THREE ESSAYS ON RATIONALITY, INTENTIONALITY AND ECONOMIC AGENCY

A dissertation submitted by

PHILIP B. FAULKNER

SELWYN COLLEGE
UNIVERSITY OF CAMBRIDGE

for the degree of

DOCTOR OF PHILOSOPHY

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

The length of this dissertation is approximately 42,000 words (including bibliography). It therefore satisfies the word limit set by the Degree Committee for the Faculty of Economics and Politics.

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

The length of this dissertation is approximately 42,000 words (including bibliography). It therefore satisfies the word limit set by the Degree Committee for the Faculty of Economics and Politics.

ABSTRACT

The central theme of this dissertation is the contribution that the theory of human ontology developed by the philosopher John Searle can make to economics. Searle's account of the cognitive functioning associated with rational behaviour provides a framework within which to analyse the role of conscious and non-conscious factors in rational behaviour; the nature and functioning of discursive and tacit knowledge; and the distinction between intentional and non-intentional states. Using this framework, each of the three essays which make up the core of this dissertation examines aspects of the conception of economic agency associated with a different field in economics; critical realism, behavioural finance and mainstream microeconomics.

The first essay, which looks at the critical realist conception of the human actor in Tony Lawson's *Economics and Reality*, argues that Lawson leaves undeveloped the notion of tacit knowledge, failing to explain important differences between knowledge that functions by virtue of conscious reflection and that which functions tacitly. From a Searlean perspective the key omission is argued to be the technical notion of intentionality, upon which Searle develops an account of tacit knowledge. I show how this notion of intentionality evades my criticism of Lawson.

The second essay examines the conception of agency associated with behavioural finance from the perspective of the human ontology proposed by Searle. The principle theme of the essay is that each of the psychological traits that behavioural finance draws on, namely prospect theory, judgmental heuristics and mental accounting, involves the interplay of both conscious and non-conscious factors. Consequently the agent of behavioural finance is a construction that is readily intelligible in Searlean terms. I argue that this finding leads to a conception of the rationality of the agent encountered in the behavioural finance literature that is quite different from the way in which it is commonly presented.

The mainstream microeconomic conception of the human actor is the focus of the final essay, in which it is argued that the treatment of human knowledge on this approach neglects a number of important factors in economic behaviour. The first half of the essay uses a simple Cournot duopoly game under conditions of complete and incomplete information in order to highlight the usual assumptions about actors' knowledge in mainstream models. On the basis of these findings the second half of the essay then considers three aspects of human agency that these models neglect: non-probabilistic forms of uncertainty and ignorance, the subjectivity of knowledge and the role of tacit knowledge.

ACKNOWLEDGEMENTS

While I wish to thank everyone who has helped me in some way with this dissertation, I shall restrict my comments here to those people who have made the largest contribution. Foremost among these must be my supervisor, Jochen Runde, who for four and a half years has been unstinting in his help and guidance with respect to all aspects of the PhD. I am enormously grateful to Jochen for his efforts. I would also like to thank members of the Cambridge Social Ontology Group, both for providing a forum within which to discuss many of the ideas contained in this dissertation, as well as for encouraging non-mainstream work of this kind. I also gratefully acknowledge the financial support of the ESRC and a Rothschild / Newton Trust Bursary.

Next I would like to thank a group of people who made the day-to-day grind of PhD-life much more enjoyable than it otherwise might have been. I have been fortunate to get to know many PhD economists during my time in Cambridge, but I want to mention in particular Mike Biggs, Jihong Lee, Karsten Neuhoff, Andreas Pick and Sang Ryoo. Their willingness to take extended coffee breaks, to give (almost) equal weight to talk of economics and football, and above all to play poker, provided many memorable times.

Finally, I thank my family and especially Ingaret, my wife, for her love, encouragement and support in all things, including this dissertation. In agreeing to proof-read the entire manuscript while at home recovering from shoulder surgery, she undoubtedly acted above and beyond the call of duty, as well as greatly improving the dissertation. My first act in repaying this debt to her will be buying the long-dreamt about drinks to celebrate completion.

Philip Faulkner
1 April 2003
Cambridge

CONTENTS

INTRODUCTION

One of the most exciting developments in contemporary economics has been the move towards explicit ontological theorising, a move most clearly seen in the projects of critical realism (Archer *at al.* 1998; Fleetwood 1999; Lawson 1997, 2003) and economic ontology (Mäki 2001), as well as many heterodox traditions in economics[1]. In this dissertation I take the call to ontology seriously by adopting an explicitly ontological perspective to examining important aspects of the conceptions of human economic agency found in three prominent projects in economics: critical realism, behavioural finance and mainstream microeconomics. In particular my focus is on the mental ontology associated with the decision-making processes of agents in each of the projects under consideration, drawing on the philosopher of mind John Searle as a source of insight.

In order to outline my own project it will prove useful at the outset to distinguish between two different aspects of the so-called 'ontological turn' in economics. The first of these is concerned primarily with economic methodology, and employs ontological reasoning in order to ground arguments about the likely usefulness of particular practices or methods in economics. Important though it is, this aspect of the ontological turn, most closely associated with the project of critical realism in economics, is not the sort of project that I pursue in this dissertation. Rather, each of my essays is an example of a second aspect of the ontological turn, which involves the use of ontological arguments as an aid to the critical assessment and development of more substantive economic theories. The form that such arguments take is captured in the notion of 'ontological elaboration', the goal of which is to uncover the presuppositions about social being that economic theories (often implicitly) rely upon, and to compare these with the actual nature of social being.

A good deal has already been achieved under the heading of ontological elaboration. Contributors have examined the ontological presuppositions of various theoretical approaches to economics and attempted to elaborate many of the forms of social structure and related categories that are otherwise taken for granted in modern

[1] Including Austrian, Evolutionary, Feminist, Institutionalist, Marxist and Post-Keynesian writers.

economics[2]. One aspect of social reality that has so far been relatively under theorised however is the human agent. With the exception of Margaret Archer's work on the concepts of culture and identity (Archer 1995, 1996, 2000), and the more recent contributions of John Davis on collective intentionality (Davis 2002), Ralph Pfouts on the ontology of the consumer (Pfouts 2002) and Geoff Hodgson on the relationship between human agency and social structure (Hodgson 2002, 2003), many aspects of human ontology have still to be properly examined. In this dissertation I concentrate on one such area: the various mental processes involved in human economic decision-making.

The importance to economics of adequately conceptualising this aspect of human agency is evident from the fact that all economic phenomena depend in part on the actions, thoughts and conceptions of individual human actors. Consequently the way in which actors are portrayed – what they are presumed to know, what motivates them to act and how they decide between the various courses of action open to them – is a key component of most economic explanations. In the three essays that make up my dissertation I examine aspects of the mental ontology associated with the conceptions of human agency and decision-making found in three important projects in economics: critical realism, behavioural finance and mainstream microeconomics. My goal in each paper is first to elaborate the mental ontology associated with the particular conception under consideration and to then offer a critical assessment of that conception based on the ontology so elaborated.

In evaluating the mental ontology of each conception I draw extensively on findings from the philosophy of mind. Since philosophy of mind is concerned in part with studying the contents and characteristics of the human mind, that is, with understanding the nature of mental states, of conscious and non-conscious processes, of tacit and discursive knowledge, of rational behaviour, and so on, this literature provides just the sort of explicit ontological account of the mind that my project requires. For reasons that I discuss in section II below, I draw particularly on John Searle's account of the ontology of the human mind.

---

[2] Including evolutionary theory (Hodgson 1999a), modern macroeconomics (Hoover 2001) and microeconomics (Rosenberg 2001), the firm and region (C. Lawson 1999), metaphor (Lewis 1996), institutions (Hodgson 2003), the relationship between social structure and human agency (Hodgson 2002), money (Ingham 1996), markets (O'Neill 2001), transactions (Pratten 1997). See Lawson (2003) chapter 2, and pages 302 and 305, for many more examples of this sort of work.

I shall structure my introductory remarks as follows. First of all I provide a brief overview of each of the three papers included in the dissertation and reflect on the sort of insights that a project of this kind can generate. Next I defend my decision to concentrate exclusively on John Searle's philosophy of mind in each essay. While a broader project, taking in competing conceptions of human agency may well prove beneficial to economics, I argue that there are good reasons for beginning with Searle's work. Finally I examine how the project that I pursue here fits into historical and contemporary developments within economics. The feature that comes through most clearly here is the tendency for economists, past and present, to have neglected issues of philosophy of mind, in particular the nature of human mental ontology.

I THREE ESSAYS ON RATIONALITY, INTENTIONALITY AND ECONOMIC AGENCY

I begin, then, by reviewing each of the essays that make up this dissertation, partly as a means of introducing my work but also as a way of indicating the manner in which an ontological project of the sort pursued here can yield useful insights. In the first essay, 'Some problems with the conception of the human subject in critical realism', I examine the portrayal of the human actor in Tony Lawson's *Economics and Reality* (Lawson 1997). Lawson is a prominent contributor to critical realism, a methodologically oriented project concerned with establishing the possibilities for economics (and social science in general) by way of examining the nature of social reality. To this end Lawson develops a theory of social ontology, one element of which, his account of the human actor, provides the subject matter of this essay[3].

According to the social ontology Lawson develops in *Economics and Reality*, social reality is a structured, emergent and inherently dynamic realm consisting in social structures (such as rules, relations and positions) and human practices. While (by definition) social reality depends, in part, on human agency, Lawson rejects both

---

[3] Other contributors to critical realism defend a similar conception of social reality to Lawson (especially Bhaskar 1978, 1979). I focus on Lawson's work in this essay because he provides the clearest and most developed account of the human subject and also because his is the work that has had most impact on economics.

voluntarist and determinist accounts of social reality in favour of a transformational model that recognises social structure as both a condition and consequence of human action, and so social structure and human social practices as ontologically distinct categories.

Turning to the portrayal of the human agent within the transformational model, one of the most important aspects of social reality that Lawson must explain is how social rules play a causal role in human behaviour. Here the observation that many social practices are performed habitually, with rules being adhered to without being consciously followed, leads Lawson to argue that many social rules are known to us only tacitly. Yet the fact that in certain circumstances social rules are consciously followed requires Lawson to distinguish between discursive and tacit knowledge of rules, where the former requires conscious attention in order for the rules to influence behaviour, while tacit knowledge does not.

It is this aspect of Lawson's account of the human subject that is addressed by the first essay. In particular I argue that Lawson leaves the crucial notion of tacit knowledge under-developed, failing to adequately explain the ontological difference between knowledge that functions by virtue of being consciously thought about, and that which functions tacitly. Having outlined the nature of the problem in the first half of the essay, I then propose that John Searle's account of human ontology provides a possible resolution. The remainder of the essay provides an overview of Searle's conception of the human mind, showing how from the distinction between intentional and non-intentional states, Searle provides an account of the functioning of both consciously reflected-upon knowledge and tacitly functioning, 'Background', knowledge which evades my original criticism of Lawson.

The second essay, 'The human agent in behavioural finance: a Searlean perspective', puts forward a Searlean interpretation of the conception of human agency found in behavioural finance. The field of behavioural finance has developed rapidly over the last twenty years, based on introducing the findings of experimental psychology into economics in order to explain aspects of financial market behaviour. The resulting conception of agency that has emerged emphasises three types of psychological traits: prospect theory, judgmental heuristics and mental accounting. My main contention in this essay is that in certain key respects the agent of behavioural finance is a recognisably Searlean human agent.

The argument here centres on one of the dominant themes of Searlean human ontology, that the mental input into rational behaviour consists of both conscious, reflected upon, reasons for action, as well as non-conscious, predispositions and knowledge that function without conscious reflection. After briefly summarising this aspect of Searle's theory at the beginning of the essay, I then examine the three sets of traits emphasised in behavioural finance. Although the question of how these traits actually function is rarely addressed by the literature, all three appear to involve the interplay of both conscious and non-conscious factors. Viewed in this light, a Searlean ontology employing the distinction between intentional and non-intentional states, together with the notion of intentional and Background causation, is a natural framework within which to conceive of the kind of agent encountered in the behavioural finance literature.

This insight has important implications for the subject of behavioural finance. An interesting feature of this literature is a lack of consensus regarding how the human actor should be conceptualised. One position to take here is simply to regard the agent of behavioural finance as irrational. This view, common both inside and outside of behavioural finance, is well expressed by Shleifer, who regards behavioural finance as the placing of 'the biased, the stupid, and the confused into competitive financial markets' (Shleifer 2000: 24). The reasoning behind such judgements is rarely fully articulated, but mostly reflects the fact that the behavioural finance conception often involves behaviour that is in conflict with the axioms of the expected utility model of mainstream economics.

I suggest a different interpretation of the agent in behavioural finance. From a Searlean perspective the reduction of rational behaviour to expected utility maximisation is mistaken[4]. While the expected utility model emphasises the role of conscious beliefs and preferences, rational human behaviour according to Searle inevitably involves non-conscious as well as conscious factors. Many of these non-conscious factors, such as our tacit knowledge of simple heuristics, aid rather than hinder rational behaviour. The traits of the kind of agent that emerges in the behavioural finance literature therefore need not

---

[4] Searle himself expresses dissatisfaction with decision theoretic models of rationality in a number of places (e.g., 1995: 138-139; 2001: 30-31), highlighting in particular the implausibility of assuming that individuals will always be prepared to trade-off risk and return, and the neglect in these models for the role that deliberation plays in determining individual's preferences.

simply reflect the stupidity or confusion of Shleifer's interpretation; rather they may be characteristic of human rationality.

The significance of this difference, between viewing a model of agency as describing various irrationalities and viewing it as capturing various common traits of rational behaviour, has implications for the applicability and therefore importance of behavioural finance. If on the one hand these traits merely constitute mistakes and anomalies, then traditional finance based on the expected utility maximising individual has little to worry about. After all, such irrational individuals will either learn from their mistakes and stop making them, or be exploited financially by other, more rational actors (Russell 1997)[5]. But if on the other hand the traits in question are characteristic of rational human behaviour, inevitable features of actors trying to do the best for themselves, then such traits are a necessary part of studying human behaviour in financial markets.

The final essay, 'Information, Knowledge and Agency', examines some of the limitations of the mainstream microeconomic approach to the issues of knowledge, ignorance and uncertainty. The essay begins with a look at the nature and extent of the knowledge that agents are assumed to have in mainstream models. This is achieved by way of a careful examination of a representative mainstream model, in this case a simple one-shot Cournot duopoly game under conditions of both complete and incomplete information. I show that even in models designed explicitly to highlight the effects of imperfections and asymmetries in actors' knowledge, the degree to which the complexities of human knowledge are reflected in these models is severely limited. In the second half of the essay I consider in some detail three aspects of human agency that are neglected by the mainstream approach: non-probabilistic forms of uncertainty and ignorance, the subjectivity of knowledge, and tacit knowledge.

The main finding of the essay is that the much vaunted information theoretic revolution in microeconomics (Stiglitz 1987, 1992, 1994, 2000) represents only a first step towards incorporating the effects of factors such as uncertainty, ignorance and subjectivity into economics. For instance, by virtue of the commitment to expected utility models of decision-making, the actors within mainstream microeconomic models

---

[5] Although there are limits to such arbitrage arguments; for instance it can pay for 'rational' actors to act in the same way as 'irrational' actors rather than attempt to exploit them (Haltiwanger and Waldman 1985; Russell and Thaler 1985) and the supply of arbitrage is anyway likely to be limited (De Long, Shleifer, Summers and Waldman 1990).

inevitably suffer only certain forms of uncertainty. There is no scope in these models for actors to be affected by ignorance of the full set of possible eventualities that might result from their actions (or indeed the options open to them), or to alter their behaviour as a result of being unable to state precise probabilities. Consequently categories such as surprise and novelty, and forms of behaviour such as convention and routine, remain outside the scope of mainstream economics.

Elements of John Searle's human ontology provide the basis for my discussion of the two remaining areas of neglect highlighted in the essay, the subjectivity of knowledge and the role of tacit knowledge in human behaviour. Other than asymmetries in what actors are presumed to know, mainstream models make no allowance for the fact that much of what human beings come to know of their situation is subjective, depending on their existing situation, knowledge, interests, experiences and expectations. One cause of this subjectivity is what Searle terms the aspectual nature of conscious experience, according to which we unconsciously impose structure on, and interpret, all of our conscious experiences. The final issue dealt with in the essay, the importance of tacit knowledge in human behaviour, again draws on Searle, this time by considering the expected utility model to be an idealised form of Searle's model of intentional causation in which the role attributed to tacit, Background, factors is suppressed.

Taken together, these three essays begin to indicate the range of insights that follow from the explicit analysis of human mental ontology, in this instance drawing on the work of John Searle. In the first essay my subject matter, the conception of the human agent in critical realism, already involves an explicit account of the agent's cognitive functioning. Yet from a Searlean perspective this account is argued to be inadequate. In contrast, while sophisticated in terms of its psychology, the behavioural finance literature has little to say about the mental ontology associated with its conception of the human agent. This time a Searlean perspective provides an ontological framework within which to make sense of the various traits of this conception of agency, leading to the possibility that the existing understanding of behavioural finance as dealing with the effects of irrationalities and stupidity in financial markets is mistaken. The third and final essay attempts to draw out the implicit ontology associated with the mainstream microeconomic conception of human agency. Once this is done, the limitations of conceiving of agency in these terms becomes clear, since factors such as the subjectivity of knowledge and uncertainty cannot be adequately treated.

II JOHN SEARLE'S THEORY OF HUMAN ONTOLOGY

Having outlined the nature of the project pursued in this dissertation and the sort of results that such a project can yield, I now want to explain my decision to concentrate exclusively on the account of human ontology developed by John Searle[6]. In part this decision undoubtedly reflects my own interests and background at the time I began writing the dissertation. In particular, to someone with existing interests in critical realism, Searle's work on social and human ontology not only represents an important competing perspective, but also one that is much more detailed than critical realism in certain respects. Most obviously, having written extensively on the philosophy of mind, Searle's account of human ontology is far more developed than that typically found in critical realist accounts.

A second reason for focusing on Searle's work in these essays is that in my opinion Searle is one of the most important writers within the field of philosophy of mind for economists to get to grips with. The reason for this is that Searle takes seriously notions that economists, albeit often only implicitly, have come to rely on. As I argue in my second and third essays, the implicit ontology associated with the most prominent conceptions of the human agent in economics is one of conscious intentional states such as beliefs and desires[7] in association with unconscious, Background dispositions, in the case of behavioural finance. Given this, Searle's conception of the mind is particularly significant to economics for the simple reason that according to Searle (see especially Searle 1994a, 1997, 2002), conscious mental phenomena such as pains, feelings, visual experiences, beliefs, intentions, emotions and the like, exist and make a difference to (among other things) human behaviour. To put the point bluntly here, Searle's account of the mind is important to economics because it emphasises the same ontological categories that economists' theories of human behaviour rely upon.

Expressed in this way, it might appear that the position Searle defends ought not to be particularly controversial within the philosophy of mind, yet writers from a variety

---

[6] Beyond his work in philosophy of mind, Searle has written extensively on subjects including philosophy of language, the nature of social reality, the nature of science and causality. I will have relatively little to say about Searle's contributions to philosophy in these areas, but see Fotion (2000) and Hirstein (2001) for two very useful recent accounts.

[7] Although as I note in section III below, some writers deny even an implicit connection between rational choice models of decision-making and human mental ontology.

of different positions take issue with it. A brief survey confers some idea of what is at stake in these debates. Against Searle's view, logical behaviourists (Hempel 1949; Ryle 1949) and identity theorists (Lewis 1966; Place 1956; Smart 1959, 1988) deny the importance (and in the limit even the existence) of consciousness by reducing mental states to being nothing more than behaviour and states of the brain respectively. Similarly functionalists (Lycan 1987; Van Gulick 1995) replace mental terms with certain relations between causes and effects. According to the strong artificial intelligence doctrine (Chomsky 1986; Fodor 1975; Marr 1982; Newell 1982), the mind is to the brain as a computer program is to the computer hardware in which it is instantiated. Consequently what matters is not the various forms of consciousness that the mind might exhibit, but rather understanding the workings of the mind as a computer program. A final group of writers that downplay or deny the existence of consciousness are the eliminative materialists (Patricia Churchland 1981, 1986, 1987; Stich 1983) who, informed by modern neuroscience, seek to reduce mental states and consciousness to the processes in the brain that cause them.

What is common to each of these doctrines is a neglect of a theory of conscious mental phenomena. For these authors, an understanding of the workings of the mind does not require entities such as beliefs and desires[8]. Searle thinks differently. According to his 'biological naturalist' conception, mental phenomena are dependent upon lower-level, physical processes in the brain. But this causal dependence of the mind on physical processes in the brain does not imply an *ontological* reduction[9]. Rather, consciousness, conscious states, thought processes and the like are emergent, higher-level, features of the brain that have causal powers of their own. Consequently rather than downplaying their role in understanding the mind, Searle takes seriously the idea of uncovering the characteristics of the conscious mind and demonstrating how consciousness is to be

---

[8] Of course there is yet another group of writers, the dualists, who share Searle's view that mental phenomena are real but then reject the idea that they have physical causes. Substantival dualists, writing in the spirit of Descartes and the Christian tradition, argue that the mind and body are fundamentally different types of substance. Searle gives short shrift to this version of dualism, arguing that it is inconsistent with modern scientific knowledge. Property dualists (Nagel 1974, 1986; McGinn 1991, 1999) meanwhile hold that mental and physical properties exclude one another by definition, so that it is an oxymoron to talk of mental phenomena as also being physical phenomena.

[9] See Searle (1994a, chapter 5) for a full account of the notion of ontological reduction.

integrated into a theory of human action[10]. Consequently a Searlean-type conception of the mind is the natural starting-point for a project of the sort I pursue in this dissertation.

Although perhaps the most well known contemporary philosopher to advance a view of the mind as emergent but irreducible to physical processes in the brain, Searle is certainly not alone in doing so[11]. This is important, since Searle's approach and many of the details of his account have generated a substantial critical literature[12] that, it could be argued, cast doubt on the arguments I make in the essays that follow. Yet while acknowledging that debates in philosophy of mind are far from settled and that many aspects of Searle's account are contested, I rely not on the fine details of Searle's account, at which many of these criticisms are directed, but instead on the broader conception of the mind that Searle's account is one particularly clear illustration of.

One final comment on the Searlean nature of my project. In choosing to focus exclusively on John Searle's account of human mental ontology, I do not mean to imply that there is little value in competing developments in contemporary philosophy of mind for economics. An example which competes with Searle's 'belief-first' conception of action is the pragmatist Hans Joas' (1993, 1996) account of human action as 'creative action'. Given the recent re-emergence of pragmatism on the philosophical landscape, Joas is a particularly pertinent example of a potentially important source for economists. Now of course there is at least one connection between pragmatist philosophy and economics that is already well established, that is the link between institutionalist writers such as Ayres, Veblen, Commons and Mitchell, and the classical pragmatism of Pierce,

---

[10] For instance, and as Searle himself observes (1994: 250), despite often putting the word 'consciousness' in the titles of their books, many writers on mind proceed to give little or no account of the nature of consciousness. In contrast, providing an account of subjective, qualitative, first-person conscious mental states is central to Searle's attempt to understand the mind and its functioning (see especially Searle 1994, 2002).

[11] I would nevertheless suggest that Searle is the most concerned with developing an explicit account of the characteristics of the conscious mind and their role in the functioning of the mind in behaviour. Among philosophers, Bindra (1976), Bhaskar (1978, 1979), Ardila and Bunge (1987), Bunge (1980), Davidson (1980) and Hebb (1949) also see the mind as a set of emergent brain functions. Notably, Searle's non-eliminative, materialist, view of consciousness also finds support within contemporary neuroscience (e.g. Crick 1994; Edelman 1989, 1992).

[12] The very tip of the iceberg here being the collections edited by Dietrich (1994), Lepore and van Gulick (1991) Preston and Bishop (2002) and the recent special issues of *Revue Internationale de Philosophie* (June 2001) and *The American Journal of Economics and Sociology* (January 2003).

James and Dewey[13]. I focus here instead on the more recent work of Joas and the important contrast between his conception of action and that of Searle.

For Joas, both the philosophy of action (including Searle) and the social sciences have neglected what he terms the 'creative' character of human action. By conceiving of cognition as independent of action, with actions following after a separate act of cognition concerning our situation, possibilities and wants, Joas argues that belief-first models of action encourage the view that our natural state is one of lethargy. Instead, Joas emphasises the idea that our natural state is really one of continuous action, with actors continually having to confront problems and act whether they want to or not. Accordingly, rather than preceding action, cognition is better conceived of as itself being a phase of action by which action is directed and redirected in its situational contexts.

Without the creative dimension of action taken into account, Joas argues that a wealth of notions involved in the concept of action, such as intention, norm, identity, role, institution, routine etc., cannot be defined consistently and in a way that does justice to what they are intended to express. He notes, for instance, that the role of situation in action immediately changes from merely being a constraint on action to being partly constitutive of action, shaping our goals and motives for acting[14]. Similarly our understanding of the notion of goal setting changes. Rather than taking place prior to actual action, goal setting is that part of action where the aspirations embodied in the body's capabilities, habits and dispositions are reflected upon and thematised. Now since the Searlean conception of action largely ignores the inter-relatedness of action and cognition, in general seeing cognition as proceeding action, these sort of insights are unlikely to be introduced into economics by a project that focuses exclusively on Searle. But this observation merely confirms the point I wish to emphasise, which is that there are undoubtedly more insights to introduce into economics from philosophy of mind than those that emerge in the essays in this dissertation.

---

[13] On which see Hands (2001), Hodgson (1998) and Hodgson *et al.* (1994).

[14] See Hodgson (2003) for a closely related argument.

III ECONOMICS AND PHILOSOPHY OF MIND

To this point I have indicated why I believe that economists can gain by paying attention to issues of human mental ontology, and secondly, why it is that John Searle is a particularly significant writer in this regard within philosophy of mind. My remaining comments in this introduction are directed at situating my dissertation within historical and contemporary developments in economics. Broadly, there are three aspects to this that I wish to consider. The first is the tendency for issues of human mental ontology to be neglected within neoclassical economics, where the concept of the utility maximising rational agent dominates discussion of agency, action and decision-making. The second is the work of a number of important heterodox writers on human agency, examining the extent to which mental ontology features explicitly in their work. The third is the isolated set of instances of economists who have brought wider issues of philosophy of mind into the discipline, and in particular John Davis' work on collective intentionality and Alexander Rosenberg's critique of economists' attachment to folk psychology.

