutionary, institutional and resource-based theorising:

‘Generative mechanisms are the underlying structures that drive processes (Pentland 1999) and in this study, they are firms’ institutional and strategic isolating mechanisms. A narrative approach illuminates how and why change occurs, by examining sequences of events (Van de Ven 1992) to reveal linkages amongst context and action (Pettigrew 1992)’.

(Jones 2001: 913)

The study makes use of a rich variety of qualitative and quantitative evidence, including firm-level archival data, published histories and industry statistics. These are used to probe the contrasting trajectories of ‘technology-driven’ and ‘content-oriented’ firms in an analytical scheme that encompasses the firms’ entrepreneurial practices, their capability-development and their co-evolutionary relationship an emerging structure of institutional rules (e.g. patent laws and artistic contracts).

Entrepreneurial social capital changes

James Bowey and Geoff Easton (2003) adopt a comparative case study approach, informed by a form of CR explanation, to examine the change of social capital in entrepreneurial network relationships. The two cases in this paper are based on contrasting business relationships involving one entrepreneur, ‘Jacques’, and two other actors. One of the cases records a process of social capital formation in a blossoming relationship, while the other traces a process of depreciation in a failing relationship. The narratives are framed using a common template that allows the researchers to probe for deeper ‘entities’ (i.e. structures), mechanisms and relationships. The research reveals similarities and differences that are not evident at the level of ‘surface’ events:
‘Both mechanisms were different because while the entities were the same the necessary and contingent relationships were not only individually different but so was their configuration. As a result they worked in different ways to cause different changes in social capital.’

(Bowey and Easton 1993: 18)

However, the authors conclude that the most important conclusion coming from their in-depth analysis concerns the difficulty in specifying causal mechanisms. They compare the role played by entities (i.e. social structures) and mechanisms in a realist paradigm with that of variables and correlations found in the ‘positivist’ research. This prompts the reflection that, though positivism’s simple ‘linear additive configurations’ are unlikely to provide useful representations of reality, ‘it is difficult to think in any other way when seeking to ascribe reasonably precise causal explanations’ (ibid: 18-19).

IS IT TIME FOR A ‘TEST-DRIVE’?

In this chapter, I have assessed the potential of critical realism as a suitable vehicle for exploring the phenomenon of entrepreneurship, with a particular reference to qualitative research. I have presented five broad arguments in support of this view. Ultimately, any methodological innovation must be subjected to a simple evaluative question. In short, to what extent can it enhance our understanding of the phenomenon we are studying?

The case for qualitative research informed by CR is that it has the potential to produce ‘better stories’, that could form the basis for more sophisticated causal explanations. Perhaps the most important limitation in narrative-based qualitative research, and one that has long been recognised in the debate between ‘models’ and ‘histories’, is that the complexity and idiosyncracy of narrative data tend to ‘crowd out’ fundamental mechanisms and relationships. One of the claims of the critical realist perspective is that it provides a basis for theoretically-informed abstraction, reflecting Marx’s earlier notion of an *histoire raisonné*. Thus, in the case of narrative-based research, CR demands a more rigorous and analytically sound periodisation of episodes than that found in much of the literature (Clark 2000: 115), with more explicit specifications of causality in the processes that it describes (Sayer 2000: 142-143). One thing is certain; the contribution of any methodology cannot be proven in the abstract. As the case examples have illustrated, there is much to gain from further testing and refinement in the field. This would be facilitated by a more creative interaction between the high ground of social theory and more earthly demands of empirical research.

**Critics and alternatives**

The principal case for CR is that offers the social scientist a distinctive methodological approach, which rejects both the naive optimism of those expecting to uncover law-like regularities from empirical data and the defeatism of those who deny any possibility of generalising our understanding of idiosyncratic phenomena such as entrepreneurship. As we have seen, CR is frequently presented as a kind of ‘third way’, providing a more sophisticated ontology than either empirical realism or postmodernism in its various forms (Ackroyd and Fleetwood 2000: 4-10, Sayer 1992: 4-7). However, the CR paradigm has also been subjected to sustained criticism, extending from its philosophical roots to the empirical studies it has inspired. The most extensive attacks have been on CR’s social theoretic propositions, which have been seen as both internally inconsistent and unoriginal (e.g. Baert 1998: 195-197, Parsons 1999, Roberts 2001). There has also been some questioning of the CR claim to provide a compelling
basis for social scientific methodology (Walters and Young 2001). The CR community has also proved to be an effective (self-)critic, mocking the ‘linguistic obscurity’ of some contributions to the CR literature (Junor 2001: 33), and warning against a common tendency to shift substantive social science issues, ‘into the terrain of philosophy.’ (Potter 2003: 163) As the writer notes, the tendency is problematic because, ‘philosophy cannot do social science’s job.’ (ibid: 163) With this thought in mind, sceptics might find themselves questioning the continuing shortage of substantive published studies that have adopted an explicitly CR methodology [13].