Human Ontology in Neoclassical Economics

As the essays that make up this dissertation demonstrate, an explicit reflection on human mental ontology has significant implications for how we conceive of human action and decision-making, raising vital questions concerning what individuals know, how they acquire this knowledge, how they make decisions, what motivates their actions, and so on. That neoclassical[15] economics has neglected, and continues to neglect these sorts of questions, and in so doing fails to address the underlying ontological issues, is hardly a new observation. After all, it provides a starting-point for many heterodox writers, some of whom I discuss later[16]. First of all though I want to be clear about precisely what is the problem in this regard with the neoclassical approach to the human agent. To do this, let me begin by identifying my target. The defining characteristic of the neoclassical model

---

[15] By neoclassical economics I mean the dominant, 'textbook', approach to economics. This approach can usefully be characterised in terms of it's 'core' assumptions, namely, that individuals are utility maximisers with stable preferences; that analysis focuses on equilibrium states; and that chronic information problems and uncertainty do not exist.

[16] Relatedly, see Lewin (1996) for a historical perspective on economist's antipathy towards psychology and the psychological foundations of rational choice theory.

of economic decision-making, and consequently its conception of the human agent, is the idea firstly, that individuals act rationally, and secondly, that rational behaviour can be represented as action that maximises some single-valued utility (or expected utility) function.

Now what interests me here about this conception of the human agent is the mental ontology associated with such 'rational choice' models. According to the standard interpretation[17] of these models, the utility function is viewed as representing an individual's subjective, introspective, evaluation of the objects of choice, with utility maximisation equivalent to the individual selecting the most preferred available alternative[18]. If such an interpretation is accepted, then rational choice models evidently make certain commitments in terms of human mental ontology. In particular, individuals must be assumed to possess well-behaved subjective preferences[19], to know the set of feasible actions and the consequences of these actions, and be able to compute their optimal choice. Now the implausibility of these assumptions is widely recognised and related themes, regarding the lack of information and knowledge, and the limited computational power of the typical human being, recur throughout this dissertation. However what I want to emphasise at this point is that these ontological commitments are also highly limited in the sense that a great deal is left unexplained. For instance, the content of an agent's preferences at the level of mental states, the form taken by an agent's knowledge of his or her environment, the manner in which preferences and knowledge function causally, and so on, are not explicitly addressed. Consequently many aspects of rational choice are unclear; whether, for instance, 'preferences' include both conscious and unconscious motives, or alternatively whether rational choice models only represent the influence of conscious, premeditated, calculation on actions.

In practice, many neoclassical economists seem content to put such questions to one side, preferring instead to get on with using the utility maximisation representation, rather than questioning quite what it is to which such a methodology commits them. Yet

---

[17] By 'standard' interpretation I mean that which is most common in modern neoclassical thought. Significantly, what I argue here to be the standard interpretation is also the typical presentation given in undergraduate and postgraduate microeconomic textbooks, for instance see Gravelle and Rees (1992), Mas-Colell, Whinston and Green (1995), Varian (1992, 2002).

[18] See Robbins (1937) and Hicks (1946).

[19] With preferences having to be assumed to conform to certain 'axioms of rational choice' in order to ensure the existence of the utility function representation.

there is also an important literature that has sought to overcome these issues by detaching rational choice models from any such ontological questions[20]. Two alternative interpretations of rational choice are particularly prominent here. The first, based on Samuelson's (1938) notion of revealed preference, argues that 'preferences' are in fact revealed in the observable choices that individuals make. Consequently rational choice models need make no commitments at all in terms of human ontology (to mental states, introspection, rational calculation and the like), but rather can sustain a theory of choice based entirely on consistent behaviour (rather than consistent preferences)[21]. A second interpretation is based on the evolutionary arguments put forward by Alchian (1950) and Friedman (1953). Here the idea is that utility maximisation contributes to an individual's evolutionary 'fitness' and that over time, a process of 'natural selection' promotes utility maximising behaviour within a population. On this interpretation questions of how individuals actually maximise utility are redundant, since however it is achieved, such individuals come to dominate.

While each of these alternative interpretations of rational choice models is highly problematic in its own right[22], the crucial point from my perspective is that on all of these accounts human mental ontology is neglected. Consequently whatever their implicit understanding of the nature of rational choice and the rationale for conceiving of the human agent in terms of utility maximisation, the work of neoclassical economists is characterised by a lack of concern for the nature of the human mind and the implications of this for human decision-making and action. Interestingly this neglect continues in the vast contemporary literature on decision-making[23], much of it in a neoclassical vein, which purports to develop more realistic models of utility maximisation.

---

[20] Indeed as Blaug and others have noted, the development of modern consumer theory from 'the introspective cardinalism of Jevons, Menger, Walras, and Marshall, to the introspective ordinalism of Slutsky, Allen, and Hicks, to the behaviourist ordinalism of Samuelson, to the behaviourist cardinalism of von-Neumann and Morgenstern' (Blaug 1992: 139-140), can be understood in terms of the desire to detach utility theory from unobservable notions such as individual introspection (and mental ontology in general).

[21] See for instance Little (1949).

[22] See in particular chapter 4 of Hodgson (1988) for a detailed discussion of the problems, as well as footnote 2 of my second essay. Significantly, these weaknesses lead Hodgson to argue that in the end, rational choice models do rely on an implicit (and incomplete) ontology of introspection, mental states and so on.

[23] See for instance Arrow et al (1996) and Camerer (1995) for recent surveys of this literature. If writers associated with critical realism and the ontological turn in economics are correct, then this continued

Human Ontology in Heterodox Economic Thought

Away from neoclassical, rational choice based, conceptions of human agency, the degree to which the various heterodox traditions in economics address issues of human ontology explicitly is mixed. In this section I intend briefly to illustrate this point by examining aspects of the work of a number of important heterodox writers and traditions. My aim in doing this is not to provide a complete survey of the treatment of human mental ontology in heterodox economic thought; such a task exceeds the scope of these comments. Rather my more modest goal is to highlight a number of the most important themes in this area and in so doing, continue to set out the context within which my dissertation emerges[24].

I want to begin these comments by considering the behavioural economics tradition, precipitated by the work of Herbert Simon on bounded rationality and the 'administrative man' conception of the economic agent and continued in the work of contemporary writers such as George Loewenstein, Daniel Kahneman, Matthew Rabin, Robert Shiller, Amos Tversky and Richard Thaler[25]. In terms of its present prominence, behavioural economics is perhaps the most significant of the heterodox traditions for me to address in these comments. Furthermore, since the defining feature of behavioural economics is a psychologically informed, descriptive account of human decision-making, one might expect issues of mental ontology to have been a central concern for behavioural writers. Yet one of the key themes of my second essay is that this has not been the case. Just as neoclassical accounts of decision-making leave unclear the mental ontology associated with utility maximisation, so too contemporary behavioural accounts leave the mental ontology underlying such notions as heuristics, mental accounting and loss aversion unexplained. As a result, while modern behavioural accounts of decision-

---

neglect of ontological considerations is closely related to the commitment of many economists to mathematical modelling. Certainly the course taken by the literature on decision-making would appear to support such an argument, where the emphasis is on generating ever more sophisticated formal models of utility maximisation rather than reflecting on the underlying mental ontology that these models represent. See, for instance, rank dependant utility theory (Diecidue and Wakker 2001; Quiggin 1982, 1993), configural weight models (Birnbaum, Patton and Lott 1999), stochastic difference models (Gonzalez-Vallejo 2002), multialternative decision field theory (Roe, Busemeyer, Townsend 2001), the componential-context model (Tversky and Simonson 1993) and range frequency theory (Parducci, 1965, 1974).

[24] See Rizzello (1997) for a closely related project.

[25] For instance, see Benartzi and Thaler (2001); Frederick, Loewenstein and O'Donoghue (2002); Gilovich, Griffin and Kahneman (2002); Lamont and Thaler (2003); Loewenstein (2001); O'Donoghue and Rabin (2000); Rabin (2002a, 2002b); Read, Loewenstein and Rabin (1999); Shiller (2003) and Thaler (2000);

making are more realistic in certain respects than the traditional rational choice models of neoclassical economics, they retain many of the latter's ontological failings.

Significantly, this tendency to neglect important issues of human mental ontology is also visible in earlier work within the behavioural tradition, most notably Herbert Simon's development of the notion of bounded rationality (Simon 1947, 1955, 1956)[26]. Consequently Simon's work nicely illustrates my criticism of behavioural accounts more generally. In describing the 'administrative man' found in actual economic situations, as opposed to the rational economic man of neoclassical theory, Simon observes that economic rationality rests on a number of highly unrealistic assumptions. He emphasises, for instance, that rather than possessing complete knowledge of all the alternatives open to them and the possible consequences of each alternative, individuals typically have to seek out available actions and on doing so know only imperfectly the possible consequences of these actions, because of both uncertainty and limited computational power. Furthermore, individuals generally have no simple way of comparing heterogeneous consequences, in contrast to the consistent utility measure of rational choice theory.

In his early work, Simon (1947) describes the kind of procedures that agents employ when faced with having to act in such circumstances. One such procedure is to look for satisfactory choices instead of optimal ones; another to replace abstract, global goals with tangible sub-goals, whose achievement can be observed and measured; a third to divide up the decision-making problem among a group, with their efforts coordinated by means of a structure of communications and authority relations. In subsequent work Simon then set about formalising these findings into the 'administrative man' conception of the human agent, describing the sort of rationality exhibited by such agents as 'bounded' rationality (Simon 1955, 1956). Two concepts are central to Simon's formal theory of bounded rationality: search and satisficing. In contrast to marginalist theories of search (e.g. Stigler 1961), Simon proposes that a decision-maker forms some aspiration as to how good an alternative he should find. As soon as an alternative that meets this level of aspiration is discovered, the search process, guided by computationally simple rules of thumb, ends and that alternative is chosen. Simon calls this mode of choice 'satisficing',

---

[26] See Augier (2001), Augier and March (2001, 2002) and Earl (2002) for recent surveys of Simon's contributions to economics.

capturing the idea that an individual seeks a satisfactory, rather than the best, available alternative.

Now the reason for citing Simon's work as a good illustration of the ontological deficiencies of the behavioural approach is the fact that Simon fails to explain how the processes involved in bounded rationality function at the mental level. In particular, Simon provides no account of how the processes of satisficing and search fit in with the hierarchical nature of consciousness and as a result, no explanation of the role of (for example) unconscious influences on behaviour, or the significance of differences between different kinds of knowledge[27]. Geoff Hodgson (1988: 100) argues that since bounded rationality still ultimately rests on a means-ends calculation, albeit a boundedly rational calculation, this would seem to imply that the processes involved in search and satisficing take place in the discursively conscious mind. Whatever is the case here, however, the important point is that while Simon's ideas represent an important advance over rational choice models of decision-making, they, like modern behavioural theory, are found wanting because of their neglect of human ontology.

Before I move on from behavioural economics, it is interesting to note two features of Simon's work that may in part explain his just-noted neglect of important aspects of mental ontology. The first of these is his commitment to empiricism[28], a methodology that emphasises the need to formalise theories and evaluate them according to empirical testing. From such a perspective the observability and measurability of phenomena is crucial, presenting obvious difficulties for providing explanations in terms of subjective, qualitative, conscious mental states and unconscious processes. A second, related observation comes from Simon's contributions to the field of cognitive science (Newell and Simon 1972; Simon 1964, 1966)[29], where he writes from a (symbolic) artificial intelligence perspective. On this view human reasoning and thinking are to be understood in terms of information processing, a consequence of which (as I noted

---

[27] However as Hodgson (1988: 287) notes, away from his formal theory of bounded rationality Simon does refer (albeit inconsistently) to the unconscious realm. Furthermore, in his brief discussion of Hayek's work on mind, Simon (1992: 26-27) clearly demonstrates his awareness of notions of tacit knowledge and subconscious know-how.

[28] For instance, see Simon's Nobel lecture (Simon 1977). The empiricism of behavioural economics continues to the present day, one sign of which is the close connection between it and experimental economics.

[29] On which see Hands (2001), chapter 4, and Sent (1997).

earlier) is that the importance of categories such as consciousness and intentionality tend to be downgraded.

Looking beyond behavioural economics, which I argued to be highly significant in this context because of the present prominence of behavioural accounts of decision-making and agency, a second writer who warrants some discussion here is Friedrich Hayek. Hayek's work in economics, political philosophy and epistemology repeatedly raises issues of human ontology[30] and a number of these themes are found in the essays in this dissertation. To take just one example, Fleetwood (1995, 1997) demonstrates that Hayek's work exhibits an increasing awareness of the concept of tacit knowledge. In his early work, while aware of the distinction between knowledge in the form of facts and knowledge in the form of skills (Hayek 1937: 50), and the existence of knowledge embedded in rules, practices, habits and institutions (Hayek 1945: 528), Fleetwood shows that Hayek initially makes no use of these notions. In 1952 Hayek then demonstrates his awareness of Ryle's distinction between 'knowing how' and 'knowing that' (Hayek 1952: 39), but it is not until the 1960s that Hayek begins to use Ryle's and Polanyi's insights in his own work (e.g. Hayek 1960, 1967, 1988). One of the most important uses here is in Hayek's account of the functioning of the social rules of conduct (Hayek 1960, 1973, 1983) that play a crucial role in his account of spontaneous order by making individual behaviour relatively predictable.

Hayek's work on these more substantive topics is clearly of interest given that these are issues that I deal with throughout my dissertation. Yet what makes Hayek particularly interesting here is that, in addition to addressing issues of human ontology in his economics, Hayek also wrote explicitly on the mind, most notably in his book on theoretical psychology *The Sensory Order* (hereafter referred to as SO). Although finally published in 1952, many of the ideas contained in the book date back to a paper written by Hayek in 1921. The book comprises Hayek's criticism of the psychology of perception and the philosophy of mind from before 1918, both of which were dominated by the ideas of Ernst Mach[31], plus Hayek's objections to more recent work in behaviourism and Gestalt theory. Hayek's starting-point is Mach's universe of sensations, which he terms the sensory order. The book's main thesis is that there are actually two 'orders' in which

---

[30] On the relation of Hayek's work on mind and human agency to that of other Austrian writers, see Boettke (1994), Cubeddy (1993), Hodgson (1988), Oakley (1997, 1999) and Runde (1988).

[31] See de Vries (1994) for a more detailed account of the relationship between Hayek's and Mach's ideas.

we arrange the objects in the world: the physical order, which classifies events as similar or different according to whether they produce similar or different other external events; and the sensory order, which classifies events according to their sensory properties. Contrary to Mach, Hayek argues that there is no simple one-to-one correspondence between the elements of the two orders; events which appear to our senses to be of the same kind may be different in the physical order, and vice versa. The relation between the sensory order of the human mind and the physical order of external events is the central problem that Hayek addresses in SO.

Of primary interest to me here is the conception of mind that Hayek develops in SO and the relation between this conception and the rest of Hayek's work, notably his economics. As Smith (1997) shows, Hayek's model of the mind is a 'connectionist' one, whereby all attributes of mental experience are explained in terms of, and are reduced to being nothing more than, the system of connections of corresponding patterns of nerve-excitations[32]. Now in advancing such an account, certain characteristics of connectionist models of mind are found in Hayek's own account. The first of these is that a connectionist stance encourages a view of human reason that emphasises its creative and contextual aspects, as well as the importance of non-propositional forms of knowledge (Dreyfus and Dreyfus 1988, Smith 1997). This is clearly in keeping with Hayek's awareness, and subsequent use, of tacit knowledge in his later economics, and as Birner (1994) notes, we can also sees aspects of Hayek's work on subjectivism, complex systems, the limitations of knowledge and evolutionary arguments in SO.

Yet Hayek's work on the mind in SO is not a perfect fit with his work elsewhere. In particular, by virtue of its eliminativist stance towards conscious experience, Hayek's connectionism, as with connectionism in general, is weak in its account of consciousness and deliberate, purposeful, conscious thinking (see Smith 1997; Weimer 1979, 1982). For instance, Smith (1997) notes the difficulty of dealing with mental causation in a connectionist framework[33], meaning the relation between reason, choice and action, which is central to Hayek's methodological individualism in economics. Similarly Hayek's account puts little emphasis on the nature of consciousness: its unity, qualitativeness, relative stability over time and so on. The overall picture, then, is mixed.

---

[32] See also Agonito (1975), Birner (1999) and Dempsey (1996) on Hayek's conception of mind.

[33] Hodgson (1988: 114-5) makes the related observation that in SO Hayek downgrades concepts such as purpose and action, in a way that goes against his (and other Austrians) use of these terms elsewhere.

While Hayek addresses issues of human ontology in his work in economics, and also writes explicitly on human ontology elsewhere, these two strands in his work are not wholly consistent, due largely to his commitment to a connectionist account of the mind.

.The two examples considered so far represent two extremes in the treatment of mental ontology in heterodox economic thought. The behavioural tradition largely neglects such issues, whereas Hayek writes explicitly but inconsistently on the mind. Between these extremes it is difficult to generalise. Many heterodox writers are explicitly concerned with decision-making and the inadequacy of the dominant, neoclassical conception. Yet often this concern, while correctly highlighting limitations with rational choice models, repeats the neoclassical neglect of ontology.

George Shackle's work on decision-making under uncertainty (Shackle 1949, 1969, 1970, 1983) is a case in point. Shackle takes as his starting-point the distinction between risk and uncertainty; for Shackle the use of probabilities (whether objective or subjective) to indicate the reliability of occurrence of specified outcomes is inappropriate where an individual has to choose an action in a situation of uncertainty. Situations of uncertainty are those in which an individual faces a 'crucial' decision, the making of which will, by definition, destroy the possibility of making that decision again. Since many, or most, significant economic decisions involve uncertainty rather than risk, rational choice models are inappropriate in economics. In proposing an alternative theory of decision-making that incorporates this insight, Shackle provides a much stronger account of the roles that time, uncertainty and expectations play in human agency. Yet as various commentators have noted (e.g., Ford 1994: 3; Oakley 2002: 154), Shackle's account makes little reference to the related psychological and philosophical literature on the mind.

In examining some of the most significant aspects of the treatment of mental ontology in heterodox economics, I have so far dwelt on largely negative themes, whether it is the absence of an explicit mental ontology or the presence of competing accounts of the mind. Now though I want to conclude this section by highlighting two instances of contemporary heterodox writers who have sought to bring issues of mental ontology to the fore. The first is Geoff Hodgson's work on habits and the second, Jon Elster's work

on emotion[34]. While writing from quite different perspectives, what each of these projects shares is the attempt to take seriously some otherwise neglected aspect of human mental ontology.

Geoff Hodgson's work on the concept of habit is a central feature of his contribution to the revival of the old institutionalist tradition in contemporary heterodox economics. For Hodgson, as for the founders of old institutionalism such as Commons, Veblen and Mitchell, habits are crucial to any portrayal of the human agent[35]. In emphasising the importance of habits, defined by Hodgson as 'a largely non-deliberative and self-actuating propensity to engage in a previously adopted pattern of behaviour' (1998: 178), Hodgson's point is not simply that the actions of human beings are often performed skilfully but with little or no conscious reflection. While undoubtedly true, if this alone were the significance of the concept of habit then the neglect of the category in mainstream economics might not be terribly important[36]. Rather as Hodgson demonstrates, the implication of recognising the role of habits can be to change radically how we conceive of categories such as human action, knowledge and institutions[37].

Perhaps the most significant aspect of Hodgson's work on habit for the project that I pursue here, especially given my earlier comments on both neoclassical and heterodox writers, is that he deals with the issue of how the particular category of interest (in this case habit) fits in with other important aspects of human ontology. In somewhat earlier work, Hodgson provides a critique of what he calls 'rationalist' conceptions of action that portray all action as the product of rational calculation (Hodgson, 1988).

---

[34] In fact, a third example of a heterodox project that deals explicitly with issues of human mental ontology has already been mentioned in this introduction, namely critical realism. Since this is considered at length in my first essay, I will not include a further summary of the project and its concerns here.

[35] In this, and in their conception of human agency in general, the founders of old institutionalism drew on the work of the instinct philosophers such as James and McDougall, and pragmatist philosophers such as Pierce. The revival in pragmatist thought has already been commented on in these introductory remarks, and it is therefore interesting to note a corresponding revival in instinct psychology, seen in the work of authors such as Cosmides and Tooby (1994a, 1994b), Plotkin (1994) and Reber 1993.

[36] Where mainstream economists have addressed the idea of habits, the tendency has been to regard them in rational choice terms, either by viewing them as the product of an earlier, ration, decision, or as a deliberate means of simplifying rational decision making (e.g. Becker 1992; Pollak 1970).

[37] For instance, on the connection between habit and accounts of institutions and social structure, see Hodgson (1998, 2002, 2003); on the role of habits and rule following in rational choice, see Hodgson (1997).

Drawing on various kinds of evidence, including philosophical and psychological arguments, Hodgson argues in favour of a hierarchical conception of the conscious mind that recognises the unconscious influences on action (such as habits) as well as the conscious, calculated, motives. Subsequently (Hodgson 1998) argues for the need to reconcile the purposeful, novel and creative aspects of human behaviour with the role of habits, rather than being forced to downgrade the significance of one or other of these aspects.

My second illustration of a contemporary writer in economics focusing explicitly on issues of human ontology is Jon Elster, whose most recent work examines the role that emotion plays in human behaviour (Elster 1998, 1999). As with the concept of habit, emotions have largely been neglected in economics, save for isolated appearances in response to specific issues (e.g., Rabin 1993; Gifford 2002; Kaufman 1999)[38]. Yet emotion would seem to play a central role in many areas of behaviour to which economists have traditionally applied rational choice models, such as decisions concerning savings and investment, gambling, insurance, bargaining, altruism and so on. Crucially, Elster argues that the role that emotions play in influencing behaviour is inadequately conceived of if they are reduced either to 'psychic costs' and 'benefits', or otherwise to sources of preferences. Rather, emotions such as anger, fear, love and the like, also affect an individual's ability to make rational choices in the first place. Therefore an adequate conception of the human agent in economics requires an ontology of the human mind that explicitly incorporates emotion.

Philosophy of Mind in Economics

In examining the situation within which my dissertation emerges, I have so far restricted my attention to the treatment in economics of issues concerned directly with human mental ontology. I now want to broaden my focus by highlighting two instances of wider issues of philosophy of mind that have been brought into economics; firstly the work of John Davis on the notion of collective intentionality, and secondly, Alexander Rosenberg's critique of folk psychology.

Davis (2002) draws on recent philosophical work on the notion of collective intentionality (Bratman 1993, 1999, Gilbert 1989; Tuomela 1991, 1995; Searle 1990,

---

[38] See also Loewenstein (2000).

1995) in order to support heterodox conceptions of agency that emphasise the 'social embeddedness' of the individual agent[39]. In contrast to the atomistic accounts of mainstream economics, in which agents are free of significant social attachments when they act, heterodox accounts stress the importance of social factors on individual activity. Yet for heterodox writers to sustain such a view requires that they demonstrate a difference between the way that individuals act when they are part of an organised group (i.e. in Davis's terms when they act socially), as opposed to when they act alone (i.e. non-socially). Davis seeks to understand this difference in terms of individuals acting in groups as having 'shared' or 'we' intentions about the groups of which they are members, using the notion of collective intentionality to sustain the idea of a 'shared' intention.

While the distinction between social and non-social actions seems relatively clear in terms of ordinary language, with people using 'I' intentions when they see themselves as acting independently of others and 'we' intentions when focusing on their actions as members of groups, the distinction has only recently found expression in the philosophical literature on collective intentionality. In this literature, 'we' intentions are intentions that individuals attribute, and believe that others attribute, to the individuals that make up the group of which they are members. When acting in a social setting each individual is therefore conceived of as possessing some form of 'we' intention (for instance, 'we have decided to adopt a following policy', or 'we are going on this trip', etc.) while believing that each member of the group possesses a similar attitude[40]. Therefore on this conception social activity is explained in terms of 'structures of reciprocally related we-intentions possessing a mutually reinforcing character. These structures of interaction...function as the skeletons of social activity and therefore of the institutions and social values that are the products of social activity' (Davis 2002: 15).

A second, quite, different engagement with issues of philosophy of mind in economics is found in the work of Alexander Rosenberg (1992, 1995, 1998 and elsewhere). In rejecting the scientific status of modern economics on the grounds of its failure to generate increasingly accurate predictions, Rosenberg argues that the cause of

---

[39] Two prominent examples of which have already been discussed: the critical realist theory of social ontology (Archer 1995; Bhaskar 1979; Lawson 1997) and the (old) institutionalist approach of Hodgson (2002).

[40] Notice that such 'shared' intentions are non-holistic; it is the individuals that make up the group that each hold a 'we' intention, not the group that has an intention over and above the intentions of the individuals that make it up.

1995) in order to support heterodox conceptions of agency that emphasise the 'social embeddedness' of the individual agent[39]. In contrast to the atomistic accounts of mainstream economics, in which agents are free of significant social attachments when they act, heterodox accounts stress the importance of social factors on individual activity. Yet for heterodox writers to sustain such a view requires that they demonstrate a difference between the way that individuals act when they are part of an organised group (i.e. in Davis's terms when they act socially), as opposed to when they act alone (i.e. non-socially). Davis seeks to understand this difference in terms of individuals acting in groups as having 'shared' or 'we' intentions about the groups of which they are members, using the notion of collective intentionality to sustain the idea of a 'shared' intention.

While the distinction between social and non-social actions seems relatively clear in terms of ordinary language, with people using 'I' intentions when they see themselves as acting independently of others and 'we' intentions when focusing on their actions as members of groups, the distinction has only recently found expression in the philosophical literature on collective intentionality. In this literature, 'we' intentions are intentions that individuals attribute, and believe that others attribute, to the individuals that make up the group of which they are members. When acting in a social setting each individual is therefore conceived of as possessing some form of 'we' intention (for instance, 'we have decided to adopt a following policy', or 'we are going on this trip', etc.) while believing that each member of the group possesses a similar attitude[40]. Therefore on this conception social activity is explained in terms of 'structures of reciprocally related we-intentions possessing a mutually reinforcing character. These structures of interaction...function as the skeletons of social activity and therefore of the institutions and social values that are the products of social activity' (Davis 2002: 15).

A second, quite, different engagement with issues of philosophy of mind in economics is found in the work of Alexander Rosenberg (1992, 1995, 1998 and elsewhere). In rejecting the scientific status of modern economics on the grounds of its failure to generate increasingly accurate predictions, Rosenberg argues that the cause of

---

[39] Two prominent examples of which have already been discussed: the critical realist theory of social ontology (Archer 1995; Bhaskar 1979; Lawson 1997) and the (old) institutionalist approach of Hodgson (2002).

[40] Notice that such 'shared' intentions are non-holistic; it is the individuals that make up the group that each hold a 'we' intention, not the group that has an intention over and above the intentions of the individuals that make it up.

this failure can be traced to economists' continued commitment to intentionalistic explanations of human action. By explaining action in terms of the preferences and expectations of actors, economists' theories of rational choice employ what philosophers term 'folk psychology', and this, according to Rosenberg, has been shown to be problematic in light of recent philosophical work.

Now I do not intend here to examine the broader context within which Rosenberg makes these arguments, in particular their relation to his earlier work or the validity of the empiricist philosophy of science that they embody[41]. Instead I want to briefly outline some of the philosophical arguments that Rosenberg draws on in order to claim that it is a commitment to folk psychology that causes economics to be inaccurate and vague in its predictions. The first of these is a problem of potential circularity in intentionalistic explanations of action. According to folk psychology, beliefs and desires provide an agent with reasons for action, and those reasons in turn provide the causal explanation for the agent's action. Now in order to employ this sort of intentionalistic explanation of an agent's action, we must know that agent's beliefs. The most obvious way to find out what these are is to ask the agent what they believe. Yet in order to connect an agents reported beliefs to their 'actual beliefs' requires an intentionalistic explanation of their speech act. The result is a potential regress, as any reported intentional states stand in need of interpretation in intentional terms.

This first line of reasoning has been powerfully challenged by Cottrell (1995) who, distinguishing both a strong and weak version of Rosenberg's argument, finds a defence against both in the form of Dennett's (1987) intentional systems theory. The second criticism that Rosenberg levels at folk psychology explanations is taken from philosophers who hold an eliminative materialist conception of the mind. Such writers (Stich 1983, Patricia Churchland 1986, 1987; Paul Churchland 1984, 1992) take folk psychology to be an unscientific explanatory model of action because the intentional vocabulary of beliefs and desires refers to things that they consider simply not to exist. Consequently, or so the argument goes, for as long as economics continues to seek to explain behaviour in intentional terms, rather than in terms of modern neuroscience, it will remain inaccurate and vague in its predictions. Yet such a position requires an acceptance of an eliminative materialist view of the mind, something that philosophers like Searle are unprepared to do. For Searle, intentional states such as beliefs and desires

---

[41] See Hands (2001: 334- 341) for a broader discussion of Rosenberg's ideas.

are caused by, but not reducible to, lower level physical processes in the brain. Consequently intentional states are legitimate entities to refer to in explanations of behaviour. For more on the implications of Searle's biological naturalism for folk psychology, see Searle (1991, 1994a, especially the appendix to chapter 2).

While these are perhaps the most philosophically significant criticisms of folk psychology that Rosenberg draws on, he also discusses a host of other problems associated with such an intentionalistic explanatory framework. For instance, Rosenberg notes the following difficulty concerning the logical relationship between reasons and causes. Within the covering law framework of scientific explanation accepted by Rosenberg, reasons are only explanatory if they bring about actions. Yet often agents have reasons for acting that are patently not the cause of what they end up doing. Rosenberg (1995: 34-35) cites the example of the jogger who runs ten kilometres a day and explains her action, truthfully, as a result of the belief that running is good for her. Yet the actual cause of her behaviour may well be an addiction to the endorphins released during the process of running. The implication of such an example is simply the warning that being able to establish the reasons that an actor has for some behaviour, even where these reasons appear 'rational', is not the same thing as establishing the causes of behaviour.

Not surprisingly, Rosenberg's analysis of the state of modern economics and the cause of its alleged plight has generated substantial criticism and debate, much of which is summarised by Hands (2001: 340-341). From the perspective I adopt in this dissertation, there are two principal problems with Rosenberg's position. The first, already mentioned, is that it ignores the Searlean-type response to criticisms of folk psychology based on notions of supervenience and emergence. The second is that in adhering to a traditional Humean philosophical framework, Rosenberg ignores developments in contemporary philosophy of science[42]. This has two notable consequences. One is that were Rosenberg to acknowledge alternative conceptions of scientific explanation, he would be able to retain many of the reservations he has about folk psychology without being forced to conclude that such explanations are necessarily unscientific. The second is that the very premise for Rosenberg's critique of economists' use of folk psychology, that economics fails to be scientific because of its poor predictive performance, is put into doubt.

---

[42] See Hands (1988).

IV SUMMARY

Having covered a large amount of ground in a relatively short space in these introductory notes, I want to conclude by returning once more to the main theme of my dissertation. The three essays that follow these notes are all concerned with one thing: uncovering and critically assessing the accounts that various approaches to economics give of the mental processes involved in economic decision-making. The form that this assessment takes is thoroughly ontological, reflecting my belief that such accounts must be judged, at least in part, in terms of how closely they correspond to actual human mental ontology (or our best accounts of that ontology). Numerous important issues are raised by such an exercise, including the nature of rational choice, of knowledge and of action, the applicability and limitations of the various conceptions of agency under consideration, and ultimately the usefulness of ontological enquiry into substantive economic topics.

SOME PROBLEMS WITH THE CONCEPTION OF THE HUMAN SUBJECT IN CRITICAL REALISM[1]

I INTRODUCTION

While many aspects of Tony Lawson's *Economics and Reality*[2] have been the subject of critical scrutiny and development, its conception of human agency appears to have received relatively little coverage in the literature. This is a pity, since Lawson advances a distinctive account of the human subject that plays an important role in his claims about society and social science.

In this essay, I examine the model of cognitive functioning associated with the human subject in *E&R*. By 'cognitive functioning' I mean the various levels of consciousness at which the human mind is said to operate and the type of knowledge that functions at each of these levels. Two realms in particular concern me: the *discursive* conscious, at which level operate the reflected upon, thought about, brain states that occur, for example, when we try to calculate the first ten prime numbers or set about planning a holiday; and the *tacit* subconscious, at which level operate the non-reflected upon, unattended to, brain states such as my knowledge of the various keyboard shortcuts in Microsoft Word that enable me, for example, to italicise words without my having to think about which keys to press.

While the idea that the human mind functions at various degrees of consciousness (from the wholly discursive to the non-conscious) is surely right, I argue in this essay that Lawson's account of tacit knowledge is unable to sustain the differences in functioning between discursive and tacit knowledge that his theory of social reality relies upon. In developing the argument, I first consider the importance of the human subject's cognitive functioning to the theory of social reality Lawson describes in *E&R*. I then outline the

---

[1] This paper has been published in the November 2002 edition of the *Cambridge Journal of Economics* (Faulkner 2002b). My thanks to Jochen Runde, three anonymous referees and to participants at the Workshop on Realism and Economics (November 2000), King's College, Cambridge, and the INEM session of the ASSA meetings (January 2001) in New Orleans, LA., for helpful comments on earlier drafts of the paper.

[2] In the text, *E&R* denotes *Economics and Reality*.

particular model of cognitive functioning Lawson attributes to the human subject. In section three, I set out my criticism of his conception of tacit knowledge.

In the second half of the essay, I propose a possible resolution of the problem, drawing on John Searle's work on the mind. The key addition from Searle turns out to be the technical notion of intentionality – a capacity by which some mental states are intrinsically about, or directed towards, some objects or states of affairs in the world. From an account of intentionality, Searle produces an explanation of both consciously reasoned behaviour and tacit knowledge that corresponds closely to the cognitive functioning Lawson attributes to the human subject, but which evades my original criticism.

II THE HUMAN SUBJECT IN *ECONOMICS AND REALITY*[3]

Lawson's ultimate aim in *E&R* is to propose an alternative model of explanation for economics and, to this end, he emphasises the need to ensure that the methods to be used are consistent with the nature of the subject matter to be investigated. For economics, this subject matter is social reality, which is why, in *E&R*, Lawson devotes considerable space to developing an understanding of the nature of society. The result is a theory of social ontology that centres on two key elements: an account of the social structures, such as rules, positions and relations that govern and facilitate human actions, together with an account of the human subject.

The particular aspect of Lawson's description of the human subject to be examined in this essay, the various levels of consciousness at which the human mind is said to operate, bears directly on two areas of his theory of social reality. The first of

---

[3] In focusing exclusively on Lawson's conception of the human subject, and specifically the notion of tacit knowledge in E&R, I do not intend to address the important issue of the genealogy of Lawson's critical realist account in relation to the work of Bhaskar (1978, 1979) and others (e.g., Archer 1995, 2000; Giddens 1976, 1984). Significantly, I do not therefore examine whether the same criticisms I make of Lawson's position can also be made of Bhaskar and other critical realist writers. On this, see Hodgson (1999b). I concentrate here on Lawson's account, rather than critical realist accounts in general, as it is Lawson's work that has impacted most directly on economics.

these areas is the capacity for purposeful[4] behaviour that Lawson ascribes to human beings, whereby action is 'governed by reasons so that it is always directed by some beliefs, or knowledge, and towards some end(s)' (Lawson 1997: 177). The process of reasoning associated with such goal-driven behaviour must take place in the discursive conscious, since it requires reflection upon our (often competing) aims and our knowledge of the world that dictates how best to achieve these aims.

The second area of Lawson's theory of social reality relating to the cognitive functioning of the human subject concerns the knowledge that human beings possess about social structures. Lawson posits the existence of various types of social structure; social rules or practices, social positions and social relations being the principal candidates. In order to be sensitive to the nature of these structures, human beings must draw on their knowledge of them whenever they act in a social context. Lawson notes that such knowledge is often imperfect and usually tacit, this latter characteristic meaning that such knowledge is typically drawn on, and shapes action without being the subject of discursive thought. Consider an example: although the shared social practice of queuing at the supermarket in order to pay for goods appears to be known to most of us, and our behaviour reveals this to be so, we tend to conform to this practice 'automatically', without conscious reflection.

The Cognitive Capacities of Lawson's Human Subject

At issue is whether the account that Lawson gives of the human subject's cognitive functioning is adequate in terms of its ability to support the forms of behaviour his theory of social ontology requires. I have identified two aspects of behaviour as being important to this theory: the functioning of beliefs, intentions and knowledge at the discursive level of consciousness that generates the reasons behind purposeful actions, and the functioning of knowledge concerning the nature of social structures that typically takes place at the

---

[4] Lawson (1997: 175) uses the word 'intentional' as a synonym for 'purposeful'. This usage, whereby intentionality refers to the directedness of *action*, contrasts with that of Searle, who uses intentionality primarily to refer to the directedness of many *mental states*. Therefore the intentionality of Lawson's human subject is quite different to the intentionality of Searle's. To avoid confusion, I will restrict my use of intentionality to Searle's meaning of the word.

tacit level of consciousness. I turn now to the model of cognitive functioning that Lawson develops in *E&R* to account for these two aspects of behaviour.

Lawson identifies three basic levels of consciousness at which the human mind operates: the discursive conscious, the tacit subconscious and the unconscious. The strategy employed to explain each of these levels is to outline the types or aspects of behaviour they make possible. At the discursive level, he tells us, operate the 'motives or needs which stimulate or prompt action and the overall plans or projects that may be formulated in line with them' (Lawson 1997: 178). These, he continues, 'tend to have direct bearing or purchase only in relatively unusual circumstances, such as moments of significant changes in possibilities and constraints' and so only 'bear on the reasons for particular actions rather than the modes in which they are chronically carried out'.

On the basis of this passage, I take Lawson to mean by the discursive conscious the 'reflected upon' or 'attended to' realm of consciousness or awareness. This would seem to follow from his mention of the 'reasons for particular actions' that function at this level, since I argued earlier that this process of reasoning involves conscious reflection upon our situation. This interpretation of the discursive realm is further supported by Lawson's characterisation of the tacit subconscious, where the 'reflected upon' (Lawson 1997: 177) nature of the discursive is contrasted with the 'without conscious reflection' (Lawson 1997: 178) nature of the tacit subconscious.

As before, the majority of what Lawson tells us about the tacit realm concerns the type of behaviour that functions at this level. In particular, he notes that 'much day to day conduct is concerned precisely with the question of possible modes in which relevant actions can be carried out' (*ibid.*). These modes, he argues, must typically be tacitly determined since the continuous nature of the events and actions that must be monitored and controlled rule out a discursive, reflected-upon, thought process. One consequence of this, he notes, is that 'much of the body of knowledge and skills drawn upon in action will be known only tacitly' (*ibid.*), being drawn upon without being thought about. What Lawson has in mind here, in terms of those things we typically know tacitly, is not only the social rules we adhere to on a regular basis, but such things as familiar geography – those routes we can navigate while 'lost in thought' – and the like. In all, I take Lawson to mean that our capacity for tacit knowledge essentially concerns knowledge of those things in the world that we are most used to acting in accordance with.

I shall concentrate on Lawson's conceptions of discursive and tacit knowledge in the remainder of this paper. But before I do so, it is necessary to mention the unconscious

level of cognitive functioning that also features in his account of the human subject. Lawson (Lawson 1997: 179) describes as unconscious those 'forms of cognition and dispositions which appear either to be inaccessible to discursive conscious because they were formed before the development of linguistic competence, or to constitute items that are repressed from discursive consciousness, being inhibited from appearing there in anything other than (possibly significantly) distorted form'. These dispositions, according to Lawson, represent unconscious 'needs and motives' (Lawson 1997: 184) which influence the actions that individuals take. In particular, Lawson cites Giddens' notion of ontological security – whereby people possess an unconscious need to engage in routine forms of behaviour in order to avoid anxiety – as a partial factor in the widely routinised nature of human behaviour.

Having highlighted the role of the unconscious realm in Lawson's account, I shall put it to one side for the remainder of this paper. Although an important part of his theory of social reality, the unconscious is not directly involved in the problem I will identify in Lawson's account of the human subject. In moving towards this problem, I now want to focus explicitly on the types of knowledge that Lawson ascribes to the human subject. His account of the realms of consciousness at which the mind operates suggests that there are two principal types of knowledge to consider: discursive and tacit knowledge.

What characterises each type of knowledge is the degree to which conscious reflection is involved, when that knowledge is drawn on in action. The sort of knowledge that gets brought to mind when we reflect upon possible courses of action to take, such as beliefs about the way things are in the world, functions causally in behaviour by virtue of being brought into the discursive conscious and its contents analysed. I term this 'discursive knowledge'. The second type of knowledge found in Lawson's account is 'tacit knowledge'. Such knowledge is drawn on without conscious reflection, as with the earlier example of 'automatically' knowing what it means to queue in the supermarket.

The only other distinction Lawson draws between discursive and tacit knowledge, other than the degree to which conscious reflection is involved in the functioning of each type, comes when he describes the process of acquiring tacit knowledge. The first stage of this process is typically the discursive acquisition of some piece of knowledge: people are taught the 'rules of (a possibly foreign) language, how to play a musical instrument or a game, how to hold cutlery or a pen, the highway code, and so forth' (Lawson 1997: 178). Such knowledge is discursive in the sense described above, since in playing the instrument or speaking the language, the rules are consciously applied at the time of

acting. With familiarity comes the acquisition of tacit knowledge of these rules, 'with time, such rules become internalised' (*ibid.*), at which point one can act according to the rules without having to think about their contents.

III THE PROBLEM – WHAT IS TACIT KNOWLEDGE?

What, then, is the problem with Lawson's account of human knowledge? Let us take a widely understood example of something that may either be discursively or tacitly known and then consider how we are to apply Lawson's model to this situation. My example will be the rules of a game, such as chess. A player thinking about what to play next in a game of chess must draw on his knowledge of the rules governing the legal moves of each piece and the definitions of such terms as 'check', 'castling', 'en passant' and so on. Consider first the novice, who possesses discursive knowledge of these rules. In deciding on his next move, he consciously applies whatever rules are relevant to the piece he is considering moving. To move the king legally, for example, means thinking through the rules to determine that such and such a move is allowed, while other moves are not.

Contrast the novice with the expert, who is sufficiently familiar with the game to know its rules tacitly. The expert's knowledge is still a causal factor in his decision-making – both players' actions are sensitive to the rules – but unlike the novice, the expert does not need to think through the content of each rule when drawing on his knowledge. Rather, while his tacit subconscious ensures that he operates within the rules of the game, his discursive conscious is likely to be concerned with aspects of strategy. The question that an account of tacit knowledge must answer is how it is that the knowledge of the rules possessed by the expert player is able to function causally in behaviour. For the novice, his knowledge of the rules functions by virtue of being brought into the discursive conscious. But the expert's knowledge of the rules also functions causally, this time without entering the discursive mind.

Lawson's explanation of the functioning of tacit knowledge is simply to say that the expert has 'internalised' the rules of the game. The trouble is that in the absence of any additional exposition of the term, all that this achieves is the renaming of tacit knowledge to internalised knowledge. What is needed, and what Lawson fails to provide, is an unpacking of the notion of tacit (or internalised) knowledge. My contention in the remainder of this essay is that John Searle's work on the human mind provides us with just this. While criticising accounts of the human subject that propose unconscious rule-

following, Searle observes: 'Our picture of unconscious mental states is that they are just like conscious states only minus the consciousness' and then asks, 'But what exactly is that supposed to mean?' (Searle 1995: 128). My criticism of Lawson takes the form of the same kind of question: what form does tacit knowledge take and how does it differ from reflected upon, discursive, knowledge? Of course, any answer to this question will be speculative to some degree, because the field of cognitive science is itself unresolved on these issues. However I think it can be shown that Searle's account of consciousness, intentionality and Background not only sheds further light on the nature of the problem identified in Lawson's account, but also offers a route to overcoming these failings.

IV SEARLE ON THE HUMAN AGENT

According to Searle, what distinguishes reflected upon, discursive, knowledge from non-discursive, tacit, knowledge is the *intentionality* of the former. By intentionality, Searle means the specific property possessed by some mental states by which they are intrinsically 'directed at, or are about or of, or refer to, or aim at, states of affairs in the world' (Searle 1999: 64). He goes on to argue that many of our conscious mental states function causally in our behaviour by virtue of being intentional. He then argues that intentional states can only function in association with non-intentional, 'Background', knowledge. What I want to suggest is that by identifying Lawson's tacit knowledge with Searle's Background, and discursive knowledge with intentional mental states, Searle's account of the human subject overcomes the deficiencies in Lawson's conception.

Intentionality

The notion of intentionality remains contentious in cognitive science. Since Searle's approach[5] to intentionality is fundamental to the other aspects of his work, we must be clear about what the Searlean perspective commits us to. If we take Searle's view to be the following, that 'it is just a plain fact [about human beings] that, for example, sometimes they get thirsty or hungry, they see things, fear things, etc.' (Searle 1994a: 79)

---

[5] See *Intentionality: An Essay in the Philosophy of the Mind* (Searle 1983), hereafter referred to simply as *Intentionality*, for a full account of the Searlean approach.

and that such states are intrinsically intentional in the sense of being directed at, or about, some thing or some state of affairs, then the key thing for Searle to explain is the apparent existence of other, non-mental, forms of intentionality.

Let us put such phenomena into one of two categories. In the first category are things like words, maps and pictures that human beings use to represent objects and events. The phrase 'The Bank of England', for example, while being non-mental does genuinely refer to an institution located in the City of London. In the second category are instances of behaviour by systems which appear to imply an element of intentionality: the thermostat that perceives changes in temperature, for example, or the leaves that arrange themselves on the branches of the tree because of their desire to absorb sunlight.

How does Searle explain the existence of these non-mental sources of intentionality? With respect to the first category, Searle (1994a: 78-82) recognises that language, pictures and maps are genuinely intentional in the sense that they literally do mean, or refer to, something else. But this intentionality is never intrinsic to the language itself; rather it is something which derives from the intentionality of language users and is consequently termed 'derived intentionality'. The second category of phenomena – the thermostat and the leaves – are not intentional at all according to Searle's account. While it is common in everyday language to speak of something's behaviour as if it had a mentalistic explanation – we may say that the thermostat *knows* what temperature to keep the room at, that it *acts* if it *perceives* an unwanted change in conditions – the intentionality here is purely metaphorical ('as-if' intentionality in Searle's scheme). There is nothing in the thermostat that possesses the intrinsic intentionality of many human mental states.

Intentional Causation

We are now in a position to consider the importance of intentionality to Searle's account of the functioning of the mind. The key thing here is the role that intentionality plays in mediating particular types of causal relations between the mind and the external world. Consider the following instances of causal contact between mind and world: my seeing an object, my remembering an event that took place last year and my desire to go on holiday being the cause of my flying abroad. In the first two cases, some aspect of the external world causes a particular mental state: the object causes a visual experience and the event causes a memory. In the third case the direction of causation is reversed; the desire causes

a change in the state of the world. What each example has in common is the presence of an intentional as well as causal component: seeing, remembering and intending all involve intentional states.

The particular form of the relationship between intentionality and the mind's causal relations to the world is what Searle terms 'intentional causation'. Because our interest here is in the way that an individual's knowledge of the world functions causally in behaviour, I shall restrict myself from now on to considering just one side of the relation: the mind-to-world direction of causation. The essence to intentional causation is the proposition that the relationship between intentionality and (mind-to-world) causation takes the following form: when an intentional state (such as my desire for a drink of water) causes some change in the world (in this case causes me to drink water) that mental state functions as the cause of the resulting behaviour by generating a conscious representation of the effect that it causes.

An additional piece of terminology, the idea of an intentional state's 'conditions of satisfaction', enables a more succinct definition of intentional causation. For those intentional states whose content is a proposition, such as my belief that it will rain later today, one can conceive of a truth condition that corresponds to that state. In the case of my belief, for example, it is true if and only if it does rain later today. Since a range of intentional modes can apply to a given proposition content – I can believe or desire or intend that it will rain later today – the notion of a truth condition must be broadened to deal with the fact that, rather than being literally true or false, my desire can either be fulfilled or frustrated and my intention, carried out or not. To capture this idea Searle introduces the notion of the conditions of satisfaction of an intentional state. These conditions state what must occur for a belief to be true, an intention to have been carried out, a desire to be fulfilled, and so on.

The notion of intentional causation can now be stated as follows: when an intentional state functions causally with respect to human behaviour, it does so because it generates a representation of, and consequently causes, its conditions of satisfaction. Consider again my desire for a drink. According to intentional causation, the desire functions as the cause of my subsequently taking a drink – bringing about its conditions of satisfaction – by means of generating a conscious representation of my taking a drink. The idea behind intentional causation is therefore that an intentional state causes the event that it represents because the person concerned seeks to bring about the conditions of

satisfaction of that state. More generally, intentional content functions because one thinks about the conditions of satisfaction it implies and acts on it.

So what type of behaviour is associated with intentional causation? In *Mind, Language and Society* (Searle 1999: 106) Searle observes that:

> Human behaviour, where rational, functions on the basis of reasons, but the reasons explain the behaviour only if the relation between the reason and the behaviour is both logical and causal.

The point here is that the notion of purposeful behaviour only makes sense if the reason – one or more mental state – not only causes the resultant behaviour but was also meant to cause that behaviour. If my desire for a drink was such that in running to the kitchen I tripped and broke my glasses, we would not say that I purposefully set out to break my glasses, even though we could cite my desire for a drink as part of the cause. 'Explanations of rational human behaviour thus essentially employ the apparatus of intentional causation' (*ibid.*), according to Searle, because if reasons are aimed at bringing about a particular effect they do so by representing that effect.

Since parts of Searle's analysis are clearly reminiscent of Lawson's discussion of purposeful behaviour, let us be clear about the similarities and differences. Searle and Lawson share a common definition of what it is for behaviour to be purposeful (Lawson (1997: 177), remember, wrote that purposeful action is 'governed by reasons so that it is always directed by some beliefs, or knowledge, and towards some end(s)'). Where Lawson's account was found wanting was in the absence of an adequate explanation of the distinction between the discursive and the tacit levels of mental functioning. More specifically, having argued that the reasons that cause behaviour operate at the discursive level of consciousness, he fails to account for the difference between discursive and tacit knowledge beyond characterising the first in terms of our 'thinking about' its contents and the second in terms of our not needing to do this. Searle, in contrast, provides us with the apparatus for understanding the functioning of reasons in behaviour, via the concept of intentionality. What remains to be specified is the precise connection between intentionality and consciousness in Searle's analysis.

Two aspects of this connection are important. The first, already implied, is that the functioning of intentional states in behaviour is a matter for the (discursive) conscious. When the desire for a drink causes me to take a drink, it does so because I am aware of my desire and I act upon it. Indeed, this is true regardless of the direction of causation:

my remembering or perceiving also implies a state of consciousness rather than unconsciousness at the time concerned. But what about those intentional states that are not, at the present time, conscious? Since at any given time most of our beliefs, intentions, hopes and so on are unconscious, Searle must explain the difference between those neurobiological structures that are unconscious intentional states and those that are simply unconscious non-intentional structures. This is the second important aspect of the connection between intentionality and consciousness in Searle's account and is referred to as the 'connection principle'.

The connection principle (Searle 1994b: 849) states that when we attribute an unconscious intentional state to a person,

> we are describing a neurobiological structure not in terms of its actual manifestation of occurrent mental properties, but rather in terms of its causal capacity to manifest those mental properties in the form of conscious thoughts and conscious behaviour.

In other words the only way to understand the notion of an unconscious intentional state S(r), according to Searle, is in terms of the neurobiological structure behind that unconscious state being capable of sustaining a conscious state of mode S and content r. Thus while we can (and do) attribute both conscious and unconscious intentional states to a person, such an attribution can only be understood in terms of the connection principle.

The Background

So intentional mental states cause behaviour through their conditions of satisfaction. Having decided I should like to read a newspaper this morning, I form the intention to go out and buy a newspaper, and this intention is correctly identified as the reason for my driving to the paper shop and buying a copy of *The Daily Telegraph*. According to intentional causation, the conditions of satisfaction of my intention both represent and cause the effect: my buying a newspaper. Now consider whether the intention alone explains my behaviour. In fulfilling the intention, I went to a particular shop, I bought *The Daily Telegraph* rather than *The Times* or *The Guardian*, and I drove rather than walked. Presumably, these aspects of my behaviour are caused by my possessing other, related,

intentional states – beliefs about the nearest shop that sells newspapers, about the easiest way to get there and about the relative merits of various papers – so that these states also played a causal role[6] in the functioning of my intention. We may say, then, that a given intentional state only functions relative to many other intentional states. Searle (1983: 141) also recognises this point. Intentional states, he writes

> do not function in an independent or atomistic fashion, for each intentional state has its content and determines its conditions of satisfaction only in relation to numerous other intentional states.

This is precisely the point of the preceding example: my intention to buy a newspaper could not have functioned without what Searle terms the 'Network' of intentional states within which my initial intention exists. The Network includes not just the most clearly related beliefs and desires but also the entire set of 'subsidiary intentions...hopes and fears, anxieties and anticipations, feelings of frustration and satisfaction', in short, 'the entire holistic network of intentionality' (*ibid.*).