In reflecting on these critiques, we should note that critical realists are not the only social theorists who promote a methodology based upon a search for the underlying generative mechanisms that connect different states or events. There is a long-standing debate in sociology, between proponents of variable-centred approaches that make extensive use of statistical modelling techniques, and those who argue for mechanism-based theorising. Advocates of the ‘social mechanisms’ approach to sociological theorising, would agree with CR on the role played by mechanisms in the routine practice of social scientific research:

> ‘The belief in explanations that provide accounts of what happens as it actually happens has pervaded the sociological literature for decades and has produced an abundance of detailed descriptive narratives but few explanatory mechanisms of any generality. It is through abstractions and analytical accentuation, however, that general mechanisms are made visible.’
> (Hedström and Swedberg 1998: 15)

However, despite some commonality of purpose, there are important differences between the methodologies adopted by ‘social mechanisms’ scholars and those associated with CR. In the former case, empirical work tends to have a much stronger quantitative orientation and to be guided by the principles of methodological individualism. The ‘middle-range’ theorising advocated in this tradition is based on the argument that sociological researchers are equipped to pursue only relatively short causal histories (cf. Layder 1993: 19-37). In addition, while sharing with CR the assumption that social mechanisms, ‘usually are unobserved’, they are treated here as analytical constructs that simply assist in the process of theorising the links between observed events. In other words, these mechanisms, though ‘generative’, do not enjoy the special ontological status that is granted to them in the work of Harré and Bhaskar (Hedström and Swedberg 1998: 7-17) [14].

‘It is in practice …’

By way of a closing comment, it seems appropriate to return to the metaphorical image of CR as potential a ‘vehicle’ for entrepreneurship research. In the course of this chapter, I have reviewed a small proportion of a substantial CR literature that has been generated in a relatively short period, stimulated by the agenda-setting philosophical writings of Harré and Bhaskar that appeared in the 1970s. The review has focused on contributions from entrepreneurship researchers, and those in related fields, rather than those of social theorists. Despite this emphasis, one over-riding impression is that researchers have invested a disproportionate amount of energy in describing CR’s elaborate ontological features, and in debating the merits of its radical epistemological styling. The necessary investment in substantive research – let us call it ‘test driving’ CR – has been correspondingly underplayed. As a philosophy of science, transcendental realism was able to reflect on many centuries of empirical practice in the natural
sciences. By contrast, CR in the social sciences draws on a much thinner body of substantive work. If CR’s critics are to be believed, the tensions inherent in critical naturalism may direct entrepreneurship researchers elsewhere. For example, they may opt for less ambitious studies of particular aspects of entrepreneurship in a ‘social mechanisms’ framework, or construct much broader geo-historical narratives that are not constrained by social scientific conventions. In any event, the final test will be an empirical one. As the renowned realist, Karl Marx (1818-1883), once observed, ‘it is in practice that we prove ... that our thought is true.’