The Network clearly covers a vast range of mental life. The sheer breadth and depth of the beliefs and desires each one of us possesses at any given time would surely be staggering if we ever stopped to think about ourselves in those terms. But while intentionality – as represented by the Network – concerns the purposeful aspects of behaviour, we have still to explain the human capacity for tacit-like functioning we earlier noted as being relevant to our knowledge of social structures and the like. Searle's 'Background thesis', the essence of which is that intentional states only function 'given a presupposed set of Background capacities that are not just more intentional states' (Searle 1999: 109), appears to explain just such a capacity. The remainder of this essay considers Searle's Background thesis and aims to substantiate the claim that the Background is a more appropriate way to conceive of tacit knowledge than that offered by Lawson.

Firstly then, what exactly is Searle's thesis? As the previous quote implies, Searle contends that there is more to the functioning of the mind in purposeful behaviour than the Network of intentionality alone. His contention is prompted by considering whether

---

[6] Since these other states are also intentional, the nature of this causal role also follows intentional causation. In other words, when I drew on my belief concerning the nearest shop, my belief functioned causally by representing its own conditions of satisfaction in my (discursively) conscious mind.

or not the Network is self-interpreting in the following sense: if we follow out each strand of the Network, identifying at each step what determines the conditions of satisfaction of the intentional state concerned[7], will a point be reached such that all conditions of satisfaction have been fully determined? Aside from possible technical difficulties presented by the fact that at any given time most of our intentional states are unconscious and that it may be difficult to individuate states, Searle argues that such a point will not be reached.

Instead, as we move out through the Network listing the things that are implicated in the determination of our intentional states' conditions of satisfaction, we shall reach a point where what we are listing seems too fundamental to be considered as yet more intentional states. For example, while it seems to make sense to say that I have a belief about the location of the nearest shop that sells newspapers and a preference for *The Daily Telegraph* over *The Times*, it seems wrong to say that I believe that the right-hand pedal of my car will make the car go faster and that the middle pedal will slow it down. Such a description is wrong because as an experienced and competent driver I do not need to think about which pedal to push when driving. When moving away from traffic lights, for example, rather than thinking to myself, 'to accelerate, push the right pedal', I simply push the appropriate pedal.

The difference here, between my knowledge of the nearest newspaper shop and my knowledge concerning the operation of the pedals in my car, is that, while knowledge of both sorts clearly plays a role in my actions, I do not hold a specific intentional hypothesis to the effect that the right-hand pedal is the accelerator and the middle pedal is the brake. That is, this second sort of knowledge is non-intentional.

The point, then, is that in addition to the Network of intentional states that contains so much of what we know about the world, Searle argues that we possess various skills, abilities, tendencies and dispositions that also embody knowledge about the world without being intrinsically intentional. It is this set of non-intentional capacities of the mind that Searle calls the Background. Consider again the sense in which I know how to

---

[7] So, starting with my intention to buy a newspaper, we found that my knowledge of the nearest shop and my preference for driving rather than walking caused me to drive to a particular shop. But what states were implicated in my driving? Well, my knowing that my car is parked next to the house, my knowledge of the local geography, knowing how to drive, and so on. But then consider the functioning of each of these: for example, what was implicated in my knowing how to drive?

or not the Network is self-interpreting in the following sense: if we follow out each strand of the Network, identifying at each step what determines the conditions of satisfaction of the intentional state concerned[7], will a point be reached such that all conditions of satisfaction have been fully determined? Aside from possible technical difficulties presented by the fact that at any given time most of our intentional states are unconscious and that it may be difficult to individuate states, Searle argues that such a point will not be reached.

Instead, as we move out through the Network listing the things that are implicated in the determination of our intentional states' conditions of satisfaction, we shall reach a point where what we are listing seems too fundamental to be considered as yet more intentional states. For example, while it seems to make sense to say that I have a belief about the location of the nearest shop that sells newspapers and a preference for *The Daily Telegraph* over *The Times*, it seems wrong to say that I believe that the right-hand pedal of my car will make the car go faster and that the middle pedal will slow it down. Such a description is wrong because as an experienced and competent driver I do not need to think about which pedal to push when driving. When moving away from traffic lights, for example, rather than thinking to myself, 'to accelerate, push the right pedal', I simply push the appropriate pedal.

The difference here, between my knowledge of the nearest newspaper shop and my knowledge concerning the operation of the pedals in my car, is that, while knowledge of both sorts clearly plays a role in my actions, I do not hold a specific intentional hypothesis to the effect that the right-hand pedal is the accelerator and the middle pedal is the brake. That is, this second sort of knowledge is non-intentional.

The point, then, is that in addition to the Network of intentional states that contains so much of what we know about the world, Searle argues that we possess various skills, abilities, tendencies and dispositions that also embody knowledge about the world without being intrinsically intentional. It is this set of non-intentional capacities of the mind that Searle calls the Background. Consider again the sense in which I know how to

---

[7] So, starting with my intention to buy a newspaper, we found that my knowledge of the nearest shop and my preference for driving rather than walking caused me to drive to a particular shop. But what states were implicated in my driving? Well, my knowing that my car is parked next to the house, my knowledge of the local geography, knowing how to drive, and so on. But then consider the functioning of each of these: for example, what was implicated in my knowing how to drive?

use the pedals in my car. According to this thesis of the Background, I have developed some sort of capacity, located in the brain, which enables me to drive. This ability to drive means that I can coordinate my use of the pedals, the gears, the indicators and so on, without having to think about what each of them does and how to operate them. It is a simple fact that I have this capacity. Notice that this ability contains my knowledge about what each pedal does – when I drive this ability manifests this knowledge – but it does so without my holding a particular belief, a particular conscious representation, that contains this knowledge.

In *Intentionality*, Searle supports the idea of the Background with arguments that consider three different areas of human behaviour: our understanding of the literal meaning of sentences, our use of metaphor and our ability to acquire physical skills. With each, the intention is to show that an explanation of the behaviour concerned requires a Background of non-intentional capacities. Consider in particular the first and third of these arguments. The first deals with the way in which we understand the literal meaning of sentences. Searle (1983: 145) asks us to consider the following sentences:

Tom opened the door.
Sam opened his book to page 37.
The surgeon opened the wound.

In each case there seems to be little scope for misunderstanding what the sentence means. That is, faced with explaining the meaning of any of the sentences, I would confidently predict a high degree of uniformity among the answers of any sample of the population. Now notice the general form of these sentences: each adheres to the pattern 'X opened Y' where the literal meaning of the word 'open' is the same in each case[8]. But while the literal meaning of 'open' – and so the semantic content it provides to each sentence – is identical across sentences, the way that that semantic content is understood clearly differs in each case. What constitutes opening a wound is not the same as what constitutes opening a door; our understanding of the word 'open' is context dependent.

So how do we correctly understand what it means to 'open' in each of these sentences? What matters is how our knowledge of the context enables the correct

[8] Since otherwise the meaning of the word 'open' is indefinitely ambiguous. See 'Literal Meaning' in Searle (1979).

interpretation. In the above examples it seems fairly clear that the relevant know-how is non-intentionalistic since, rather than thinking about the nature of doors, books and wounds we immediately know what it is to open the object concerned. But if books were unfamiliar objects, perhaps being something we had been told about but were not used to dealing with, it would be intentional knowledge that enabled our understanding of the phrase 'opened the book'. We would reflect on our knowledge of books and then interpret the sentence accordingly. Notice though that even in this case there must still be a Background. If our understanding of a sentence relies on other intentional knowledge, then our understanding of the semantic content of this additional knowledge ultimately requires the non-intentional know-how contained in the Background.

Now consider another area of human behaviour, the nature of physical skills such as being able to drive or play a particular sport well. Using a favourite example of his, the ability to ski, Searle contrasts the experience of the learner skier with the expert. The beginner receives a series of verbal and visual instructions telling him how to ski. These instructions rely on representations, which is to say that when he skis these instructions are part of the intentional content that determines his behaviour: he tries to keep his knees bent, thinks about putting his weight on the downslope ski and so on. As such, the instructions and the skier's behaviour are a case of intentional causation, with the instructions having a word-to-world direction of causation. Contrast this with the expert or 'skilled' skier, who skies without thinking about the instructions he once learnt. What from the point of view of mental functioning, asks Searle, distinguishes the two cases?

The cognitivist answer[9], Searle tells us, would be to suppose that whereas the beginner is conscious of the instructions while he skies, the expert has managed to 'internalise' the instructions, such that they now function unconsciously. This idea of the unconscious functioning of intentional content seems to be what Lawson has in mind in *E&R*[10]. In commenting on the fact that people first acquire their knowledge of the rules of

---

[9] See, for example, Polanyi's (1958) conception of tacit knowledge, Chomsky (1975) on the unconscious following of the rules of Universal Grammar and Fodor (1975) on our unconscious knowledge of the language of thought.

[10] If Lawson's position on tacit knowledge is something other than cognitivist, that is, if the meaning of 'internalised' is not the unconscious functioning of intentional content, then my criticism is that this alternative position is left unexplained. In either case, Searle's notion of intentionality is offered as the key to understanding the difference between the functioning of discursive and tacit knowledge. One phrase in particular invites a non-cognitivist reading of Lawson's notion of tacit knowledge. Having suggested that

languages, games, social situations and so on, discursively (that is, intentionalistically), and that over time this knowledge becomes tacit, Lawson (1997: 178) proposes that the rules become internalised:

> even where social rules are tacitly held many of them were once discursively acquired, people are initially taught rules [of languages, games etc.]...even if, with time, such rules become internalised and, temporarily at least, forgotten from discursive consciousness, being only tacitly known.

Searle gives a different answer. Rather than the expert skier being someone who has managed to internalise the instructions and who can now interpret them without explicit thought, Searle suggests that for the expert, the instructions are no longer a factor: 'repeated experiences create physical capacities, presumably realised as neural pathways, that make the rules simply irrelevant' (Searle 1983: 150). Which is the more plausible, he asks: that each physical skill is underlain by a number of unconscious mental representations or that 'practice and training in a variety of situations eventually makes the causal functioning of representation unnecessary in the exercise of the skill' (Searle 1983: 151)? If it is the latter, then it is the distinction between intentional (Network) and non-intentional (Background) knowledge that explains what Lawson refers to as tacit knowledge, rather than the consciousness or unconsciousness of given intentional contents.

These two arguments, concerning our capacity for understanding literal meaning and the acquisition of physical skills, provide strong support for our conceiving of the tacit realm along the lines of the Background, as Searle argues we should. Searle's strategy in putting forward these considerations is not to show either formally or decisively that his position is the correct one. Rather, he is concerned to show that, given the Background as a working hypothesis, evidence for the idea 'piles up everywhere' (Searle 1983: 152). It is noteworthy therefore that Searle suggests an alternative argument that represents a more formal 'proof' of the existence of the Background.

---

social rules may be represented in terms of an injunction of the form 'if x do y under conditions z', he writes, 'I have not suggested that agents acting on social rules typically represent them in this manner or indeed hold any explicit representation at all' (Lawson 1997: 172). We are then referred to chapter 13, in which Lawson develops the account of tacit knowledge I reviewed in section two of this essay, taking us no further than the notion of internalised knowledge.

The argument begins by supposing that all cognitive capacities are intentional; that is, that the opposite of what would be the case if the Background thesis is true. In this situation, mental processes would consist in moving from one semantic content to the next, calculating conditions of satisfaction and then acting. The problem with this picture is that semantic contents are never fully self-interpreting, that is, we have to know how to apply them, and that knowledge cannot consist of further representations without infinite regress. Take the example of a person's ability to walk (Searle 1983: 152). If this ability consists of various representations, a set of internalised rules for walking, we would have various rules such as 'first, move the left foot forward, then the right foot, then the left' and so on. As we saw earlier, however, our interpretation of such a rule is not fully determined by its semantic content: what exactly counts as a 'foot', as 'movement', as 'forward'? If the correct interpretation is fixed by further representations, then the problem will recur at the next step because these additional representations will require interpretation. The only way out, argues Searle, is to recognise that at some point 'we just act'.

In the case of walking, we do not need rules. We just have the ability to walk. In the cases where we do act according to rules, where rules function intentionalistically in our behaviour, we do not need any further rules for interpreting these rules. In this way, the Background of non-intentionalistic capacities grounds our intentionalistic mental life, preventing the possibility of circularity and infinite regress in our original conception of intentionality and intentional causation.

The Background and Social Rules

Having established the Background as:

1) functionally equivalent to the capacity for tacit knowledge possessed by the human subject in *E&R*, in the sense of it enabling the functioning of knowledge in behaviour without that knowledge being 'thought about'; and,

2) conceptually preferable to Lawson's account of tacit knowledge, since the notion of intentionality provides the missing element in his discussion of discursive and tacit knowledge,

we can now apply the idea of the Background to an individual's knowledge of social rules. The idea that much of what we know about society's structures is located in the Background is not without precedent. In *The Construction of Social Reality* (Searle 1995), Searle himself sets out a theory of social ontology which incorporates the Background as the means by which we are largely unaware of the nature of the institutions within which we act. The central point about such Background knowledge will by now be familiar: it is that 'one can develop, one can evolve, a set of abilities that are causally sensitive to specific structures of intentionality without actually being constituted by that intentionality' (Searle 1995: 142).

Consider the institution of money as it exists in the UK. We may think of the existence of notes and coins that possess the power to function as a means of exchange in terms of various rules: for example, 'notes and coins with a certain physical structure count as Sterling', 'Sterling is accepted in exchange for goods and services in the UK' and so on. While it is clear that our behaviour is sensitive to these rules – we possess knowledge of them – it also seems clear that when we use money we do not think about the rules that govern its use. According to the Background hypothesis, rather than either consciously or unconsciously following the rules, the explanation (Searle 1995: 142) of people's behaviour is that they

> have developed a set of dispositions that are sensitive and responsive to the specific content of those rules...they must have the ability to use money as a medium of exchange, and they have to be responsive to the distinction between counterfeit money and real money, even though they might not be able to tell the difference without expert help. Their behaviour has to be responsive to the fact that currency is valuable not because of the paper it is printed on but because it functions as a medium of exchange...those sorts of abilities, this type of know-how, that become ingrained are in fact a reflection of the sets of constitutive rules.

Such 'Background Causation' therefore involves a type of explanatory inversion away from the cognitivist conception of unconscious intentional causation. Rather than saying that a person behaves the way he does because he is following the rules of an institution, we should say that at the causal level, the person behaves that way because his Background disposes him towards behaving that way; at the functional level he has come to be disposed that way because that is the way that conforms to the rules of the

institution. In other words, the causal structures of the Background explain the behaviour and the structures are explained by the system of rules, but the Background need not itself be a system of rules.

V CONCLUSION

I have argued that the conception of tacit knowledge employed by Tony Lawson in *E&R* is inadequate in that it fails to account for the difference between knowledge that functions in behaviour by virtue of being thought about and knowledge that functions without being thought about. The notion of knowledge being 'internalised' is question begging, since we are not provided with an explanation of how such internalised knowledge differs from currently unconscious, but discursively-functioning, knowledge. The notions of intentionality, intentional causation and the Background, provide us with an answer: those neurobiological features of the brain that correspond to intentional mental states explain the discursive functioning of knowledge, while non-intentional brain structures account for our capacity for tacit knowledge.

THE HUMAN AGENT IN BEHAVIOURAL FINANCE: A SEARLEAN PERSPECTIVE[1]

I INTRODUCTION

One of the main themes in John Searle's philosophy of mind is that the mental input into rational behaviour consists of more than our conscious, reflected upon, reasons for acting. In addition there are Background states, predispositions, skills and capacities that function causally but non-consciously. As Runde (2002) argues, the dominant conception of rational agency in economics, expected utility theory, ignores the existence and importance of the Background. For expected utility theory, behaviour is all about preferences and beliefs[2]. In this paper I examine the Searlean credentials of a second conception of agency found in economics, that associated with the literature on behavioural finance.

The model of the economic agent assumed in behavioural finance, call him BFA, makes for an interesting object of study for at least two reasons. The first comes from the fact that the field of behavioural finance is concerned with the impact of various psychological influences on human behaviour in financial markets. As a consequence its conception of human agency is an explicitly descriptive one, informed by psychologist's research into how people actually form judgments and take decisions. While unquestionably rational in the sense that his actions are caused by reasons, the BFA is also influenced by factors that are less obviously to do with reason and conscious

---

[1] This paper has been published in the Spring 2002 edition of the *Journal of Economic Methodology* (Faulkner 2002a). My thanks to Jochen Runde, Tony Lawson, Clive Lawson, Leon Montes and Steve Pratten for helpful comments on earlier drafts of the paper.

[2] Despite the apparently clear ontology (based on beliefs, states of the world and preferences) associated with expected utility theory (on which, see section V of my third essay), alternative interpretations of utility maximisation (e.g., Chipman *et al.* 1971; Friedman 1953; Machlup 1946; Samuelson 1937, 1938, 1947, 1948) that emphasise consistency of behaviour or the evolutionary benefits of utility maximisartion suggest that beliefs and preferences are not essential to rational choice conceptions of action. See Ross and Nisbett (1991) for a review of the evidence on behavioral consistency; Sen (1993) for criticisms of the 'consistency' view; and Winter (1964) and Hodgson (1994) for problems with the evolutionary

reflection. It seems natural to enquire whether some of these factors are Searlean Background states. The second reason for studying the BFA is the present prominence of the field of behavioural finance. Since its emergence in the mid-1980s in the work of economists such as Shiller, Summers and Thaler, behavioural finance has seen a surge in research output over the past decade[3].

The paper proceeds as follows. In the next section I set out Searle's account of rational human behaviour, central to which is his theory of intentional mental states (mental states that are intrinsically about, or directed towards, something or someone) and the non-intentional Background. Section three discusses some aspects of human agency emphasised in the behavioural finance literature. I characterise these as traits falling into three categories: prospect theory, heuristic-based judgment and mental accounting. Section four examines the BFA from a Searlean point of view. I find that each of the three categories of traits seems to involve both conscious and non-conscious elements, consistent with the Searlean account of rational behaviour. I conclude with some comments on my findings.

II THE SEARLEAN AGENT: RATIONAL CHOICE AND THE BACKGROUND

From the point of view of its rational decision-making, two propositions lie at the heart of Searle's conception of human agency[4]:

- Where behaviour is rational, actions are caused by reasons. These reasons are intentional mental states that function in the discursive, reflected upon, conscious mind.

---

interpretation. Steedman (1989, 2000) and Hodgson (1988) argue that mainstream accounts have a strong implicit attachment to an intentionalistic ontology.

[3] See Shefrin (2000) for a recent account.

[4] This section draws on Searle's work on intentionality (Searle 1983), consciousness (Searle 1994a) and social ontology (Searle 1995, 1999).

- When intentional mental states function causally, they always do so in association with non-intentional, Background capacities and dispositions that function in the non-conscious, non-reflected upon, mind.

In this section I briefly expand on these two propositions. This requires a number of definitions, the first of which concerns the notion of a mental state. According to Searle, a brain state is only a mental state if it is '...in principle capable of causing that state in a conscious form' (Searle 1999: 86). The key point about this definition is that it enables a distinction to be drawn between unconscious mental states (such as my presently unconscious, but often conscious, belief that marriages are time-consuming events to organise) and non-conscious, non-mental, brain states (such as the secretion of enzymes in the brain). This distinction, for Searle, turns on accessibility to consciousness. The second important definition concerns intentionality. An intentional mental state is a mental state that is intrinsically about, or directed towards, something or someone. Beliefs, for example, have this intrinsic 'aboutness' property, since to have a belief one must believe such and such a thing to be the case. The same is true of wants, needs, desires, memories and so on.

While Searle recognises the existence of non-intentional mental states, such as a conscious feeling of undirected optimism or anxiety, he rules out the possibility that non-mental brain states may be intentional. This latter possibility is eliminated as a result of a set of arguments he collectively terms the connection principle (Searle 1994a, 1994b), the essence of which is that intentionality assumes accessibility to consciousness which in turn implies that the state be mental. The most important set of non-intentional, non-mental, brain states for our purposes are the contents of what Searle calls the Background, to which I return below.

With these definitions out of the way, I can return to my two initial propositions. The first deals with Searle's understanding of rational behaviour. Searle considers behaviour to be rational where it 'functions on the basis of reasons' (Searle 1999: 106), and where the relationship between these reasons and the behaviour they cause 'is both logical and causal' (*ibid.*). By requiring a logical relationship between a reason and the action it causes, Searle defines as rational only those actions that were meant to occur on the basis of the reasons held. So for example, if because of my desire to maximise my wealth and my belief that stock markets offer the best combination of risk and return I

decide to invest in stocks, my subsequent purchase of a stock portfolio would be considered rational.

The identification of rational behaviour with reasons that not only cause, but were intended to cause, the subsequent behaviour, suggests that the causal functioning of these reasons must involve the conscious mind. Reasons function causally because they are thought about, reflected upon and scrutinised as part of the decision-making process. Searle recognises this by identifying the reasons that cause rational behaviour with intentional mental states, such as the beliefs and desires that caused me to invest in stock. His theory of intentional causation then formalises the notion of reasons being logically as well as causally related to the subsequent behaviour.

According to intentional causation, when an intentional mental state is the reason for an action, that state functions as the cause of the action by generating a conscious representation of the effect that it causes. Suppose in the above example that I form the intention to purchase a portfolio of stocks and that this intention is the reason for my subsequent behaviour. Intentions are themselves intentional mental states and according to intentional causation, when the intention caused my behaviour it did so by sustaining a conscious representation of my fulfilling that intention. But this is precisely the point to identifying the intention as the reason for my behaviour: I possessed the intention, I reflected upon what it meant I should do, I acted. Rational behaviour, to summarise the point, involves intentional mental states that generate conscious thoughts in the process of functioning causally.

The second proposition relates to the role of non-intentional, Background, states in rational behaviour. Searle's argument here is essentially a logical one that proceeds from the idea of intentional causation. According to intentional causation, an intentional mental state functions causally by virtue of being consciously scrutinised and its 'condition of satisfaction' determined. The conditions of satisfaction of a mental state define what needs to be the case for an intention to be fulfilled, a belief to be true, a desire to be sated, and so on. Intentional states rarely function in isolation, in the sense that the determination of the conditions of satisfaction of any one state typically relies on the contents (and so the determination of the conditions of satisfaction) of many other intentional states.

Searle's point is that this process of intentional causation, the conscious determination of conditions of satisfaction, could not and does not continue without end. At some point, rather than another set of conditions of satisfaction being consciously

determined, the Background takes over. Having earlier defined the Background to be a set of non-intentional, non-mental, brain states, I can now introduce them in more detail. Being non-intentional, when Background states function causally they do so non-consciously. What this means is that in contrast to thinking through the contents of another intentional state, the Background taking over marks the stage at which things are simply accepted or taken for granted without further conscious enquiry.

Perhaps the most effective demonstration of Background states and their functioning is provided by the following example (Searle 1983: 145). Searle asks how we understand the following sentences:

Tom opened the door.
Sam opened his book to page 37.
The surgeon opened the wound.

The general form of each sentence is the same: each conforms to the pattern 'X opened Y', with the literal meaning of the word 'open' being the same in each case. However what it is to open each thing clearly differs; that is, the X and Y terms partly determine our understanding of each sentence. Now how do we achieve this? Do we reflect on the nature of each term, in the manner of intentional causation, determining a series of conditions of satisfaction? For these examples I suspect not. We understand each sentence without any conscious reflection of this sort; that is, the Background takes over straight away, in this case enabling our understanding through its non-conscious causal functioning. This needn't have been the case. Had one of the terms been unfamiliar, had we had to think about what it is for a surgeon to open a wound, then one or more rounds of conscious, intentional, reflection would have been necessary. Even in such a case though, the Background would have taken over at some point.

Finally then, just what are these Background states? Searle views them as the various skills, capacities, abilities and dispositions we acquire that function without conscious reflection or deliberation. He cites the Background, for example, as the location of much of our knowledge of the rules that constitute society (Searle 1995). More generally, the Background contains our know-how about the most familiar, most frequently encountered, aspects of the world.

III THE HUMAN AGENT IN BEHAVIOURAL FINANCE

I turn now to the conception of the human agent in behavioural finance. Although the literature here draws on a wide range of findings from psychology, three types of traits represent the most prominent characteristics of the BFA. These relate either to prospect theory, heuristic-based judgement or mental accounting.

Prospect Theory

Kahneman and Tversky's (1979, 1984, 1992) prospect theory conception of risky choice is central to behavioural finance and its associated conception of human agency. In common with expected utility theory, prospect theory adopts a consequentialist approach to choice, which is to say that in making decisions, people are assumed to be concerned with the likely outcomes of their actions. In particular, they evaluate possible courses of action based on i) the desirability, and ii) the likelihood, of each action's possible outcomes. Where prospect theory differs from expected utility theory is in its explicit incorporation of certain well-established psychological traits[5].

Kahneman and Tversky break down the decision-making process into two phases, an initial editing phase, followed by the evaluation phase. The editing phase involves a number of cognitive operations that transform the outcomes and probabilities associated with an action into a representation that then enters the evaluation phase. Among the most prominent editing operations are the following:

- The coding of outcomes into gains and losses.
- The simplification of prospects.
- The detection of dominance.
- The cancellation of common components.

The first operation represents one of the most important characteristics of the decision-maker under prospect theory: that outcomes are perceived in terms of gains and losses relative to some reference point. The reference point might be the status quo, such

---

[5] In addition to the sources referred to in the text, Camerer (1995) provides an excellent survey of the literature on the evidence for the psychological traits incorporated into the BFA.

that the gains and losses associated with a simple lottery coincide with the actual amounts that can be won or lost. Alternatively, the reference point may come from the framing of the problem (see below) or be suggested by the expectations or history of the decision-maker (the purchase price of a share, for example). In such a situation, the gains and losses associated with a lottery need not correspond to the actual amounts won or lost.

The remaining editing operations allow for other commonly observed traits. The simplification of prospects refers to the rounding of probabilities and outcomes, one effect of which can be to discard extremely unlikely events at this stage of the decision. Where dominance can easily be detected, dominated choices also tend to be discarded before the evaluation phase. The cancellation of components that are common to two or more prospects reflects a tendency for decision-makers to ignore the impact of shared, equally likely, outcomes in choosing between alternatives.

In the subsequent evaluation phase, each of the edited prospects is then valued. The overall value of a prospect depends on two scales:

- The subjective value, $v(x)$, associated with each possible outcome, where x is the gain or loss relative to the reference point.
- The decision weight, $\pi(p)$, associated with each possible outcome, where p is the probability of that outcome.

The first scale, the value function, has the following properties:

- Defined over gains and losses rather than final wealth.
- Concave over gains and convex over losses.
- Displays loss aversion.

Each property relates to a basic principle of human perception or judgement. By defining the value function over gains and losses, Kahneman and Tversky incorporate the finding that changes rather than end states tend to be the carriers of value. The curvature assumption reflects diminishing sensitivity over both gains and losses. This means that the difference in value between a gain of £10 and a gain of £20 is greater than the difference in value between a gain of £1,000 and a gain of £1,010.

The third property, loss aversion, captures the fact that people tend to feel the pain associated with the loss of a given amount more than they would enjoy a similar sized gain. That is:

$$-v(-x) \approx \lambda v(x)$$

where the loss aversion parameter, λ, is typically found to be in the order of two for small and medium-sized gambles (see Kahneman and Tversky (1992), for example, who fit prospect theory preferences to a sample of twenty five students using fictional lotteries that have prizes in the hundreds of dollars range).

The second scale, the weighting function, captures the effect that the likelihood of an outcome has on that outcome's contribution to the value of a prospect. As with the value function, the properties of the weighting function correspond to prominent judgmental and perceptual traits:

- Boundary effects: at either end of the probability scale, changes in probability are overweighted relative to changes in the middle of the scale. This means that the effect on the value of a lottery of a move from zero probability to positive probability, or from positive probability to certainty, is greater than the effect of a similar sized increase in probability that takes place away from either endpoint.
- The weighting function is unstable for low probability events, such that these events tend either to be overweighted or ignored.
- Away from the endpoints, decision weights are regressive with respect to probability.

Having sketched its main features, I now want to highlight three important implications of the prospect theory conception of decision-making. The first concerns the attitude to risk taking of the decision-maker under prospect theory. There are two points of note here. First of all, the curvature of the value function implies that the prospect theory decision-maker is risk averse over gains and risk seeking over losses. Secondly, with gains and losses being defined relative to a reference point and the reference point not being fixed for any given lottery, the risk attitudes of the prospect theory decision-maker are sensitive to context. To take one example, an investor's attitude to a share price rise depends on

whether the rise equates to a larger capital gain or a smaller capital loss; that is, whether he is risk averse or risk loving depends on the price at which the share was purchased.

Second, loss aversion plays a major role in behavioural finance, as demonstrated by the tendency (Shefrin and Statman 1985; Shefrin 2000) for investors to keep their loss making positions open for longer than the objective of maximising returns would suggest. Prospect theory tells us that investors will tend to focus on gains and losses when they make decisions. In editing these gains and losses the price at which the asset was purchased represents a natural reference point. Loss aversion means that the pain associated with a loss is felt much more than the pleasure associated with an equivalent gain. But investors do not suffer a loss until they sell their holding. Consequently, when faced with deciding which share in their portfolio to sell (whatever the reason for selling), loss aversion causes investors to sell their profitable positions before their losing positions in spite of the outlook for these shares. Indeed, as Odean (1998) and Heisler (1994) show, the 'winners' that get sold tend to subsequently outperform the 'losers' that are kept.

The final phenomenon that follows directly from prospect theory concerns the sensitivity of preferences to the framing of decision problems. The following example (Kahneman and Tversky 1984) demonstrates the effect:

1. An outbreak of a disease is expected to kill 600 people. Scientists propose two possible programs:

   Program A, according to which they estimate that 200 people will survive, or
   Program B, according to which they estimate that the probability that all 600 people will survive is 1/3 and the probability that none will survive is 2/3.

Kahneman and Tversky report that the majority of respondents choose program A. Such a response is typical of, and consistent with, prospect theory. Since the implied reference point for this problem is the loss of 600 people, the two programs deal with gains. According to prospect theory, people tend to be risk averse over gains, consistent with the majority favouring program A.

Now consider the same scenario presented in a different frame:

2. An outbreak of a disease is expected to kill 600 people. Scientists propose two possible programs:

   Program C, according to which they estimate that 400 people will die, or
   Program D, according to which they estimate there the probability of nobody dying is 1/3 and the probability that all 600 people will die is 2/3.

Faced with this problem, Kahneman and Tversky report that a majority now favour program D although the underlying scenario – the outbreak and the possible programs – is identical. So why does changing the frame lead to a change in preferences? The difference between the two problems is the implied reference point used for the calculation of gains and losses. In problem two the implicit reference point is zero deaths. The outcomes are therefore evaluated as losses. Prospect theory tells us to expect risk-seeking behaviour over losses, which is the observed result in this problem. Therefore the framing effect – the switch in preferences between two identical programs – is jointly caused by 'topical' (gains and losses) rather than 'comprehensive' editing and evaluation on the part of the decision maker and a frame that suggests a different reference point for each problem.

Heuristic-Based Judgement

Whereas prospect theory provides an account of the editing and evaluation of lotteries that are defined in terms of given outcomes and probabilities, the second set of traits relating to the BFA concern the way that people arrive at their beliefs about possible outcomes and their probabilities. People are found to rely on a variety of heuristics, or rules of thumb, in forming their judgements, the classic reference here being the collection of papers in Kahneman *et al.* (1982). Heuristics are like back-of-the-envelope calculations, reducing the amount of computation to be performed while providing a reasonable estimate of the required probability or quantity. Their interest to behavioural finance derives from the various systematic violations of the rules of probability that they introduce into people's judgements. Computational manageability, it seems, comes at the cost of bias. I will discuss three prominent heuristics in behavioural finance: representativeness, availability and anchoring-and-adjustment.

2. An outbreak of a disease is expected to kill 600 people. Scientists propose two possible programs:

   Program C, according to which they estimate that 400 people will die, or
   Program D, according to which they estimate there the probability of nobody dying is 1/3 and the probability that all 600 people will die is 2/3.

Faced with this problem, Kahneman and Tversky report that a majority now favour program D although the underlying scenario – the outbreak and the possible programs – is identical. So why does changing the frame lead to a change in preferences? The difference between the two problems is the implied reference point used for the calculation of gains and losses. In problem two the implicit reference point is zero deaths. The outcomes are therefore evaluated as losses. Prospect theory tells us to expect risk-seeking behaviour over losses, which is the observed result in this problem. Therefore the framing effect – the switch in preferences between two identical programs – is jointly caused by 'topical' (gains and losses) rather than 'comprehensive' editing and evaluation on the part of the decision maker and a frame that suggests a different reference point for each problem.

Heuristic-Based Judgement

Whereas prospect theory provides an account of the editing and evaluation of lotteries that are defined in terms of given outcomes and probabilities, the second set of traits relating to the BFA concern the way that people arrive at their beliefs about possible outcomes and their probabilities. People are found to rely on a variety of heuristics, or rules of thumb, in forming their judgements, the classic reference here being the collection of papers in Kahneman *et al.* (1982). Heuristics are like back-of-the-envelope calculations, reducing the amount of computation to be performed while providing a reasonable estimate of the required probability or quantity. Their interest to behavioural finance derives from the various systematic violations of the rules of probability that they introduce into people's judgements. Computational manageability, it seems, comes at the cost of bias. I will discuss three prominent heuristics in behavioural finance: representativeness, availability and anchoring-and-adjustment.

The representativeness heuristic, famously used by De Bondt and Thaler (1985) to explain their winner-loser effect, refers to people's reliance on similarity or representativeness when judging the likelihood that an event or object *A* belongs to class or process *B*. An investor, for example, may have to assess the likelihood that a company (*A*) has a profitable future (*B*) or not; or decide whether or not an innovative new product will prove commercially viable. According to representativeness, a common way for people to make such judgements is to rely on the degree to which *A* resembles the stereotypical *B*. When *A* is highly representative of *B*, the assessed probability is high; when *A* appears unrepresentative of *B*, the assessed probability is low.

The rationale for basing a probability judgement on representativeness comes from the fact that the stereotype captures some of the information on which a correct assessment of likelihood should be based. Yet to the extent that representativeness does govern judgement, certain types of bias are inevitable. Firstly, representativeness is independent of factors such as prior probabilities, sample size and predictability, which while relevant to the probability of *A* being an instance *B*, do not influence the resemblance of *A* to the stereotypical *B*.

Secondly, the representativeness heuristic sustains certain misconceptions about the nature of probabilistic events. For instance, people tend to believe that a sequence of events generated by a random process will be representative of the key characteristics of that random process, even over short sequences. Such an expectation is mistaken. Sequences of ten coin tosses, for example, will contain many more runs of consecutive heads or tails than representativeness would lead one to expect. A belief in such 'local representativeness' also sustains the 'gambler's fallacy', the mistaken belief that after a run of heads, a tail is 'due' on the next toss.

A second important heuristic in the behavioural finance literature is the availability heuristic, according to which people tend to assess the frequency of a class, or the probability of an event, by the ease with which instances of that class or event can be brought to mind. For example, the risk of death from a particular disease, or the likelihood of a given business venture collapsing, is judged on the basis of how many fatal occurrences of that disease can be recollected, or how many different failure scenarios can be imagined. The value of the heuristic comes from the fact that instances of larger classes or more likely events tend to be more easily recalled than instances of smaller classes, and that likely events are more easily imagined than unlikely events. Yet,

as with representativeness, the tendency to rely on availability leads to certain systematic biases in judgement, since factors other than frequency and probability affect availability.

Most obviously, the retrievability of instances also depends on factors such as familiarity, salience and experience. Two classes may be of equal size, but if the instances of one class are more familiar (due to press coverage, personal experience, and so on) than that class will be judged the larger. A second source of bias stems from the need to imagine, rather than remember, the instances of some types of classes. These frequency judgements will then depend on the ease with which relevant instances can be constructed, which need not be well correlated with actual frequency.

The final heuristic device, known as 'anchoring and adjustment', occurs when people's judgements are anchored to a particular value. When new information arises from which to re-evaluate a judgement, the adjustment process is typically insufficient (Lichtenstein and Slovic 1971). This 'conservatism' leads to judgements that are biased towards the anchor and as a consequence the source of this initial value becomes significant. The sort of bias that anchoring can introduce into judgement is discussed by Kahneman and Tversky (1974). They demonstrate, for example, the tendency for people to anchor their estimates of uncertain quantities to unrelated, but salient, values. They also show that the conservatism of the adjustment process leads to systematic over- and under-estimation of the likelihood of certain types of compound events. Shefrin (2000) uses the notion of conservatism to explain why earnings surprises tend to be positively serially correlated. Analysts and investors tend to underweight the latest earnings information in their earnings forecasts. This under-adjustment means that continued strong or poor performance is then under-anticipated.

Shiller (2000) looks at the role that anchors play in the day-to-day movement of stock prices. Noting that prices tend to be stable over relatively long periods but liable to sudden shifts, he distinguishes two types of anchor:

- Quantitative anchors, which indicate the 'appropriate' level of the market.
- Moral anchors, which determine the strength of reason for buying stocks.

Each type of anchor functions differently. With quantitative anchors, investors weigh numbers against prices in order to judge whether or not a stock is worth buying. The most likely anchors for individual stocks are salient price levels; for example, the

most recently remembered price or a notable price from the past, such as a recent high or low. Anchoring may also be to price changes or price-earnings ratios.

While quantitative anchors are particular values, moral anchors are the stories and reasons that people have for holding their financial assets relative to other possible uses of their wealth. The 'intuitive force' or strength of these stories relative to how much of their wealth is currently invested in the market determines whether a person invests more or less. These anchors often have cultural or ethical roots. Shiller notes the association between holding stocks and being thought of as sensible or prudent, an anchor and motive for market behaviour that applies regardless of the current level of the market.

Just as the stability of anchors goes some way to explaining the day-to-day stability of stock markets, their occasional fragility can be the cause of dramatic shifts in the market. Rather than continually assessing the validity of their anchors, people seem only to examine them in the face of exceptional events. This means that when anchors do change, they do so abruptly and well after the time it would, with hindsight, have made sense to adjust them. Perhaps something akin to this explains the nature of the 'dotcom' collapse in 2000. Despite the market rising and rising, the majority of people appear to have continued to believe the stories behind the hi-tech shares and remained invested, even when an examination of these companies' fundamentals would have shown their incredible valuations. They were anchored to their 'dotcom' judgements. Only when the market stopped rising and began to fall could people begin to comprehend that their anchors were hopelessly out of date.

Mental Accounting

In addition to prospect theory and heuristic-based judgement, the third set of traits drawn on extensively by behavioural finance relate to mental accounting. Thaler defines mental accounting to be the 'set of cognitive operations used by individuals and households to organise, evaluate, and keep track of financial activities' (Thaler 1999: 241). These operations matter because the rules that people adopt are not neutral with respect to decision-making. The assignment of a purchase to one account rather than another, or the decision to combine a number of outcomes within one category, or how often to 'balance the books', affects the decisions that people make.

Thaler distinguishes three components of mental accounting. The first deals with the accounting of individual transactions, in particular the way that outcomes are

perceived and experienced, decisions made and subsequently evaluated. In this respect, the mental accounting literature aligns itself with the editing and evaluation phases of prospect theory. Having already discussed these in some detail, I want simply to highlight some of the additional factors that mental accounting introduces into the editing phase. Consider the following decision problem:

Choose one alternative from each situation:

Situation 1: A: £2,400 gain with certainty, or
B: £10,000 gain with probability 0.25, nothing with probability 0.75.

Situation 2: C: £7,500 loss with certainty, or
D: £10,000 loss with probability 0.75, nothing with probability 0.25.

The most common response (Thaler 1999) to this 'concurrent choice' problem is the pair (A, D), which is equivalent to the lottery (-£7,600, 0.75; £2,400, 0.25). Yet the pair (B, C), which is equivalent to the lottery (-£7,500, 0.75; £2,500, 0.25), strictly dominates (A, D). The reason most people choose the dominated pair appears to be the tendency to evaluate each element of the choice separately. Taken alone, options A and D are consistent with the typical value function of a prospect theory decision-maker.

The failure to conceive of these two situations as one choice, or rather the tendency to evaluate a choice in the terms that it is presented (in this case, as two separate decisions), is a common finding in the mental accounting literature. Thaler and Johnson (1991) demonstrate a number of similar effects caused by the framing of payoffs. They show, for example, that preferences depend on whether part of a payoff is framed as a previous gain or loss. They also find that people are more likely to gamble after a win, a result that is well illustrated by the tendency for gamblers to keep their winnings in a separate pocket and be far more risk seeking in their use of money from the 'winnings' pocket.

The second component of mental accounting concerns the maintenance of mental accounts, or categories, and the assignment of activities to these accounts. Expenditures, wealth and income are all labelled in this way. The division of expenditures into different

accounts serves two main purposes. Firstly it enables 'rational' trade-offs to be made between competing uses for funds, at the level of types of spending rather than individual purchases. Secondly, it encourages self-control by allowing budgets to be operated for different types of consumption. Heath and Soll (1996) consider the mental processes involved in the maintenance of such consumption accounts. They propose a two-stage process; stage one involves expenses being noticed ('booked'), at stage two these expenses are then assigned to their correct accounts ('posted').

Wealth accounts also appear to serve as a self-control device, making certain funds less accessible than others. Shefrin and Thaler (1988) propose a hierarchy of wealth locations along these lines, arranged by how tempting it is for a household to spend the money in each. At the 'most easily spent' end of the scale lies the 'current assets' category which consists of cash in hand and money in deposit accounts; this wealth is routinely spent every period. Next comes 'current wealth', which consists of assets such as stocks, bonds and funds, which are less tempting to spend and are normally thought of as savings. Home equity comes next, followed finally by 'future income', which comprises future earnings and retirement savings.

The primary importance of this component of mental accounting to behaviour lies in the non-fungibility of money between different accounts. For expenditures this occurs when the budget for one category of consumption has been fully spent or recently drawn on. For wealth accounts Shefrin and Thaler estimate that the marginal propensity to spend a dollar of wealth in the current income account is near unity, while for wealth in the future income account this figure is close to zero. The 'mental location' of wealth therefore matters for consumption, in contrast to a simple life-cycle consumption model where the effect on consumption of a change in wealth ignores the type or source of this wealth.

Income accounting also affects behaviour. O'Curry (1997) examines the effect on consumption of income from different categories, by distinguishing 'serious' from 'frivolous' sources and uses of funds. She finds a tendency for people to match the seriousness of the source of a cash windfall with its use. Kooreman (1997) reports a similar finding, which is that the source of income partly determines how it is spent. Investors also appear to be affected by this phenomenon, apparently distinguishing between capital gains and dividends as types of income. Hatsopoulos *et al.* (1989), for example, report evidence of a 'mailbox' effect, according to which capital gains are spent much less readily than dividends, which tend to be spent as soon as they arrive.

The final component of mental accounting concerns the closing of mental accounts. To illustrate this aspect, consider Kahneman and Tversky's (1979) observation that betting on long shots at horse races tends to increase for the last race of the day. Why might this be? Thaler (1999) notes that the average bettor enters the final race sitting on a loss and with the desire to break-even on the day. That is, the typical gambler opens a racing account for each days racing, and closes it after the final race. Faced with a loss coming into the final race, the average gambler places additional value on the longshot because it offers the greatest chance to break even on the day, and to therefore close the account in the black.

While some situations offer a natural choice of 'bracket' for the mental account – the race day for the gambler – most investment scenarios take place over a long period of time. This provides an individual with some discretion as to the choice of bracket and the period over which gains or losses are evaluated. A significant result in this light is that people who are loss averse are more willing to take risks if they can combine bets together rather than considering them one at a time; in other words, if they can take a long-term view of their investment decisions[6]. This idea of 'myopic loss aversion' sustains Benartzi and Thaler's (1995) explanation of the equity premium puzzle (Mehra and Prescott 1985).

IV A SEARLEAN PERSPECTIVE ON CONCEPTIONS OF AGENCY IN BEHAVIOURAL FINANCE

While the move to incorporate aspects of human psychology into its conception of the human agent is the defining feature of behavioural finance, the literature is silent on what might be termed the 'Searlean dimension' to these traits, that is, on the role of consciousness in their operation. At the same time, it seems clear that a number of them relate to, or involve, non-conscious, Background dispositions as well as conscious, reflected upon, reasons and reasoning. In this section I develop the Searlean perspective.

Let me begin by reiterating what Searle's analysis does and does not say. As discussed in section two, Searle proposes that we understand rational behaviour in terms

---

[6] See Thaler (1999) for a derivation of this result. Experimental work (Gneezy and Potters 1997; Thaler *et al.* 1997) also supports the finding that longer time horizons induce greater risk taking.

of behaviour caused by conscious reasons that are aimed at bringing about certain actions. These reasons function intentionalistically but do so only in association with non-intentional, Background 'know-how', which functions without being part of the discursive conscious. The force of Searle's position is therefore to show that wherever there is intentionality, it must 'bottom-out' in non-intentional Background states. What Searle's analysis does not do is dictate which traits are to be understood in either intentional or non-intentional terms. This is a matter to be settled by an analysis of the traits themselves.

Prospect Theory

According to the Searlean definition, actions are rational where they are caused by reasons and where these reasons are both causally and logically related to the subsequent action. For Searle these reasons are intentional states that function in the conscious mind. Is the decision-maker under prospect theory rational in this Searlean sense? Reflection on the two stages of prospect theory suggests that he is, with the reasons of the prospect theory decision-maker functioning during the evaluation phase of the decision. In support of this, consider the conscious, reflected upon, processes involved in decision-making. The principle processes to consider relate to the subjective valuations of possible outcomes, the weighting of outcomes and the subsequent combining of these two scales into a valuation of each prospect.

As an example, consider the process of valuing an outcome. Here the decision-maker must reflect upon his likes and dislikes and infer from these the value of each possible outcome. The Searlean notions of determining conditions of satisfaction, of scrutinising the contents of our intentional states and of intentional causation seem appropriate here. The same can be said concerning the decision over how to weight and combine each possible outcome into a final valuation for each prospect. These refer to conscious, intentional, processes. In light of this, the key characteristics of the value and weighting functions (loss aversion, diminishing sensitivity, the overweighting of certain outcomes etc.) are to be thought of as prominent characteristics of human preferences, by which I mean that they are typical features of the intentional states that people possess.

If the evaluation phase of prospect theory deals essentially with the intentionalistic aspects of behaviour, what can be said of the editing phase? Why is it, to take its most important aspect first, that people focus on gains and losses rather than comprehensive

outcomes when they take decisions? Reflecting on this tendency, it rarely seems to be the result of a conscious decision to edit outcomes in this way. Rather it is something that is typically done without reflection, a way of proceeding that we are simply predisposed towards. In support of this, consider the earlier framing effect. The observed difference in preferences between the two identical disease-control programs was understood in terms of people valuing changes, combined with a shift in reference point. It seems wrong to associate these effects with decision-makers' conscious deliberation. Rather, once the deliberative stage is reached, the editing has already occurred.

It is interesting in this light that Kahneman and Tversky report the responses of the experimental subjects who faced this particular problem. When confronted with their conflicting answers, the subjects reported that they still wished to be risk averse in the 'lives saved' version of the problem, risk seeking in the 'lives lost' version, but that they also wished to give consistent answers. Kahneman and Tversky conclude that 'in their stubborn appeal, framing effects resemble perceptual illusions more than computational errors' (Kahneman and Tversky 1984: 5). If their cause is a Background disposition towards evaluating changes relative to some reference point, then Kahneman and Tversky are right to reject the comparison to computational errors, since these would seem to imply conscious rather than non-conscious cognitive processes.

Although the tendency to focus on gains and losses is the most important of the prospect theory editing operations, the others discussed earlier – the detection of dominance, the simplification of prospects and the cancellation of common components – may also suit a Background, rather than intentionalistic, interpretation. In ignoring obviously dominated or inferior alternatives, for example, it is clear that we need not be consciously deciding to disregard such options. Rather, the experienced actor may become predisposed towards ignoring them, leaving his intentionality to evaluate the alternatives that need to be decided between. I would argue that the same applies for the observed tendencies to round or simplify probabilities and outcomes and cancel common components; rather than discursively thinking about and implementing such operations, we find that some people simply have simplified and cancelled prior to the conscious evaluation process.

Heuristics

In putting forward a Searlean interpretation of heuristic-based judgement, the first question to address would seem to be whether or not heuristics function like rules, which people consciously apply in situations requiring a probability judgement. In principle heuristics certainly could take this form. For example, the representativeness heuristic could refer to a person who consciously reflects to himself, 'the way to calculate the likelihood that event A belongs to class B is to consider how closely A resembles my stereotype of B'.

In general, however, this does not seem to be what the idea of a heuristic is getting at. Rather than consciously deciding to form judgements on the basis of things such as availability or representativeness, it seems better to think of heuristics as Background states that predispose us towards relying on these things as the basis of our probability judgements. This fits with Kahneman and Tversky's link between the operation of heuristics and the making of what they call 'intuitive judgements' (Kahneman and Tversky 1974: 18), by which they mean judgements that do not involve full conscious deliberation at each step.

While heuristics themselves may reside in the Background, this does not prevent certain intentional processes being involved in their operation. The availability heuristic, for example, relies on the intentional operations of retrieval, construction and recall. But this point simply emphasises the difference and interaction between the intentional and non-intentional levels of functioning. The decision-maker, faced first with the need to form a probability assessment, is guided by Background dispositions to focus on availability. This occurs unthinkingly. What takes place in the conscious mind is the working through of instances – whether by recall or construction – that fit whatever event is under consideration. The result is a judgement that overweights certain pieces of information while ignoring others.

Finally on the subject of heuristics, what are we to make of the quantitative and moral anchors referred to by Shiller? Shiller himself comments that 'psychological anchors for the market hook themselves on the strangest of things along the muddy bottom of our consciousness' (Shiller 2000: 147), which while being a graphic description does no more than hint at their actual functioning. Let us first consider quantitative anchors. A typical quantitative anchor might be the lowest price that a particular share has reached in recent times. As an anchor, this price is the level below which I judge that the price will not fall. It therefore functions as a reason for my

purchasing the share when that price is reached. There are two separate aspects to consider here.

The first is that when the anchor causes me to make a purchase, it does so intentionalistically. That is, the anchor provides the conscious and intentional reason for my buying these shares at this time. But what about the 'decision' to use this anchor in the first place? Are anchors such as recent price highs and lows, or recent percentage changes, used by choice or are they things that we are predisposed to relying on in our judgements, perhaps because of their prominence in the absence of anything better to base a judgement on? On this question there is no doubt some truth to Shiller's earlier comment. The anchors we find ourselves basing market judgements on may be the outcome of conscious reflection, but what is more likely is that they are the result of something less considered, something more akin to a Background disposition.

The same sort of analysis seems to apply to moral anchors, although an important consideration here is the idea that much of the human thinking that leads to action is in the form of storytelling, narrative and justification. According to Shiller, moral anchors have the effect that they do because convincing or compelling stories can lead to action, rather than requiring an analysis of the most preferred possible action. Consider how easy it is to act on the basis of a share tip from a friend, or to purchase the shares of companies that one has personal knowledge of. In each case one can tell vivid stories and these alone can provide sufficient reason to invest. Here again, the intentionalistic aspect of anchors when they cause behaviour is evident.

Finally, consider the insensitivity of anchors to market conditions in anything other than exceptional circumstances. This seems to reinforce the non-intentional, unquestioned, nature of anchors. For example, if one of the factors in the recent 'dotcom' boom really was a well-established moral anchor concerning the importance of being invested in technology stocks, why did people fail to question themselves about the validity of the anchor? If anchors were in some sense decided upon intentionalistically, they would surely be consciously scrutinised in a way that they appear not to be. This would suggest that to be an anchor, is to be part of the Background.

Mental Accounting

Third and finally are the cognitive operations behind mental accounting. The first component of mental accounting, concerning the accounting of individual transactions,

has largely been dealt with under prospect theory. One additional element, represented by the concurrent choice problem analysed earlier, provides further evidence on the Searlean nature of the prospect theory editing operations. Consider again the interesting feature of this problem. By framing a given set of outcomes as the result either of two simple choice situations, or one compounded one, people are found to switch their preferences. The mental accounting explanation of this phenomenon is that people tend to retain and use the division suggested by the frame of the problem, rather than compound the choices. I would suggest that this reflects a Background disposition towards what might be termed 'keeping things simple'. Certainly it seems to be something we do in an unthinking way, as I argued were many of the prospect theory editing operations, rather than being a conscious decision to keep separate accounts for each part of the decision.

The second component of mental accounting concerns the maintenance and operation of categories, budgets and accounts. Heath and Soll (1996) describe the cognitive operations involved in maintaining expenditure accounts. Their analysis of a two stage accounting process finds that the first stage, the 'booking' of expenditures, relies on attention and memory, while stage two, the 'posting' of expenditures to accounts, relies on judgements and categorisations. These operations, which have counterparts in income and wealth accounting too, are based on conscious, intentional, processes. Perhaps the more interesting issue concerns the effects associated with the maintenance of these accounts. Is the finding that spending tends to be matched to the income source, or that the dividends are spent much more freely than capital gains, a result of intentional or Background functioning? Since being influenced in spending by the 'label' attached to income is not an obviously beneficial trait, it is unlikely to be the result of a conscious decision to link spending to income in this way. What seems more likely is that such behaviour is the result of certain acquired Background dispositions.