SUGGESTIONS FOR FURTHER READING

It is good practice to trace ideas to their original source, but as Junor (2001: 30) has noted, the philosophical texts that underpin CR, such as Harré (1972) and Bhaskar (1975, 1979) are not necessarily the best starting-point for most empirical researchers. The alternative is to turn to other philosophers, such as Collier (1994), who offers a fairly accessible and critical introduction to Bhaskar’s ideas, or to scholars from other disciplines, who can mediate between the high plateaux of philosophy and the practical challenges of social science research. Leading contributors in the latter group include: Sayer (1992, 2000), who discusses the methodological implications of CR in various fields, with a particular focus on geo-historical research; Lawson (1997), whose broad methodological critique of ‘mainstream’ economics methodologies is grounded in a CR perspective; and Archer (1995), who develops a distinctive ‘morphogenetic’ approach, placing particular emphasis on the time dimension and engaging in a strong critique of structuration theory [15]. Much of this work is summarised by Danemark et al. (2002), who provide a clear and coherent introduction to critical realism in the social sciences, with particular emphasis on methodology and practical application. As noted previously, there are relatively few published accounts of CR as it has been applied in the field of entrepreneurship. However, the volumes edited by Ackroyd and Fleetwood (2000) and Fleetwood and Ackroyd (2004) contain recent conceptual and empirical work in related areas of management and organisational studies (e.g. Porter’s (2000) CR ethnography), while the study by Whittington (1989) includes a CR-inspired methodological discussion. In order to develop a balanced view of CR, readers may also wish to pursue some of the leading critics, or to seek comparisons between CR and rival perspectives (e.g. Miri and Watson 2001, Mutch 2002). As noted above, there are many critiques of CR as social theory, but these tend to lack the methodological application that is of more immediate interest to empirical researchers. Walters and Young (2001) is an exception, with a challenge to CR’s methodological claims that is based on recent applications in the field of economics; the ‘debates’ section of Fleetwood (1999) contains some counter-arguments.

ACKNOWLEDGEMENTS

The author would like to thank the organisers of the first Interdisciplinary European Conference on Entrepreneurship Research (IECER), University of Regensburg, Germany (February 2003), conference participants, Stephen Ackroyd, Peter Clark, Michael Rowlinson, the editors of this volume and two anonymous referees for their support, comments and suggestions. The usual caveat applies.
REFERENCES


Critical realism has had relatively limited exposure in the entrepreneurship literature, in contrast to its profile in related areas, such as organisation studies and economic geography. However, as the editors of this handbook have indicated, there is much to gain from a more critical discussion concerning the often-implicit methodological choices that are made by researchers in this field.

This concern to develop ‘more open, context-dependent and plural’ accounts of the social world was echoed by the rise of postmodernism and post-structuralism, but argues that the development of CR in fields such as urban and regional studies happened ‘largely independently’ of their emergence (Sayer 2000: 5).

‘Necessary relations’ (or ‘necessity’) refer to things that must go together, given the nature of objects in the natural or social world, while ‘contingent relations’ (or ‘contingency’) refer to things that might go together. Researchers adopting a CR approach see the pursuit of questions about necessity as fundamental to the practice of theorising in the social sciences, forcing researchers to sharpen their conceptualisations of their objects of study.

Sayer (2000: 33-35) mounts a strong defence of critical realism against the sceptical implication that social phenomena cannot be treated in the same way as their natural world counterparts (i.e. as ‘intransitive’ objects of study). However, as one of the reviewers has suggested, while analogies from nature may help to clarify CR’s unfamiliar ontological and epistemological claims, they should always be used sparingly and with careful ‘translations’ to social examples.

Bhaskar (1975) presents a philosophical case for stratification and emergence, based on the existence of scientific practice. Realists argue that this proposition also corresponds to evidence from the natural and social worlds, giving examples such as the evolution of life on earth (Collier 1994: 46).

Sayer (2000) provides an insight into the way this tension between the need for interpretive understanding and that of causal explanation was experienced in the course of his own empirical work:

‘The empirical context was the prosaic one of studies of the development of urban and regional systems [...]. In attempting to develop an understanding of these that was both dynamic and spatial, it slowly dawned on me that social systems were necessarily open, and that they evolved rather than equilibrated, not least because people have the capacity to learn and change their behaviour. Consequently, I realized the goal of finding rough regularities, let alone laws, to describe social systems, was a pipe dream. At the same time, realist philosophy was beginning to challenge the regularity or secessionist theory of causation, and to analyse the explanation of change in open systems, so that it became clear that abandoning hopes of finding regularities in no way meant abandoning explanation.’ (Sayer 2000: 4-5)

Retroduction breaks with Popper’s hypothetical-deductive form of scientific inference (Sayer 1992: 169-174). Popper’s falsificationist criterion of science, according to which a theory must be at least potentially falsifiable by empirical observation, was famously illustrated with reference to the appearance of a single black swan. Observation of this event was deemed sufficient to falsify the
theory that ‘all swans are white.’ Some scholars distinguish ‘abduction’ as a fourth mode of scientific inference, involving the theoretical redescription of underlying structures and relationships (Danermark et al. 2002: 88-95), while others incorporate into accounts of the retroductive process (cf. Lawson 1997).