The final component of mental accounting, dealing with the closing, or balancing, of mental accounts, was argued to be of considerable importance to investor behaviour given its role in determining attitudes to risk. Ultimately, one always has discretion over the choice of time frame or boundary for the transactions within an account. The interesting distinction is between those activities where a natural choice of frame suggests itself, such as the betting day for the gambler, and those activities where there is no obvious frame, such as the investor who continually holds a portfolio of assets and can choose how often to evaluate his gains or losses.

In the case of the gambler, when the end of the betting day is the natural time to 'balance the books', it seems clear that this frame functions through the Background. The gambler, unless intentionally trying to do otherwise, finds that drawing a line under the day's bets is the natural way to proceed. As a result of this, in combination with loss aversion, horses with long odds will tend to be backed in the final races of the day when in the absence of the daily frame, a gambler would not feel obliged to take on this greater risk. Notice that as with the other areas I have analysed, this Background functioning is combined with the intentional causation associated with deciding upon which horses to back, maintaining the daily betting account, and so on.

For the investor, typically without a natural time frame to operate within, the case is even more interesting. What does determine how often a portfolio is evaluated, with the losses suffered or the gains enjoyed? Given that whatever it is that causes the investor to value his portfolio frequently or otherwise directly influences his investment behaviour, the importance of this question is clear. In the absence of a conscious effort to work to some decided upon time scale, context would once again seem to be the crucial factor. For example, investors with low costs to time, information and trading are likely to track their performance more closely then investors who cannot afford the time to do so or to whom trading is expensive. The perverse result here of course is that being time constrained, for example, may be beneficial if it prevents myopic loss aversion.

What about investors who realise that they suffer myopic loss aversion and decide to force themselves to watch their portfolio's less closely? The interesting Searlean observation in this context is that the investors who best manage to control themselves are likely to be those that can acquire the time frame they wish to use as a Background disposition. The investor, in effect, has to try and move portfolio monitoring away from the intentional brain into the Background. The development of such Background dispositions can usefully be thought of as part of the learning process of the investor.

V IMPLICATIONS

My analysis suggests that the psychological traits of the human agent in behavioural finance admit a Searlean interpretation in two ways. Firstly, a number of the individual traits considered in the behavioural finance literature seem to be best conceived of as Background dispositions. Secondly, the various mental processes involved in the

decision-making of the BFA – evaluation, judgement, mental accounting etc. – involve both conscious and non-conscious elements, just as Searle's theory of intentionality and the Background would suggest.

By way of final remarks, I want to comment on some implications that seem to follow from my analysis. First of all, the Searlean perspective reinforces a certain view of the rationality of the BFA that contrasts with that often given in the literature. Within traditional finance, the agent whose behaviour conflicts with the predictions of expected utility theory and Bayesian probability theory tends to be labelled 'irrational' and 'unsophisticated', or as being representative of the behaviour of non-rational 'noise traders' (Kyle (1985), Black (1986)). Even within the behavioural finance literature, Shleifer talks of behavioural finance as being the study of 'human fallibility in competitive markets' (Shleifer 2000: 24), by which he means the placing of 'the biased, the stupid, and the confused into competitive financial markets' (*ibid.*).

Two comments seem appropriate here. The first is that rationality should not be confused with expected utility maximisation. Expected utility theory and prospect theory are both accounts of rational choice and so the decision-maker under prospect theory cannot be regarded as irrational, stupid or confused. But what about the non-conscious traits of the BFA, such as the heuristics that people fall back on to form judgements. Do these represent human irrationality? From the Searlean point of view the answer is that they need not. The central message of Searle's analysis is that rational behaviour necessarily involves both conscious reasoning and non-conscious dispositions. A conception of rational choice that excludes the non-intentional is therefore an incomplete one. It is in this light that I commend Shiller's (2000: 136) assessment of the traits of the behavioural finance agent:

> These patterns of human behaviour are not the result of extreme human ignorance, but rather of the character of human intelligence....Investors are striving to do the right thing, but they have limited abilities and certain natural modes of behaviour...

For Shiller's 'natural modes of behaviour' I propose we read Searle's Background. A Searlean perspective can therefore ground claims about the rationality or otherwise of a particular conception of agency. A second area to consider, albeit one that I can only scratch the surface of here, concerns the importance to finance of recognising

the non-intentional, Background, factors in rational behaviour. The key point is that the Background always matters. The non-deliberative elements of behaviour are not infrequent anomalies that one's analysis can ignore. To drive home the point, I suggest the following as important questions within finance that will not be properly addressed until their non-intentional aspects are understood:

- Why do different investors have such different models of what represent good investments?
- How do people perceive risk and what determines their attitudes to risk-taking?
- Do people adopt investment 'styles'? Can they learn to be more successful investors, that is, can harmful traits be overcome?
- Is technical analysis ('charting') ever a profitable strategy? If so, do traits (such as the anchoring of expectations to salient price levels) explain patterns such as the tendency for price trends to reverse?

Finally, I want to emphasise that Searle is not alone in pursuing an account of human ontology that incorporates the sort of dispositions and capacities found in the Background into human behaviour. Within economics, authors such as Hayek (1967), Polanyi (1967) and Lawson (1997) have developed similar ideas. While there exist differences that remain to be resolved between these writers' accounts and Searle's[7], this paper has hopefully demonstrated the importance to economics of acknowledging theories of human ontology of this type.

---

[7] See Faulkner (2002b) for a Searlean perspective on Lawson's critical realist human agent.

INFORMATION, KNOWLEDGE AND ECONOMIC AGENCY[1]

I INTRODUCTION

The relationship between the information available to economic actors and what they can actually be presumed to know, and then how and in what form this knowledge informs their actions, raises some of the most difficult questions in economics. In this paper I concentrate on how the relationship between information and knowledge is treated in mainstream microeconomics. My aim is to establish that even in models that relax the standard assumption that actors have 'perfect' knowledge, this typically amounts to no more than introducing small 'black-spots' into the otherwise unlimitedly sharp and comprehensive knowledge that they are assumed to have of their model 'world'. I then show how this approach, while useful in some respects, fails to address some important aspects of the relationship between information and knowledge.

My argument begins with a brief account of what is meant by information and knowledge. The paper then splits into two halves. The first half explores the mainstream economics perspective on information and knowledge by way of a detailed examination of a representative game theoretic model. The second half considers three aspects of the relationship between information and knowledge that are largely neglected on the mainstream approach: non-probabilistic forms of uncertainty and ignorance, the subjectivity of knowledge, and tacit knowledge.

II INFORMATION AND KNOWLEDGE

I begin, then, by explaining what I mean by information and knowledge in what follows. Central to my use of these terms is the need to distinguish between what people could, in principle, know about their situation (information) and what it is, in fact, that they do

---

[1] My thanks to Jochen Runde for first suggesting the outline of the argument developed in this paper, as well as for many helpful comments on various drafts. I also wish to thank John Davis, Alain Marciano and participants at the Cambridge Realist Workshop (November 2002), Newnham College, Cambridge, for their comments on an earlier draft of the paper.

know (knowledge). To capture this distinction, I define information to mean any sort of fact about the world, where a fact is simply a property of the world, for example, that something exists, has occurred, takes a particular form, and so on[2]. Knowledge then refers to the beliefs and capabilities which people possess that embody information.

Now since beliefs and capabilities are quite different forms of knowledge, the two must be treated separately. Let me begin with beliefs. Knowledge in the form of beliefs corresponds to propositional knowledge (knowledge that). Such knowledge, acquired through various forms of exposure to information (studying, communicating, introspection, and so on), is directly available to the conscious mind. The important issue with this type of knowledge concerns the criterion to be used in determining which of our beliefs count as knowledge. In the philosophical literature a common assumption is that true belief is a necessary condition for knowledge, while some philosophers go further and argue that justification, when added to true belief, is a necessary and sufficient condition for knowledge[3].

Now certainly I wish to rule out as knowledge those beliefs that are plainly false, since they do not embody information. Yet true belief may not provide a necessary condition for beliefs to be considered knowledge. The point here is that our beliefs are often only accurate in some regards, whether because they are partial or fragmentary, or because they are only approximately true. Now in many cases such beliefs ought to be recognised as knowledge, since they still provide a reliable basis for action. Suppose I believe that the time is ten o'clock, when in actual fact the true time is five minutes past ten. If I am planning to catch the two minutes past ten train, then my belief that it is ten o' clock is sufficiently inaccurate that it ought not be regarded as knowledge. Yet for deciding whether it is time for my morning cup of coffee, my belief is likely to be more than good enough. The problem resides in the notion of there being an absolute standard by which we can judge a belief to be knowledge or not, since it would seem that what counts as knowledge depends on the situation. In what follows then I intend to regard

---

[2] My definition of information therefore differs from that of some authors who draw a distinction between facts and information, where information consists of facts that have already been systematised in some way.

[3] See Hands (2001: 144-150) for a more detailed account of debates surrounding belief-based definitions of knowledge. As will become clear later, mainstream microeconomic theory tends to equate knowledge with true belief, to the point of assuming a perfect correspondence between actors' beliefs about their world and the world itself.

beliefs as knowledge when they are true relative to the context in which they play a part, recognising that even this definition leaves the possibility of substantial grey areas.

In addition to beliefs, I also identified capabilities as being a form of knowledge. While beliefs correspond to propositional knowledge, capabilities that embody information correspond to tacit, sub-conscious, functioning knowledge (knowledge how)[4]. Here I am thinking of such capabilities as our tacitly functioning ability to ride a bicycle or to adhere to the rules of grammar when communicating, without the need for conscious reflection. Such capabilities are again the result of exposure to information, in this case via example, practice, repetition, and so forth. In section VII I consider the impact of the neglect of tacit knowledge in standard mainstream microeconomic models of the sort considered in the following two sections. I therefore defer further comment on tacit knowledge until then.

III A GAME OF COMPLETE INFORMATION

In this section and the next I consider the treatment of information and knowledge in mainstream microeconomic theory. I do so using the example of a one-shot Cournot duopoly game, first under conditions of what game theorists call 'complete information' and then, in the next section, under conditions of 'incomplete information'. I have chosen this example, partly because it is simple and compact enough to be considered in full, but mainly because it is highly representative of the kind of analytical approach widely adopted in contemporary economics. My aim in this section is to demonstrate the conventional assumption that actors' knowledge of the contents of their 'model world' corresponds exactly to the formal expression of that world. That is, knowledge and information in these models coincide perfectly. In the section that follows, I argue that even in models that explicitly recognise that information and knowledge are not coextensive, this amounts only to introducing small 'black-spots' into the otherwise perfect knowledge attributed to the actors concerned.

---

[4] While I concentrate in this essay on John Searle's understanding of capabilities as a form of knowledge, a similar point is made in modern pragmatist thought (Joas 1993, 1996; Plotkin 1994; Putnam 1995). See Hands (2001) for an account of the pragmatist revival in modern philosophy of science.

A game is one of complete information if the structure of the game (the moves, the payoffs, and so on) is common knowledge to all of the players[5]. The particular game I shall consider here involves two firms, A and B, in competition as the sole suppliers of a homogenous good. The firms simultaneously choose their preferred output (q_A and q_B), each knowing that total demand for the good is characterised by the inverse demand function:

$$P(Q) = a - Q$$

where $Q = q_A + q_B$. The firms employ identical technologies, such that the total cost to firm i of producing a quantity q_i is:

$$C(q_i) = cq_i$$

I assume that $a > Q$ in equilibrium (such that the market clears at a strictly positive price) and that $c < a$. The structure of the game is assumed to be common knowledge, something I return to after reviewing the solution to the game.

The standard game-theoretic solution to the one-shot Cournot game with complete information is the Nash equilibrium. This involves determining a pair of strategies, one for each firm, such that each firm is maximising its payoff given the strategy being played by the other. Strategies for each firm are simply quantities, and the payoffs are the firms' profits. Expressed as a function of the strategies chosen by each firm, profits for firm i given firm j's output are:

$$\pi_i(q_i, q_j) = q_i[P(q_i + q_j) - c] = q_i[a - (q_i + q_j) - c]$$

The strategy pair (q_A^*, q_B^*) is a Nash equilibrium if, for each firm i, q_i^* solves the maximisation problem:

$$\max \pi_i(q_i, q_j^*) = \max q_i[a - (q_i + q_j^*) - c]$$

---

[5] A proposition is common knowledge in some social system if everyone in that system knows it, everyone knows that everyone knows it, everyone knows that everyone knows that everyone knows it, and so on.

The first-order condition for this problem, which is both necessary and sufficient given the assumptions made about demand and technology, is:

$$q_i = \tfrac{1}{2}(a - q_j^* - c)$$

This means that for the strategy pair (q_A^*, q_B^*) to be a Nash equilibrium, the firms' quantity choices must satisfy:

$$q_A^* = \tfrac{1}{2}(a - q_B^* - c)$$
$$q_B^* = \tfrac{1}{2}(a - q_A^* - c)$$

Simultaneously solving this pair of equations yields the Nash equilibrium pair of strategies:

$$q_A^* = q_B^* = \tfrac{1}{3}(a-c)$$

The model having been set up and solved, my interest now is in the assumptions about information and knowledge that the model makes (leaving aside the meaning and significance of the Nash solution). Remember that information was defined to be any fact about the world. In the context of the 'model world' I have just considered, then, information consists exactly of no more and no less than the formal statement of the world:

1) The Players

The players consist of two identical firms that:

- produce an identical good;
- are fully described by their cost function; and
- are rational in the sense that they are single-minded in their goal of profit maximisation and perfect reasoners in the sense that they do not make any slips or other mistakes of reasoning, choose consistently, and are able to deduce the full logical consequences of all the knowledge that is attributed to them; and
- possess common knowledge of their situation.

2) The Environment

The environment consists of:

- the demand conditions for the good being produced, that is, the inverse demand function; and
- the restrictions that $a > Q$ in equilibrium and $c < a$.

1) and 2) together exhaust the 'world' of the Cournot game. Considering now what the two players know about their world, I said earlier that the game is one of 'complete information' in the terminology of game theory, meaning that the structure of the game is common knowledge to both players. This means, in the first place, that both firms know not only their own characteristics, but also those of their competitor and the market they are operating in. Knowing, here, amounts to a perfect correspondence between actors' beliefs about each aspect of their model world and the formal expression of those aspects. That is to say, actors are assumed to know their world in terms of exactly the same mathematical expressions in which it is written down (i.e. there are no ambiguities, misinterpretations, and so on, which it is of course one of the purposes of formal analysis to ensure). The model reviewed is thus very much in accordance with the so-called 'perfect knowledge' assumption of standard neoclassical theory[6]. In the second place, it means that what each firm is assumed to know about the other firm's knowledge takes a particular form. Specifically, not only does Firm *A* know everything about the world it operates in, it also knows that Firm *B* knows everything about its world, and that Firm *B* knows that Firm *A* knows everything, and so on. The same applies with respect to Firm *B*'s knowledge of what Firm *A* knows.

It will be apparent that the assumption of complete information is a demanding one, and the last thirty years or so have seen a great deal of work in economic theory aimed at investigating the consequences of relaxing it (see Stiglitz's (1994) review of the so-called 'information theoretic' approach to economic analysis). I shall attempt to

---

[6] Notwithstanding that the game is one of 'imperfect information' in the terminology of game theory, since both firms move simultaneously without knowing, and being unable to deduce, the level of output the other firm will choose. Common knowledge of the game, then, even in conjunction with common knowledge of rationality, is not enough to ensure that the firm will play their Nash strategies (although being perfect reasoners they will know that if they do not play their Nash strategies they will not be in equilibrium). But this is a wider problem that need not detain us here.

convey the flavour of this approach and the modelling strategy it adopts, by introducing a restriction on the knowledge of one of the firms in the Cournot game.

IV A GAME OF INCOMPLETE INFORMATION

I now assume that both firms know everything about their 'world' that they knew before, *except* that Firm *A* does not know Firm *B*'s costs (perhaps because *B* has developed new technology or is a recent entrant into the market). This is what I shall call a 'black-spot'. The game becomes one of incomplete information in the terminology of game theory, since the structure of the game is no longer common knowledge to both firms. As will become apparent, this version of the game breaks with the assumption of a perfect correspondence between information and knowledge in the game of complete information considered in the preceding section.

How does this informational asymmetry change the game theoretic analysis of the Cournot duopoly? In order for game theory to get going, some specific form must be given to Firm *A*'s ignorance of Firm *B*'s costs, so that in turn some form can be given to *A*'s knowledge of the payoffs it faces. The standard move here, originally suggested by Harsanyi (1967), is to suppose that while Firm *A* does not know Firm *B*'s true costs, it does know that Firm *B* is one of a given set of possible 'types' of firm, where each type is distinguished by its level of costs. For example, in the simple case to be considered here, Firm *B* is one of two possible types, either low cost or high cost. Whereas Firm *A* knew Firm *B*'s type in the example of the last section, Firm *A*'s knowledge has now been slightly reduced (the 'black-spot'). The crucial assumption here is that the contents of the 'black-spot' can be expressed probabilistically over the set of possible types. For in this way Firm *A*'s payoffs (profits) can be constructed in terms of expected profits conditional on Firm *B*'s type.

To show how this change alters the game theoretic analysis of the Cournot game, I shall now derive the Bayesian Nash equilibrium of the Cournot game under conditions of asymmetric information[7]. It is common knowledge that Firm *B* knows both its costs and

---

[7] The Bayesian Nash equilibrium requires a strategy to be specified for each possible type of player, although the reason for this need not detain us greatly. Essentially a Bayesian Nash equilibrium still requires that each player play their best response to the strategies played by their opponents. Since not all

those of Firm A. Firm A however only knows that Firm B's marginal costs are c_H with probability p and c_L with probability $(1 - p)$. This too is common knowledge. The rest of the model is as before. The firms compete by simultaneously choosing their output and the inverse demand function is as it was before:

$$P(Q) = a - Q$$

Given that Firm A's cost function is:

$$C_A(q_A) = cq_A$$

and Firm B's cost function:

$$C_B(q_B) = c_H q_B \text{ with probability } p$$
$$C_B(q_B) = c_L q_B \text{ with probability } (1 - p)$$

we can determine the maximisation problem that each type of firm faces[8]. Let $q_B^*(c_H)$ and $q_B^*(c_L)$ denote Firm B's quantity choices as a function of its cost, and let q_A^* denote Firm A's quantity choice.

If Firm B's cost is high, it chooses $q_B^*(c_H)$ to solve:

$$\max\ [(a - q_A^* - q_B) - c_H]q_B$$

If Firm B's cost is low, it chooses $q_B^*(c_L)$ to solve:

---

players know the type of their opponents, strategies for some players will be conditional on the strategies of all possible types of each of their opponents. Therefore for the equilibrium to consist of best response strategies by each player, some (or all) player's strategies must specify a plan for each of their possible types, since otherwise we cannot know whether or not all players are playing their best response. In my example, if the solution concept were not to specify a plan for each possible type of Firm B then we could not determine whether or not Firm A is playing its best response to Firm B's plan, since Firm A's profits (and so optimal choice) depend on the plans of both low and high cost type Firm Bs.

[8] With $c_L < c_H$.

$$\max\ [(a - q_A^* - q_B) - c_L]q_B$$

Firm A knows that Firm B's cost is high with probability p and should anticipate that Firm B's quantity choice will be $q_B^*(c_H)$ or $q_B^*(c_L)$, depending on its costs. Thus Firm A chooses q_A^* to solve:

$$\max\ p[(a - q_A - q_B^*(c_H)) - c]q_A + (1 - p)[(a - q_A - q_B^*(c_L)) - c]q_A$$

so as to maximise expected profit. The first-order conditions for these three optimisation problems are:

$$q_B^*(c_H) = (a - q_A^* - c_H)/2$$
$$q_B^*(c_L) = (a - q_A^* - c_L)/2$$
$$q_A^* = (a - c - pq_B^*(c_H) - (1 - p)q_B^*(c_L))/2$$

Assuming that the two firms' costs are not too different – if they are, then the high-cost firm produces nothing – these first-order conditions characterise the solutions to the earlier maximisation problems. Solving the first-order conditions gives:

$$q_B^*(c_H) = (a - 2c_H + c)/3 + (1 - p)(c_H - c_L)/6$$
$$q_B^*(c_L) = (a - 2c_L + c)/3 - p\,(c_H - c_L)/6$$
$$q_A^* = (a - 2c + pc_H + (1 - p)c_L)/3$$

which corresponds to the Bayesian Nash equilibrium of the game. Comparing the solution of this Cournot game with incomplete information, to the earlier Cournot game with complete information, we can say that a 'high-cost' ('low-cost') type Firm B produces more (less) in equilibrium when Firm A is ignorant of Firm B's type as compared with the complete information case, since Firm B knows that Firm A is maximising expected, rather than actual, profit.

As before, I wish to examine the information and knowledge contained in the game. To save repetition however, I will focus simply on what has changed now that there is an asymmetry in what the firms know. In terms of the information in the model, we simply observe that for whatever reason, Firm A conceives of there being two possible

types of Firm *B* and attaches a certain probability to the likelihood of each. The set of types and the probability distribution are information according to my definition, as is the fact that Firm *A* is an *expected* profit maximiser.

If we now consider the knowledge that the game presumes the firms to possess, Firm *B* continues to have full knowledge, in the sense that its knowledge corresponds exactly with the information in the game. Yet the scenario was designed to allow for Firm *A*'s knowledge to be limited in some regard, in this case with respect to its knowledge of Firm *B*'s costs. It is now clear how games of incomplete information restrict the limitations that can be imposed on a player's knowledge (and ignorance). Although we started with the assumption that Firm *A* was ignorant of Firm *B*'s costs, in solving the game we have that Firm *A* is far from ignorant. Rather, Firm *A* knows accurately the set of possible cost-types of Firm *B* and the probability distribution with which nature draws the particular type. All that Firm *A* suffers from is a 'black-spot' in its knowledge of its world, not knowing which type Firm *B* actually is.

This kind of approach is widely employed in the economics of information, namely to begin with a world in which all information is transparent and available to the actors concerned and then to study what happens when there is some perturbation, namely when one or more of the actors involved suffers a 'black-spot' (and where the usual move is to replace knowledge of a certain outcome or state of affairs with certain knowledge of a list of possible outcomes or states of affairs with their associated probabilities). While models of this kind are illuminating in some respects, they gloss over the differences between information and knowledge and, accordingly, the ways in which these differences may matter for the way that we think about economic phenomena. My claim then is that such a modelling approach necessarily limits the mainstream microeconomic analysis of imperfect knowledge. To develop this argument, I now move on to discuss three important aspects of economic agency that are neglected or obscured on the mainstream approach.

V UNCERTAINTY AND IGNORANCE

The first feature of the mainstream approach to information and knowledge that I consider concerns its treatment of uncertainty. Now since the points I want to raise here relate to the expected utility (EU) model of decision-making under uncertainty widely

employed in mainstream economics, I begin by briefly looking at the main features of the EU model. Within this framework an actor's decision problem is reduced to three basic elements: a set of available acts a_i (i = 1, 2, ..., m) between which an actor must choose; a set of mutually exclusive 'states of the world' s_j (j = 1, 2, ..., n), one of which must occur once a choice has been made; and a set of possible consequences, one for each act/state pair, c_{ij}. Together, these three elements give the familiar Acts/States/Consequences matrix associated with the EU model:

$$
\begin{array}{cc|cccccc|}
 & & & & \text{States} & & & \\
 & & s_1 & s_2 & . & . & . & s_n \\
\hline
 & a_1 & c_{11} & c_{12} & . & . & . & c_{1n} \\
 & a_2 & . & . & & & & . \\
\text{Acts} & . & . & . & & & & . \\
 & . & . & . & & & & . \\
 & . & . & . & & & & . \\
 & a_m & c_{m1} & c_{m2} & . & . & . & c_{mn} \\
\hline
\end{array}
$$

The preferred action of an actor within the EU model, subject to that actor satisfying certain axioms of 'rational choice'[9], is the act (or acts) which yields at least as high an expected utility as any other act, where the expected utility of an act is simply the mathematical expectation of the utility of the consequences of the act:

$$U(a_i) = \sum_{j=1}^{n} p_j u(c_{ij})$$

where p_j (j = 1, 2, ..., n) are the probabilities representing the actor's beliefs about the states of nature and $u(c_{ij})$ is a von Neumann-Morgenstern utility function, defined over all possible consequences of the decision.

Although the EU model continues to dominate mainstream analyses of decision-making under uncertainty, such a simple model is inevitably a highly stylised

[9] Economists have established various different sets of axioms that are sufficient for an actor's decision-making to be described by the EU model (Marschak 1950; Savage 1954; von Neumann and Morgenstern 1947). See Bacharach (1976), chapter 2, for a proof of how such axioms lead to the EU representation.

representation of actual human behaviour. Here I will restrict myself to commenting on the conception of uncertainty associated with the EU model, in particular focusing on the extent to which it captures the nature and severity of the kinds of uncertainties decisions-makers actually face in their day-to-day lives. I intend to concentrate on two points that are often raised in this respect: the particular way that that the EU model represents uncertainty and its neglect of an important form of ignorance.

Uncertainty in the EU model arises from the decision maker not knowing which of the possible states of the world will be realised (except in degenerate cases where the probability of some state is 1) and, hence, which consequence will actually follow from each act under consideration. This uncertainty is quantified by way of a classical probability distribution defined over the set of states, with questions concerning the nature and source of these probabilities typically ignored. In many EU models the probabilities in question are simply assumed to correspond to objective frequencies, while in other cases, where objective frequencies are not available, theorists fall back on a subjectivist or 'Bayesian' conception of probability. On the subjectivist view rational actors are simply assumed to have numerically definite personal probabilities about random events at the back of their minds at all times, which they update in a Bayesian fashion.

While some of the decision-situations we encounter in day-to-day life clearly are of the sort that involve obvious (or at least calculable) objective frequencies, such situations are actually rather rare. Much more common is the case in which people do not have access to such probabilities and are forced to fall back on their own judgements. Now this raises two main problems for the EU model and consequently for the mainstream approach that employs it. The first concerns the implausibility of the assumption that people always act 'as-if' they possess numerically definite probabilities. As Ellsberg (1961) first illustrated, even faced with choices between simple lotteries, many people exhibit behaviour that contradicts the assumption that they possess numerically definite subjective probabilities[10]. The second problem here is that many

---

[10] Two distinct theoretical approaches to accommodating such choices have emerged in the economics literature. The first, the so-called 'multiple prior' approach, drops the assumption that people possess unique subjective probabilities, in favour of supposing that they possess interval-valued probabilities (Bewley 1986; Casadesus-Masanell *et al.* 2000; Chen and Epstein 2002; Epstein and Wang 1994; Gärdenfors and Sahlin 1982; Gilboa and Schmeidler 1989, 1993; Kelsey 1993, 1994; Kelsey and Milne

important aspects of economic behaviour appear to be the consequence of people not always having reliable probabilities on which to act, and therefore not acting in the manner envisaged by the EU model. Here I have in mind certain forms of convention (Keynes 1973); rule-following, habits and behavioural routines (Cyert and March 1992; Hodgson 1997; Vanberg 1994); as well as various phenomena concerned (in the first instance at least) with betting behaviour, such as ambiguity aversion (Camerer and Weber 1992, Epstein 1999) and the competence hypothesis (Heath and Tversky 1991).

The second issue I want to raise in relation to the treatment of uncertainty in the EU model concerns the assumption that actors always know the full set of mutually exclusive possible states of the world relevant to their decision problems[11,12]. Such a requirement is likely to be difficult to fulfil in practice. For one thing, actors are unlikely to possess the degree of imagination required in order to determine the full set of possible outcomes associated with a decision problem. To compound this problem, some authors note that the consequences of a decision may not even exist as possibilities at the time of taking the decision (Shackle 1972) and therefore cannot be conceived of by actors at the time of acting. But even in situations in which actors have a reasonably good idea about the possible eventualities that may follow from their actions, by assuming that all actors know the full set of possible states of the world, the EU model ignores the fact that there remains an inevitable degree of arbitrariness where actors will draw the boundaries between one state and the next. Consequently actors are likely to 'divide up' the possible

---

1999; Kelsey and Quiggin 1992; Kyburg 1990; Levi 1986). The second approach incorporates the reliability of, or weight of evidence for, a judgement by replacing subjective probability with a non-additive capacity or weight (Dow and Werlang 1992a, 1992b; Eichberger and Kelsey 1999; Fox and Tversky 1998; Gilboa 1987; Schmeidler 1989). See Runde (2001) for a fuller discussion of some of the issues raised by these models.

[11] In game theory, at least in games of incomplete information that employ the Harsanyi transformation to convert the game into one of imperfect information, as in section IV, the states of nature consist of the possible types of each player. And as we saw in the case of the Cournot game, the full set of possible states of nature and associated probabilities are common knowledge to all players, a practice that is not unproblematic due to possible difficulties with the use of Bayesian rationality (i.e. the Savage axioms) in the context of strategic interaction (see Mariotti 1995, 1997).

[12] In principle the EU model need not require an actor to correctly know the full set of possible states of the world, merely that he have some conception of the possible states and to act on the basis of this conception. Yet this is not the way models of this sort usually proceed, since all actors are assumed to have the same, correct, conception of the set of possible states.

eventualities in quite different ways (see the discussion on the Background in the following section), contrary to the usual assumptions of the EU model.

The implication of the mainstream's neglect of these forms of non-quantifiable uncertainties and ignorance is to radically understate the extent to which indeterminacies are a feature of economic life. Now mainstream models do not eliminate indeterminacy altogether; for instance, even the simplest of games can have multiple equilibria[13]. Yet by virtue of the possibilities in these models always being pre-specified and therefore limited, the mainstream approach does rule out indeterminacies both on the scale, and of the sort, envisaged here. Contrary to the EU model, actors are not typically able to determine all possible states of the world. Rather each comes up with a more or less accurate view of what might happen, where these views are in part contingent on each actor's own situation and background. Opinions about what may happen, and about what other actors think may happen, vary considerably and are liable to sudden, unpredictable, changes. Consequently models of the sort under consideration here, and more specifically the standard assumptions of the EU model that they embody, ignore phenomena such as vague and unreliable probabilities, novelty and surprise . In turn this means that these models do not have the resources to accommodate phenomena in which vague probabilities, novelty and surprise are important: previously unimagined moves by firms in a competitive environment, the notion of entrepreneurial 'alertness' associated with the Austrian conception of the market process, the growth of knowledge, and so on[14].

VI THE SUBJECTIVITY OF KNOWLEDGE

An important characteristic of mainstream models of the sort reviewed earlier is that the actors within them know their world as, and in exactly the same mathematical terms in which, it is written down by the economist, save of course where one or more of the actors suffer some or other 'black-spot'. With this exception, actors in these models are

---

[13] The game of Heads and Tails is a good example. This game involves two players independently choosing Heads or Tails, where both of them know they will receive a prize if they both make the same choice and nothing otherwise. This game has two Nash equilibria in pure strategies (Heads, Heads) and (Tails, Tails), but there is nothing in the formal structure of the game to give the players a reason to choose either Heads or Tails.

endowed with a God's-eye view of their world, there being effectively a one-to-one relation between that world as it is and what they know about it.

One thing that this conception ignores is that our knowledge of our circumstances is in general highly subjective, being dependent on our intentions and desires, past experiences and knowledge, geographical and temporal location, physical and emotional state, and so on[15]. As a result, our knowledge of the world rarely corresponds to the world as it actually is in a simple one-to-one way. While the 'black-spot' manoeuvre could be interpreted as a first step towards acknowledging this subjectivity, since it at least allows the knowledge of actors in these models to differ, the nature, causes and effects of the subjectivity of knowledge do not significantly contribute to mainstream analyses. Now the range of possible issues I might consider here is vast and as such I intend to focus exclusively on one important factor in the subjectivity of knowledge, namely the structured nature of all conscious experience. To this end I draw on the philosopher John Searle's (1983, 1994a, 1995, 1999) work on consciousness and intentionality[16]. While not uncontroversial in some quarters, I regard Searle's work as providing a clear and commonsensical account on which to base a discussion of the subjectivity of knowledge.

Searle's account of the structured nature of conscious experience starts from the idea that many conscious mental states are intentional[17], in the sense that they are intrinsically about, or directed towards, something or someone. Thus you may intend to

---

[14] On this, see in particular Kirzner (1973, 1979, 1985, 1989).

[15] Within economics, this issue is closely related to the much broader issue of subjectivism (see Horwitz 1994, Kirzner 1995). A key tenet of the Austrian school and the writings of Menger, von Mises, Hayek, Knight and Shackle, subjectivists argue that any adequate explanation of human interactions and institutions must start with the subjective meaning that individuals attach to their actions, that is, must begin from the subjective mental states of the actors being studied. From this starting-point, Kirzner distinguishes three quite different expressions of subjectivism in modern economic thought; that found in mainstream economics, the radical subjectivism of Shackle and Lachmann, and the more moderate subjectivism of modern Austrian thought. The particular issue under consideration here, based on the distinction between actors perceptions of reality and that reality itself, and subsequently the roles played by context and interpretation in economic affairs, is central to both the second and third of these expressions, although it need not imply some of the more nihilistic findings associated with the Shackle-Lachmann view.

[16] A similar perspective is supported by work on perceptual schema and cognitive framing in cognitive theory (e.g. Bruner 1973, Lloyd 1972) and recent work in psychology (Clark 1997; Cosmides and Tooby 1994b; Margolis 1994; Plotkin 1994), which emphasises the role that context plays in human reasoning.

[17] But not all. For instance, one may have conscious feelings of anxiety that have no specific focus or target.

purchase a share, remember that you purchased a share, wish that you had purchased a share, and so on. Intending, remembering and wishing are all intentional states by virtue of the fact that they are necessarily directed at something, in this case the purchase of a share. Many intentional states, particularly those whose content is a proposition, have what Searle calls conditions of satisfaction, namely the conditions that must be met for an intention to be fulfilled, a belief to be true, and so on. Thus the condition of satisfaction of my belief that I will buy a new bicycle tomorrow is that I do in fact do so.

By virtue of intrinsically referring to, or being about, other things, intentional states are considered to be crucial by Searle in mediating the mind's causal relations with the external world. One aspect of this is the mechanism by which intentional mental states function causally in rational behaviour, what Searle terms the mechanism of intentional causation. The central proposition here is that intentional states function during rational behaviour by virtue of their contents being consciously scrutinised so as to establish their conditions of satisfaction. For example, a desire functions causally as a result of being consciously examined to determine what actions must be carried out in order for it to be fulfilled.

Searle then notes that in general intentional states never function in isolation. Instead they depend firstly on a network of other intentional states. My desire to purchase a new bicycle for instance only functions in association with my knowing the location of a nearby bicycle shop and how much I can afford to spend. But the functioning of these intentional states also depends on non-intentional capabilities or what Searle terms an actor's Background: a reservoir of non-intentional 'capacities, abilities, tendencies, habits, dispositions, taken-for-granted presuppositions and "know-how" generally' (Searle 1999: 107-108). In contrast to intentional states, Background capacities operate without conscious intervention or reflection. As Searle would have it then, conscious intentionality consists of thought processes that work in the way they do relative to a Background of unthought capacities, skills and know-how.

Rather than prove the existence of the Background, Searle (1983: 145) instead prefers to demonstrate his thesis by the accretion of examples. His favourite ones are linguistic in nature, and specifically deal with how the Background enables linguistic interpretation. Consider the three statements:

Tom opened the door.

Sam opened his book at page 37.

The surgeon opened the wound.

While each sentence shares the same basic form, 'X opened Y', what we understand by the word 'opened' in each case clearly differs. To be told that Tom has opened the door does not lead us to imagine Tom in an operating theatre carefully cutting into the door with a scalpel. Nor do we imagine the surgeon looking for a door handle. Instead, we correctly determine the meaning of each sentence without any conscious act of interpretation. Searle proposes that this ability is one instance of our Background capabilities.

The functioning of the Background is not restricted to fixing the semantic content of sentences. Searle maintains that all intentional states only function against a set of non-intentional Background skills. While the determination of the conditions of satisfaction of an intentional state firstly requires other intentional states, the process of drawing consciously on additional intentional states does not, and cannot, go on forever. At some point the process of intentional causation always 'bottoms out' on non-intentional, non-conscious, Background capacities which contribute to the determination of the intentional states conditions of satisfaction but that require no conscious intervention.

Turning now to consider the structured nature of conscious experience, Searle argues that all intentional mental states possess 'aspectual shape' (Searle 1994a: 156-157), which refers to the fact that whenever we perceive or think about something, we necessarily do so under particular aspects and not others. But to experience something *as* something, that is experience something under a certain aspect, we must already be familiar with the set of categories under which we experience that aspect. Thus our ability to recognise a piece of paper as an invoice presupposes that we already have the category of 'invoice'. The same goes for most of the objects and events that fill our everyday life, which we recognise immediately and without apparent effort. Searle regards the knowledge, and ability to apply, such categories as a Background ability.

The structuring of consciousness by the Background is not limited to individual objects or momentary events. The Background also conditions our experiences in a dynamic way, by providing perceptual and linguistic categories that extend over sequences of events and which allow us to structure those sequences into comprehensible narrative shapes (Searle 1995: 134-135). These 'dramatic' categories allow us to form

what Searle calls 'scenarios of expectations', our taken-for-granted conceptions of how certain kinds of situations typically unfold, such as our negotiating to purchase a bicycle or our paying for a meal at a restaurant. Thus we become accustomed to and take for granted certain courses of events as following on from each other.

That the structuring of conscious experience is a significant cause of the subjectivity of knowledge follows from the fact that many of the categories we draw on are learned rather than being innate. Consequently these categories are a function of the social and cultural environment in which we have grown up and the way in which these past experiences have shaped our Background. As a result, although we all have many shared sets of categories, we each of us have also acquired significantly different sets of categories that we draw on. Now once we recognise that each of us interprets or structures the information we receive in different ways, a range of phenomena become possible that the mainstream approach excludes by virtue of assuming a strict one-to-one relationship between information and knowledge.

Perhaps the simplest illustration of the kind of effect that I have in mind here is provided by the example of two actors who may have quite different experiences of the same event. For instance suppose the event concerned is a controversial incident at a football match or the latest stock market announcement of a listed company. In the first case an actor's view of the event will in part depend on the football team that they support; while in the second the knowledge derived from the company's statement is likely to depend on whether an actor holds a short or long position in the company concerned, or otherwise has prior expectations about the outlook for the company.

Recently, insights of this kind have begun to attract the attention of some economists, albeit those from outside of the mainstream. The notion that individuals engage in mental accounting (Thaler 1980, 1999), operating a set of rules by which they organise, evaluate and keep track of their financial activities; or that individuals implement a set of basic editing operations when faced with simple gambles (Kahneman 2000; Kahneman and Tversky 1979) prior to evaluating the available alternatives; incorporates the idea that the set of rules and categories that people implement is partly subjective and that their knowledge of these rules and implementation of them is in part a tacit, Background operation. The related finding that individuals tend to be susceptible to 'framing effects', being unconsciously influenced by the presentation of a decision-situation (Camerer 2000; Shafir, Diamond and Tversky 1997; Tversky and Kahneman

1981, 1986) also appears to presuppose an account of conscious experience similar to that portrayed here (Faulkner 2002a).

VII TACIT KNOWLEDGE

The final feature of the mainstream modelling approach that I wish to highlight is that the actors within these models are assumed to form comprehensive, conscious representations of their decision-situations and, on the basis of these representations, to deliberate consciously (and perfectly) about the best way to proceed in those situations. The associated conception of knowledge is of that which resides in propositional form, that is, a form suited to the direct application of discursive reasoning. Indeed, as I showed above, this is exactly the form in which knowledge is 'given' to the economic actors in the models concerned.

By interpreting mainstream models in this way, as descriptions of the way economic actors actually behave, I recognise that I am reading more into these models than many economists would be prepared to. Some economists hold the view that formal models such as these should only be viewed in terms of their predictive adequacy and that they are not meant as descriptions of what people actually do (Friedman 1953). Others view such models in normative terms, as showing what a rational actor should do in a given situation. Yet even if these alternative interpretations are appropriate to certain models in certain situations, it is hard not to interpret the sort of models that I consider here in a descriptive way. The objective of much of game theory is to model the sorts of situations that real economic actors find themselves in and to analyse them by reconstructing the discursive reasoning of actors about what constitutes their best course of action given their expectations about what the other actors will do. The Cournot game examined earlier is just such a model, where the focus is on how a firm engaged in Cournot competition will react when its knowledge of its competitors costs is less than perfect.

The key point that I wish to make in this section is that economists' obsession with formal rational choice models, and the exclusive appeal to conscious, propositional knowledge that this entails, has led to a neglect of tacit knowledge in mainstream economics. By tacit knowledge I mean knowledge in the form of the many skills, capacities and dispositions that human actors routinely draw on without conscious

reflection. In contrast to propositional knowledge, which operates at the foreground of the mind and tends to be readily codifiable and communicable, tacit knowledge is often difficult to express and need not be consciously thought about when enacted. In my view mainstream rational choice theory deflects attention from, first, the extent to which human behaviour relies on tacit rather than conscious functioning, and second, the fact that conscious, propositional knowledge always functions in conjunction with tacit knowledge. I shall consider these points in turn, beginning with the idea that human behaviour is driven by non-conscious skills, routines and habits as much as it is by conscious beliefs and desires. The point is quite widely recognised in non-mainstream contributions (e.g. Hayek 1948; Polanyi 1967; Nelson and Winter 1982; Reber 1993; Nonaka and Takeuchi 1995; Lawson 1997; Ambrosini 2003) but rather than reviewing this literature, I shall instead focus in some detail on one particular example, that of heuristic based judgement.

The idea that actors frequently employ heuristics, or simple rules of thumb, to guide their behaviour first emerges in economics in Simon's work on satisficing (Simon 1956). More recently the idea of heuristic-based judgement has received considerable attention in behavioural finance as an explanation for the persistent biases observed in individuals' probability judgements (Kahneman, Slovic and Tversky 1982; Shefrin 2000; Warneryd 2001). One common heuristic is availability (Tversky and Kahneman 1982; Slovic, Fischhoff and Lichtenstein 1982), according to which actors tend to judge the likelihood of an event based on the ease with which instances of that event can be brought to mind. An investor assessing the likelihood of a particular company going bankrupt for example, is often guided by the number of similar company failures that he can recall.

The notion of availability serves as a useful example for my purposes because such heuristics function tacitly and, as such, exemplifies a number of points about tacit knowledge in general. The tacit nature of the heuristic lies in the fact that, rather than consciously deciding to focus only on information concerning how easy it is to recall or imagine possible instances of some class of event, the actor's conscious mind is 'automatically' focussed on information relating only to availability. In effect heuristics, and tacit knowledge in general, serve to economise on the conscious effort required of an actor, by concentrating the conscious mind on only the most important, rather than all, aspects of behaviour.

The benefit to heuristic-guided judgement comes from the fact that the more complicated the actor's environment, the better the actor is likely to perform by following

a heuristic rather than by consciously reasoning in accordance with the Savage axioms or some such[18]. The reason for this lies in the boundedness of human mental capacities. Suppose that an actor has complete information on which to make a judgement. Given limited computational power (be this for lack of time or ability) an actor may make more accurate judgements on average by being predisposed towards focussing only on information concerning availability rather than performing an incomplete or inaccurate Bayesian calculation, attempting but failing to correctly incorporate all information[19]. The situation becomes even more favourable towards the heuristic-guided actor when complete information is not freely available. For, before any calculation, actors must decide how much effort to put into searching for information. The heuristic-guided actor immediately focuses on availability, whereas the Bayesian actor must decide in which direction to search by comparing the costs of search relative to its (unknown) benefits.

Against the advantages associated with economising on conscious effort, the disadvantages to tacit knowledge follow from the elimination of complete conscious reflection. Again our example of heuristic-guided judgement demonstrates the point. The disadvantage to relying on heuristics is that the resulting judgements will tend to be biased. The availability heuristic, for instance, is susceptible to the fact that factors other than frequency and probability affect the ease with which instances of an event can be brought to mind. The likelihood of an event which has received widespread media attention, or that is particularly salient to an individual by virtue of past experience, is likely to be consistently overstated by an individual forming judgements in accordance with availability.

The example of judgmental heuristics also suggests a number of additional observations that may be made about tacit knowledge and rational choice models. First, and as they are usually presented, the actors in these models would gain nothing from

---

[18] Evolutionary psychology (e.g., Cosmides and Tooby 1994a) notes that since certain of the heuristics that we employ may reflect inherited computational devices tailored to problems faced by past generations, not all the heuristics we now employ need be beneficial with respect to our present circumstances.

[19] Such a situation might be termed the curse of information, since an actor presented with all relevant information may make worse judgements on average than an actor who only has available to him a subset of this information. A related but distinct phenomena is the curse of knowledge (Camerer, Loewenstein and Weber 1989; Camerer 1995) in which better informed actors in asymmetric information models such as the earlier Cournot game are unable to accurately imagine being in the position of the less well-informed actors because the extra knowledge cannot simply be ignored.

employing heuristics since they face no computational constraints. Indeed with the costs (if any) to acquiring information known and unlimited conscious computational abilities, actors have no need for any type of tacit knowledge since they would gain nothing by restricting their conscious effort. Therefore mainstream economists neglect of tacit knowledge is at least understandable given the (admittedly unrealistic) assumptions of their models.

A more sophisticated variant of rational choice model (Heiner 1983) acknowledges the possible utility to actors of adhering to simple rules, by portraying rule-following as an actor's rational response to their own cognitive limitations. In effect these models simply increase the number of constraints faced by an actor to include search and processing costs, with the result that a (boundedly) rational actor may choose to adopt a heuristic. The problem with this approach remains the fact that all knowledge in these models is propositional and all functioning, intentionalistic. Yet if an actor is to be guided by a heuristic he must do so unthinkingly in each decision-situation, which requires an account of tacit functioning. So while it is appropriate to identify heuristics as being beneficial to an actor because they enable conscious effort to be automatically directed, these models must also recognise heuristics as functioning tacitly in order for them to perform this function.

The second theme I pursue in this section concerns the role of tacit knowledge in the functioning of propositional knowledge. Here I again refer to John Searle's work on intentionality to illustrate the argument. In section VI, I introduced the notion that many conscious mental states, including those here conceived of as propositional knowledge, are intentional and that these states function causally in rational behaviour in accordance with the mechanism of intentional causation. According to this mechanism, an intentional state functions causally by virtue of being consciously scrutinised. Now I wish to make two points in relation to the idea of intentional causation. The first is to recognise that mainstream rational choice models deal with an idealised form of intentional causation (Runde 2002), since if we accept that these models portray an actor consciously choosing a preferred course of action given their wants and beliefs then that actor is behaving intentionalistically. My second point here is that these models eliminate the role that Searle attributes to the Background in the functioning of intentional states.

As discussed earlier, Searle finds that rather than functioning in isolation, intentional states always function in conjunction with a Network of other intentional states and a Background of non-intentional capacities. In effect the Background, which

consists of non-conscious capacities, skills and know-how, provides the backdrop against which all intentionalistic functioning, including rational decision-making, takes place. In the case of my desire to purchase a quantity of shares in a particular company, a considerable amount of tacit knowledge concerning how to conduct share transactions is implicated in the functioning of my desire, just as is my propositional knowledge concerning the particular market on which that share is listed. By ignoring the role of tacit knowledge in the functioning of propositional knowledge, the mainstream approach takes our attention away from the importance in purposeful, intentionalistic, behaviour of the things that we typically have tacit knowledge of, such as the rules and structures that influence our social interactions (Lawson 1997), the complex technical skills that we draw on (Nelson and Winter 1982) or the habits we develop (Hodgson 1988).

VIII CONCLUSION

Prior to the information theoretic revolution, mainstream microeconomics paid scant attention to what actors could actually be presumed to know about their environment. The shift in emphasis brought about by the information revolution has led to explicit consideration of the effects of imperfect and asymmetric knowledge on the functioning of markets, as a result of which mainstream views regarding the properties of such markets have changed dramatically.

Yet if the position I defend in this paper is correct, then despite having successfully focused the attention of mainstream economists on certain aspects of human knowledge, the nature of the mainstream project limits the extent to which issues concerning information, knowledge and ignorance can be adequately attended to within a mainstream framework. My argument here has proceeded by looking at three particular areas of economic agency that, while apparently central to the issue of information and knowledge, I believe to be neglected by the mainstream approach. First of all I drew attention to the expected utility model of decision-making, which I argued led to a conception of uncertainty in mainstream models that necessarily understates the role of novelty and surprise in economic behaviour. I then showed that the subjectivity that characterises much of what human beings know of their circumstances is excluded from these models by virtue of the perfect correspondence between information and knowledge, which, barring 'black-spots', these models assume. Finally I argued that by

restricting all knowledge to be of an intentional form (beliefs and preferences), the role that tacit knowledge plays in behaviour is suppressed. In short then, while revolutionary in terms of the mainstream microeconomic project, the information theoretic revolution represents only a partial step towards the incorporation of the full complexities of human knowledge into economics.

BIBLIOGRAPHY

Agonito, R. (1975), 'Hayek Revisited: Mind as the Process of Classification', *Behaviorism*, Vol. 3, No. 2, 162-71.

Alchian, A. (1950), 'Uncertainty, Evolution and Economic Theory', *Journal of Political Economy*, Vol. 58, No. 3, 211-22.

Ambrosini, V. (2003), *Tacit and Ambiguous Resources as Sources of Competitive Advantage*, Basingstoke: Palgrave Macmillan.

Archer, M. (1995), *Realist Social Theory: The Morphogenetic Approach*, Cambridge: Cambridge University Press.

Archer, M. (1996), *Culture and Agency: The Place of Culture in Social Theory*, Cambridge: Cambridge University Press.

Archer, M. (2000), *Being Human: The Problem of Agency*, Cambridge: Cambridge University Press.

Archer, M., Bhaskar, R., Collier, A., Lawson, T., and Norrie, A. (1998), *Critical Realism: Essential Readings*, London: Routledge.

Ardila, R. and Bunge, M. (1987), *Philosophy of Psychology*, New York: Springer.

Arrow, K, Colombatto, E., and Schmidt, C. (1996), *The Rational Foundations of Economic Behaviour*, IEA Conference Vol. 114, London: Macmillan.

Augier, M. (2001), 'Sublime Simon', *Journal of Economic Psychology*, Vol. 22, No. 3, 307-334.

Augier, M. and March, J. (2002), 'A Model Scholar', *Journal of Economic Behavior and Organisation*, Vol. 49, No. 1, 1-17.

Bacharach, M. (1976), *Economics and the Theory of Games*, London: Macmillan.

Benartzi, S. and Thaler, R. (1995), 'Myopic Loss Aversion and the Equity Premium Puzzle', *Quarterly Journal of Economics*, Vol. 110, No. 1, 73-92.

Benartzi, S. and Thaler, R. (2001), 'Naïve Diversification Strategies in Defined Contribution Saving Plans', *American Economic Review*, Vol. 91, No. 1, 79-98.

Bewley, T. (1986), 'Knightian Decision Theory: Part 1', Cowles Foundation Discussion Paper, No. 807.

Bhaskar, R. (1978), *A Realist Theory of Science*, Hemel Hempstead: Harvester Press.

Bhaskar, R. (1979), *The Possibility of Naturalism*, Hemel Hempstead: Harvester Press.

Bindra, D. (1976), *A Theory of Intelligent Behavior*, New York: Wiley Interscience.

Birnbaum, M., Patton, J. and Lott, M. (1999), 'Evidence Against Rank-Dependent Utility Theories: Tests of Cumulative Independent, Interval Independence, Stochastic Dominance, and Transitivity', *Organizational Behaviour and Human Decision Processes*, Vol. 77, 44-83.

Birner, J. (1994), 'Comment', in Colonna, M., Hagemann, H. and Hamouda, O. (eds.), *The Economics of F. A. Hayek: Volume 2*, 201-206, Aldershot: Edward Elgar.

Birner, J. (1999), 'The Surprising Place of Cognitive Psychology in the Work of F.A.Hayek', *History of Economic Ideas*.

Black, F. (1986), 'Noise', *Journal of Finance*, Vol. 41, No. 3, 529-43.

Blaug, M. (1992), *The Methodology of Economics*, 2nd ed., Cambridge: Cambridge University Press.

Bratman, M. (1993), 'Shared Intention', *Ethics*, Vol. 104, 97-113.

Bratman, M. (1999), *Faces of Intention*, Cambridge: Cambridge University Press.

Bruner, J. (1973), *Beyond the Information Given*, Boston: Martinus Nijhoff.

Bunge, M. (1980), *The Mind-Body Problem: A Psychobiological Approach*, Oxford: Pergamon.

Camerer, C. (1995), 'Individual Decision Making', in Kagel, J and Roth, A. (eds.) *The Handbook of Experimental Economics*, Princeton: Princeton University Press, 587-703.