Recent studies pursuing a CR-inspired research agenda in this area (e.g. Blundel 2002, Bowey and Easton 2003), can be contrasted with those adopting the more ‘mainstream’ approach of inductive inference, based on the analysis of empirical regularities (i.e. the regression of variables derived from a questionnaire-based survey of owner-managers against concurrent firm-level measures of growth) (e.g. Adams and Hall 1993, Barkham et al. 1996).

Edith Penrose asserts her realist credentials in a later and more light-hearted comment that seems to anticipate more recent critiques of postmodernist positions:

‘Now none but the most philosophically sophisticated businessman will accept the proposition that the opportunities for the expansion of his firm are simply his ideas about what his firm can do; he will insist that the opportunities he sees reflect the “facts” of the world, facts that may be known with indifferent accuracy to be sure, but facts none the less’. (Penrose 1959: 216)

This long-standing case for greater integration is reflected in Penrose’s (1953) critique of early evolutionary theorising by economists in relation to the growth of firms. Her argument anticipates CR’s rejection of deterministic explanations, on the grounds that they tend to abstract away the essential interplay between human cognition, agency and their environment:

‘Once human will and motivation are recognized as important constituents of the situation, there is no a priori justification for assuming that firms, in their struggle for profits, will not attempt as much consciously to adapt the environment to their own purposes as to adapt themselves to the environment’. (Penrose 1953: 10)

Other recent network-related studies adopting elements of the CR paradigm include Neergaard (1999), where the author explores the role of networks in the internationalisation of small furniture manufacturers, and Blundel (2002), which addresses the interplay between inter-organisational networks and institutional-level changes in relation to contrasting growth processes in artisanal firms.

The realist orientation of this analysis is reflected in its capacity to probe intermediate mechanisms and context-specific interactions. For example, in Best’s (2001) account of regional growth dynamics in Northern Ireland, poor performance in innovation and productivity is traced across several levels of analysis, to reveal a lack of growth engines. By layering the analysis in a dynamic, open-systems framework, Best is able to make connections between these mechanisms (e.g. the long-run shortage of new entrepreneurial firms can be related to the limited development of technology management capabilities at a regional level.).

The shortage was evident from my own search for studies to illustrate this chapter, which yielded mostly implicit or quasi-realists. The editors of a leading CR text have reflected on this experience. While acknowledging that three (out of six) of their contributors illustrating contemporary realist practice ‘do not
make any explicit reference to realism’, they argue that in practice the field contains ‘much more work that is implicitly realist than that which is implicitly or explicitly postmodernist’. Fleetwood and Ackroyd (2000: 19)

14 The authors cite Bhaskar (1975) in support of the argument that mechanism-based explanations usually invoke some sort of ‘causal agent’ that generates an observed relationship (Hedström and Swedberg 1998: 11). However, the ‘social mechanisms’ school take a contrary position to CR on the operation of these mechanisms, arguing that social world phenomena must always be explained on the basis of individual actors. This fundamental difference is clarified in the following statement:

‘In the natural sciences, causal agents come in a variety of forms such as organic reactions in chemistry and natural selection in biology. In the social sciences, however, the elementary “causal agents” are always individual actors, and intelligible social science explanations should always include explicit references to the causes and consequences of their actions. This principle of methodological individualism is intimately linked to the core idea of the mechanism approach: Understanding is enhanced by making explicit the underlying generative mechanisms that link one state or event to another, and in the social sciences, actions constitute this link.’ (Hedström and Swedberg 1998: 11-12)

15 Archer (1995: 102) encapsulates her detailed critique of structuration theory by stating that it involves ‘sinking’ rather than ‘linking’ the essential differences between structure and agency (cf. Giddens 1984). By contrast, Archer’s procedure of ‘analytical dualism’ is based on the CR concept of emergence: social structures and human agency are different strata (hence ‘dualism’), whose interactions are only open to social scientific (hence ‘analytical’) inquiry. Danermark et al. (2002: 178-182) is a helpful summary of Archer’s arguments.