Camerer, C. (2000), 'Prospect Theory in the Wild: Evidence from the Field', in Kahneman, D., and Tversky, A. (eds.) *Choices, Values and Frames*, Cambridge: Cambridge University Press, 288-300.

Camerer, C., Loewenstein, G., and Weber, M. (1989), 'The Curse of Knowledge in Economic Settings: An Experimental Analysis', *Journal of Political Economy*, Vol. 97, No. 5, 1232-54.

Camerer, C. and Weber, M. (1992), 'Recent Developments in Modelling Preferences: Uncertainty and Ambiguity', *Journal of Risk and Uncertainty*, Vol. 5, 325-370.

Casadesus-Masanell, R., Klibanoff, P., and Ozdenoren, E. (2000), 'Maxmin Expected Utility over Savage Acts with a Set of Priors', *Journal of Economic Theory*, Vol. 92, No. 1, 35-65.

Chen, Z. and Epstein, L. (2002), 'Ambiguity, Risk, and Asset Returns in Continuous Time', *Econometrica*, Vol. 70, No. 4, 1403-43.

Chipman, J., Hurwicz, L., Richter, M. and Sonnenschein, H. (eds) (1971), *Preferences, Utility and Demand*, New York: Harcourt Brace Jovanovich.

Chomsky, N. (1975), *Reflections on Language*, New York: Pantheon.

Chomsky, N. (1986), *Knowledge of Language: Its Nature, Origin and Use*, New York and Philadelphia: Praeger Special Studies.

Churchland, Patricia S. (1981), 'Eliminative Materialism and the Propositional Attitudes', *Journal of Philosophy*, Vol. 78, 67-90.

Churchland, Patricia S. (1986), *NeuroPhilosophy: Toward a Unified Science of the Mind-Brain*, Cambridge, MA: MIT Press.

Churchland, Patricia S. (1987), 'Epistemology in the Age of Neuroscience', *The Journal of Philosophy*, Vol. 84: 544-53.

Churchland, Paul M. (1981), 'Eliminative Materialism and the Propositional Attitudes', *Journal of Philosophy* Vol. 78: 67-90.

Churchland, Paul M. (1984), *Matter and Consciousness*, Cambridge, MA: MIT Press.

Churchland, Paul M. (1992), *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*, Cambridge, MA: MIT Press.

Clark, A. (1997), *Being There: Putting the Brain, Body and World Together Again*, Cambridge, MA: MIT Press.

Cosmides, L. and Tooby, J. (1994a), 'Better than Rational: Evolutionary Psychology and the Invisible Hand', *American Economic Review*, Vol. 84, No. 2, 337-332.

Cosmides, L. and Tooby, J. (1994b), 'Beyond Intuition and Instinct Blindness: Toward an Evolutionarily Rigorous Cognitive Science', *Cognition*, Apr. – June 1994, 41-77.

Cottrell, A. (1995), 'Intentionality And Economics', *Economics and Philosophy*, 11: 159-176.

Crick, F. (1994), *The Astonishing Hypothesis: The Scientific Search for the Soul*, New York: Scribner.

Cubeddu, R. (1993), *The Philosophy of the Austrian School*, London: Routledge.

Cyert, R. and March, J. (1992), *A Behavioural Theory of the Firm*, Oxford: Blackwell.

Davidson, D. (1980), *Essays on Actions and Events*, Oxford: Oxford University Press.

Davis, J. (2002), 'Collective Intentionality and Individual Behaviour', in Fullbrook, E. (ed.) *Intersubjectivity in Economics: Agents and Structures*, London: Routledge.

Dietrich, E. (ed.) (1994), *Thinking Computers and Virtual Persons*, San Diego: Academic Press.

De Bondt, W. and Thaler, R. (1985), 'Does the Stock Market Overreact?', *Journal of Finance*, Vol. 40, No. 3, 793-805.

De Long, B., Shleifer, A., Summers, L., and Waldman, R. (1990), 'Noise Trader Risk in Financial Markets', *Journal of Political Economy*, Vol. 98, No. 4, 703-38.

Dempsey, G. (1996), 'Hayek's *Terra Incognita* of the Mind', *Southern Journal of Philosophy*, Vol. XXXIV, No. 1, 13-41.

Dennett, D. (1987), *The Intentional Stance*, Cambridge, MA: MIT Press.

Dennett, D. (1991), *Consciousness Explained*, Boston: Little, Brown and Company.

De Vries, R. (1994), 'The Place of Hayek's Theory of Mind and Perception in the History of Philosophy and Psychology', in *Hayek, Co-ordination and Evolution*, ed. J. Birner and R. van Zijp, 311-322, London and New York: Routledge.

Diecidue, E. and Wakker, P. (2001), 'On the Intuition of Rank-Dependent Utility', *Journal of Risk and Uncertainty*, Vol. 23, 281-298.

Dow, J. and Werlang, S. (1992a), 'Excess Volatility of Stock Prices and Knightian Uncertainty', *European Economic Review*, Vol. 36, No. 2-3, 631-638.

Dow, J. and Werlang, S. (1992b), 'Uncertainty Aversion, Risk Aversion and the Optimal Choice of Portfolio', *Econometrica*, Vol. 60, No. 1, 197-204.

Earl, P. (2002), 'Introduction', in P. Earl (ed.), *The Legacy of Herbert Simon in economics: Volume 1*, Aldershot: Edward Elgar.

Edelman, G. (1989), *The Remembered Present: A Biological Theory of Consciousness*, New York: Basic Books.

Edelman, G. (1992), *Bright Air, Brilliant Fire: On the Matter of the Mind*, New York: Basic Books.

Eichberger, J. and Kelsey, D. (1999), 'E-Capacities and the Ellsberg Paradox', *Theory and Decision*, Vol. 46, No. 2, 107-40.

Ellsberg, D. (1961), 'Risk, Ambiguity and the Savage Axioms', *Quarterly Journal of Economics*, Vol. 65, No. 4, 643-669.

Elster, J. (1998), 'Emotions and Economic Theory', *Journal of Economic Literature*, Vol. 36, 47-74.

Elster, J. (1999), *Alchemies of the Mind: Rationality and the Emotions*, Cambridge: Cambridge University Press.

Epstein, L. (1999), 'A Definition of Ambiguity Aversion', *Review of Economic Studies*, Vol. 66, No. 3, 579-608.

Epstein, L. and Wang, T. (1994), 'Intertemporal Asset Pricing Under Knightian Uncertainty', *Econometrica*, Vol. 62, No. 2, 283-322.

Faulkner, P. (2002a), 'The human agent in behavioural finance: a Searlean perspective', *The Journal of Economic Methodology*, Vol. 9, No. 1, 31-52.

Faulkner, P. (2002b), 'Some problems with the conception of the human subject in critical realism,' *Cambridge Journal of Economics*, Vol. 26, No. 6, 739-751.

Fleetwood, S. (1995), *The Political Economy of Hayek: The Socio-economics of Order*, London: Routledge.

Fleetwood, S. (1997), 'Hayek III: The Necessity of Social Rules of Conduct', in Frowen, S. (ed.), *Hayek: Economist and Social Philosopher*, 155-178, Basingstoke: Macmillan Press.

Fleetwood, S. (1999), *Critical Realism in Economics: Development and Debate*, London: Routledge.

Fodor, J. (1975), *The Language of Thought*, New York: Crowell.

Ford, J. (1994), *G.L.S Shackle: The Dissenting Economist's Economist*, Aldershot: Edward Elgar.

Fotion, N. (2000), *John Searle*, Teddington: Acumen Publishing.

Fox, C. and Tversky, A. (1998), 'A Belief-Based Account of Decision Under Uncertainty', *Management Science*, Vol. 44, No. 7.

Frederick, S., Loewenstein, G. and O'Donoghue, T. (2002), 'Time Discounting and Time Preference: A Critical Review', *Journal of Economic Literature*, Vol. 40, No. 2, 351-401.

Friedman, M. (1953), 'The Methodology of Positive Economics', in *Essays in Positive Economics*, Chicago: University of Chicago Press, 3-43.

Gärdenfors, P. and Sahlin, N. (1982), 'Unreliable Probabilities, Risk Taking and Decision Making', *Synthese*, Vol. 53, 361-386.

Giddens, A. (1976), *Central Problems in Social Theory*, Berkeley: University of California Press.

Giddens, A. (1984), *The Constitution of Society*, Cambridge: Polity Press.

Gifford, A. (2002), 'Emotion and self-control', *Journal of Economic Behavior and Organization*, Vol. 49, 113-130.

Gilbert, M. (1989), *On Social Facts*, London: Routledge.

Gilboa, I. (1987), 'Expected Utility Theory with Purely Subjective Non-Additive Probabilities', *Journal of Mathematical Economics*, Vol. 16, No. 1, 141-153.

Gilboa, I. and Schmeidler, D. (1989), 'Maxmin Expected Utility with a Non-Unique Prior', *Journal of Mathematical Economics*, Vol. 18, No. 2, 141-153.

Gilboa, I. and Schmeidler, D. (1993), 'Updating Ambiguous Beliefs', *Journal of Economic Theory*, Vol. 59, No. 1, 33-49.

Gilovich, T, Griffin, D. and Kahneman, D. (2002), *Heuristics and Biases: The Psychology of Intuitive Judgment*, New York and Melbourne: Cambridge University Press.

Gneezy, U. and Potters, J. (1997), 'An Experiment on Risk Taking and Evaluation Periods', *Quarterly Journal of Economics*, Vol. 112, No. 2, 631-646.

Gonzalez-Vallejo, C. (2002), 'Making Trade-Offs: A Probabilistic and Context-Sensitive Model of Choice Behaviour', *Psychological Review*, Vol. 109, 137-154.

Gravelle, H. and Rees, R., (1992), *Microeconomics*, London: Longman.

Haltiwanger, J. and Waldman, M. (1985), 'Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity', *American Economic Review*, Vol. 75, No. 3, 326-40.

Hands, W. (1993), *Testing, Rationality, and Progress: Essays on the Popperian Tradition in Economic Methodology*, Lanham, MD: Rowman & Littlefield.

Hands, W. (1998), 'Scientific Explanation' in J. Davis, W. Hands and Mirowski. P (eds.), *The Handbook of Economic Methodology*, 439-43. Cheltenham: Edward Elgar.

Hands, W. (2001), *Reflection without Rules: Economic Methodology and Contemporary Science Theory*, Cambridge: Cambridge University Press.

Harsanyi, J. (1967), 'Games with Incomplete Information Played by Bayesian Players Parts I II and III', *Management Science* Vol. 14, 159-82, 320-34, 486-502.

Hatsopoulous, G., Krugman, P. and Porterba, J. (1989), 'Overconsumption: The challenge to U.S economic policy', Paper presented at the American Business Conference 1989.

Hayek, F. (1937), 'Economics and Knowledge', *Economica*, Vol. 4, 33-54.

Hayek, F. (1945), 'The Use of Knowledge in Society', *The American Economic Review*, Vol. xxxv, no. 4.

Hayek, F. (1948), *Individualism and Economic Order*, Chicago: University of Chicago Press.

Hayek, F. (1952), *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*, London: Routledge and Kegan Paul.

Hayek, F. (1960), *The Constitution of Liberty*, London: Routledge.

Hayek, F. (1967), 'Rules, Perception and Intelligibility', in *Studies in Philosophy, Politics and Economics*, London: Routledge and Kegan Paul.

Hayek, F. (1973), 'Rules and Order', in *Law, Legislation and Liberty*, London: Routledge and Kegan Paul.

Hayek, F. (1983), *Knowledge, Evolution and Society*, London: Adam Smith Institute.

Hayek, F. (1988), *The Fatal Conceit*, London: Routledge.

Heath, C. and Soll, J. (1996), 'Mental Accounting and Consumer Decisions', *Journal of Consumer Research*, Vol. 23, 40-52.

Heath, C. and Tversky, A. (1991), 'Preference and Belief: Ambiguity and Competence in Choice Under Uncertainty', *Journal of Risk and Uncertainty*, Vol. 4, 5-28.

Hebb, D. (1949), *The Organisation of Behavior*, New York: Wiley

Heiner, R. (1983), 'The Origin of Predictable Behaviour', *American Economic Review*, Vol. 73, No. 4, 560-95.

Heisler, J. (1994), 'Loss Aversion in a Futures Market: An Empirical Test', *The Review of Futures Markets*, Vol. 13, No. 3, 793-822.

Hempel, C. (1949), 'The Logical Analysis of Psychology', in H. Feigl and Sellars, W. (eds.), *Readings in Philosophical Analysis*, 373-84, New York: Appleton Century Crofts.

Hicks, J. (1946), *Value and Capital*, 2nd ed., Oxford: Clarendon.

Hirstein, W. (2001), *On Searle*, Belmont, CA: Wadsworth/Thomson Learning.

Hodgson, G. (1988), *Economics and Institutions*, Cheltenham, Polity Press.

Hodgson, G. (1994), 'Optimisation and Evolution: Winter's critique of Friedman revisited', *Cambridge Journal of Economics*, Vol. 18, No. 4, 413-430.

Hodgson, G. (1997), 'The Ubiquity of Habits and Rules', *Cambridge Journal of Economics*, Vol. 21, No. 6, 663-684.

Hodgson, G. (1998), 'The Approach of Institutional Economics', *Journal of Economic Literature*, Vol. 36, No. 1, 166-192.

Hodgson, G. (1999a), *Evolution and Institutions*, Cheltenham, Edward Elgar Press.

Hodgson, G. (1999b), 'Structures and Institutions: Reflections on Institutionalism and Critical Realism', unpublished paper presented at the Workshop on Realism and Economics (January 1999), King's College, Cambridge.

Hodgson, G. (2002), 'Reconstitutive downward causation: Social structure and the development of individual agency', in Fullbrook, E. (ed.) *Intersubjectivity in Economics: Agents and Structures*, London: Routledge.

Hodgson, G. (2003), 'The hidden persuaders: institutions and individuals in economic theory', *Cambridge Journal of Economics*, 27: 159-175.

Hodgson, G., Samuels, W. and M. Tool (1994), *The Elgar Companion to Institutional and Evolutionary Economics*, Aldershot: Edward Elgar.

Hogarth, R. (1987), 'Decision Making under Ambiguity', in Hogarth, R. and Reder, M. (eds.) *Rational Choice: The Contrast Between Economics and Psychology*. Chicago: Chicago University Press, 41-67.

Hoover, K. (2001), 'Is macroeconomics for real?', in Mäki, U. (ed.) *The Economic World View: Studies in the Ontology of Economics*, Cambridge: Cambridge University Press.

Horwitz, S. (1994), 'Subjectivism', in P. Boettke (ed.), *The Elgar Companion to Austrian Economics*, 17-22, Aldershot: Edward Elgar.

Ingham, G. (1996), 'Money is a Social Relation', *Review of Social Economy*, Vol. 54, No. 4, 507-30.

Joas, H. (1993), *Pragmatism and Social Theory*, Chicago: University of Chicago Press.

Joas, H. (1996), *The Creativity of Action*, Cambridge: Polity Press.

Kahneman, D. (2000), 'Preface', in Kahneman, D and Tversky, A. (eds.) *Choices, Values and Frames*, ix-xvii, Cambridge: Cambridge University Press.

Kahneman, D., Slovic, P., and Tversky, A. (1982), *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge, England: Cambridge University Press.

Kahneman, D. and Tversky, A. (1974), 'Judgment Under Uncertainty: Heuristics and Biases', *Science*, Vol. 185, 1124-1131.

Kahneman, D. and Tversky, A. (1979) 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica*, Vol. 47, No. 2, 263-291.

Kahneman, D. and Tversky, A. (1984), 'Choices, Values and Frames', *American Psychologist*, Vol. 39, No. 4, 341-350.

Kahneman, D. and Tversky, A. (1992), 'Advances in Prospect Theory: Cumulative Representation of Uncertainty', *Journal of Risk and Uncertainty*, Vol. 5, No. 4, 297-324.

Kahneman, D. and Tversky, A. (1996), 'On the Reality of Cognitive Illusions', *Psychological Review*, Vol. 103, No. 3, 582-591.

Kaufman, B. (1999), 'Emotional Arousal as a Source of Bounded Rationality', *Journal of Economic Behavior and Organization*, Vol. 38, 135-144.

Kelsey, D. (1993), 'Choice Under Partial Uncertainty', *International Economic Review*, Vol. 34, 297-308.

Kelsey, D. (1994), 'Maxmin Expected Utility and Weight of Evidence', *Oxford Economic Papers*, Vol. 46, No. 3, 425-444.

Kelsey, D. and Milne, F. (1999), 'Induced Preferences, Nonadditive Beliefs, and Multiple Priors', *International Economic Review*, Vol. 40, No. 2, 455-77.

Kelsey, D. and Quiggin, J. (1992), 'Theories of Choice Under Ignorance and Uncertainty', *Journal of Economic Surveys*, Vol. 6, No. 2, 133-153.

Keynes, J.M. (1973), 'A Treatise on Probability', *The Collected Writings of John Maynard Keynes*, vol. 8. London: Macmillan.

Kirzner, I. (1973), *Competition and Entrepreneurship*, Chicago: University of Chicago Press.

Kirzner, I. (1979), *Perception, Opportunity and Profit*, Chicago: University of Chicago Press.

Kirzner, I. (1985), *Discovery and the Capitalist Process*, Chicago: University of Chicago Press.

Kirzner, I. (1989), *Discovery, Capitalism and Distributive Justice*, Oxford: Basil Blackwell.

Kirzner, I. (1995), 'The Subjectivism of Austrian Economics', in G. Meijer (ed.), *New Perspectives on Austrian Economics*, 11-24, London and New York: Routledge.

Knight, F.H. (1921), *Risk, Uncertainty and Profit*. Chicago: University of Chicago Press.

Kooreman, P. (1997), 'The Labelling Effect of a Child Benefit System', Unpublished working paper, University of Groningen.

Kreps, D. (1988), *Notes on the Theory of Choice*, Boulder: Westview.

Kyburg, H.E. (1990), *Science and Reason*, New York: Oxford University Press.

Kyle, A. (1985), 'Continuous Auctions and Insider Trading', *Econometrica*, Vol. 53, No. 6, 1315-36.

Lamont, O. and Thaler, R. (2003), 'Can the Market Add and Subtract? Mispricing in Tech Stock Carve-Outs', *Journal of Political Economy*, Vol. 111, No. 2, 227-68.

Lawson, C. (1999), 'Towards a Competence Theory of the Region', *Cambridge Journal of Economics*, vol. 23, no. 2.

Lawson, T. (1997), *Economics and Reality*, London: Routledge.

Lawson, T. (2003), *Reorienting Economics*, London: Routledge.

Lepore, E and R. van Gulick (eds) (1991), *John Searle and his Critics*, Oxford and Cambridge, MA: Blackwell.

Levi, I. (1986), 'The Paradoxes of Allais and Ellsberg', *Economics and Philosophy*, Vol. 2, 23-53.

Lewin, S. (1996), 'Economics and Psychology: Lessons for Our Own Day From the Early Twentieth Century', *Journal of Economic Literature*, Vol. 34, No. 3, 1293-1323.

Lewis, D. (1966), 'An Argument for the Identity Theory', *Journal of Philosophy* 63, 1:17-25.

Lewis, P. (1996), 'Metaphor and Critical Realism', *Review of Social Economy*, vol. LIV, no. 4.

Lichtenstein, S. and Slovic, P. (1971), 'Reversals of Preference Between Bids and Choices in Gambling Situations', *Journal of Experimental Psychology*, Vol. 89, No. 1, 46-55.

Little, I. (1949), 'A Reformulation of the Theory of Consumer's Behaviour', *Oxford Economic Papers*, Vol. 1, 90-9.

Lloyd, B. (1972), *Perception and Cognition*, Harmondsworth: Penguin.

Loewenstein, G. (2001), 'The Creative Destruction of Decision Research', *Journal of Consumer Research*, Vol. 28, No. 3, 499-505.

Loewenstein, G. (2000), 'Emotions in Economic Theory and Economic Behavior', *American Economic Review*, Vol. 90, No. 2, 426-32.

Lucas, R. (1981), *Studies in Business Cycle Theory*, Cambridge, Massachusetts: MIT Press.

Lycan, W. 1987), *Consciousness*, Cambridge, MA: MIT Press.

Machlup, F. (1946), 'Marginal Analysis and Empirical Research', *American Economic Review*, Vol. 36, No. 3, 519-54.

Mäki, U. (2001), *The Economic World View: Studies in the Ontology of Economics*, Cambridge: Cambridge University Press.

Margolis, H. (1994), *Paradigms and Barriers: How Habits of Mind Govern Scientific Beliefs*, Chicago: University of Chicago Press.

Mariotti, M. (1995), 'Is Bayesian Rationality Compatible with Strategic Rationality?', *Economic Journal*, Vol. 105, No. 432, 1099-109.

Mariotti, M. (1997), 'Decisions in Games: Why There Should be a Special Exemption From Bayesian Rationality', *Journal of Economic Methodology*, Vol. 4, No. 1, 43-60.

Marr, D. (1982), *Vision*: San Francisco, CA: W. H. Freeman.

Marschak, J. (1950), 'Rational Behaviour, Uncertain Prospects, and Measurable Utility', *Econometrica*, Vol. 18, 111-41.

Mas-Colell, A., Whinston, M. and Green, J. (1995), *Microeconomic Theory*, New York and Oxford: Oxford University Press.

McGinn, C. (1991), *The Problem of Consciousness*, Oxford: Basil Blackwell.

McGinn, C. (1999), 'Can We Ever Understand Consciousness?', *New York Review of Books*, Vol. XLVI, No. 10, 44-8.

Mehra, R. and Prescott, E. (1985), 'The Equity Premium Puzzle', *Journal of Monetary Economics*, Vol. 40, No. 2, 145-161.

Mitchell, W. (1918), 'Bentham's Felicific Calculus', *Political Science Quarterly*, Vol. 33: 161-83.

Nagel, T. (1974), 'What is it Like to be a Bat?', *Philosophical Review*, Vol. LXXXIII, No. 4, 435-50.

Nagel, T. (1986), *The View from Nowhere*, New York: Oxford University Press.

Nelson, R. and Winter, S. (1982), *An Evolutionary Theory of Economic Change*, Cambridge, Massachusetts: Harvard University Press.

Newell, A. (1982), 'The Knowledge Level', *Artificial Intelligence*, Vol. 18, 87-127.

Newell, A. and Simon, H. (1972), *Human Problem Solving*, Englewood Cliffs, NJ: Prentice Hall.

Nonaka, I. and Takeuchi, H. (1995), *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation*, Oxford and New York: Oxford University Press.

Oakley, A. (1997), *The Foundations of Austrian Economics from Menger to Mises*, Cheltenham: Edward Elgar.

Oakley, A. (1999), *The Revival of Modern Austrian Economics*, Cheltenham: Edward Elgar.

Oakley, A. (2002), *Reconstructing Economic Theory: The Problem of Human Agency*, Cheltenham: Edward Elgar.

O'Curry, S. (1997), 'Income source effects', Unpublished working paper, DePaul University.

Odean, T. (1998), 'Are Investors Reluctant to Realise Their Losses?', *Journal of Finance*, Vol. 53, No. 5, 1775-1798.

O'Donoghue, T. and Rabin, M. (2000), 'Risky Behavior Among Youths: Some Issues from Behavioral Economics', University of California, Berkeley, Department of Economics Working Paper: E00/285.

O'Neill, J. (2001), 'Essences and markets', in Mäki, U. (ed.) *The Economic World View: Studies in the Ontology of Economics*, Cambridge: Cambridge University Press.

Parducci, A. (1965), 'Category Judgement: A Range-Frequency Theory', *Psychological Review*, Vol. 72, 407-418.

Parducci, A. (1974), 'Contextual Effects: A Range-Frequency Analysis', in L. Carterette and M. Friedman (eds.), *Handbook of Perception (Vol. II)*, 127-141, New York: Academic Press.

Pfouts, R. (2002), 'On the Need for a More Complete Ontology of the Consumer', in Fullbrook, E. (ed.) *Intersubjectivity in Economics: Agents and Structures*, London: Routledge.

Place, U (1956), 'Is Consciousness a Brain Process?', *British Journal of Psychology*, Vol. 47, 44-50.

Place, U. (1988), 'Thirty Years On – Is Consciousness Still a Brain Process?', *Australasian Journal of Philosophy*, Vol. 66, No .2, 208-219.

Plotkin, H. (1994), *Darwin Machines and the Nature of Knowledge: Concerning Adaptations, Instinct and the Evolution of Intelligence*, Harmondsworth: Penguin.

Polanyi, M. (1958), *Personal Knowledge: Towards a Post-Critical Philosophy*, Chicago, University of Chicago Press.

Polanyi, M. (1967), *The Tacit Dimension*, London: Routledge and Kegan Paul.

Pratten, S. (1997), 'The Nature of Transaction Cost Economics', *Journal of Economic Issues*, Vol. 31, No. 3, 781-803.

Preston, J. and M. Bishop (eds.) (2002), *Views into the Chinese Room, New Essays on Searle and Artificial Intelligence*, Oxford and New York: Oxford University Press.

Putnam, H. (1995), *Pragmatism*, Oxford: Blackwell.

Quiggin, J. (1982), 'A Theory of Anticipated Utility', *Journal of Economic Behavior and Organisation*, Vol. 3, 323-43.

Quiggin, J. (1993), *Generalised Expected Utility Theory: The Rank-Dependent Model*, Boston: Kluwer Academic.

Rabin, M. (1993), 'Incorporating Fairness into Game Theory and Economics', *American Economic Review*, Vol. 83, 1281-1302.

Rabin, M. (1998), 'Psychology and Economics', *Journal of Economic Literature*, Vol. 36, No. 1, 11-46.

Rabin, M. (2002a), 'A Perspective on Psychology and Economics', *European Economic Review*, Vol. 46, No. 4-5, 657-85.

Rabin, M. (2002b), 'Inference by Believers in the Law of Small Numbers', *Quarterly Journal of Economics*, Vol. 117, No. 3, 775-816.

Read, D., Loewenstein, G., and Rabin, M. (1999), 'Choice Bracketing', *Journal of Risk and Uncertainty*, Vol. 19, No. 1-3, 171-97.

Reber, A. (1993), *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*, Oxford and New York: Oxford University Press.

Rizzello, S. (1997), *The Economics of the Mind*, Cheltenham: Edward Elgar.

Robbins, L. (1937), *An Essay on the Nature and Significance of Economic Science*, London: Macmillan.

Roe, R., Busemeyer, J. and Townsend, J., (2001), 'Multialternative Decision Field Theory: A dynamic Connectionist Model of Decision-Making', *Psychological Review*, Vol. 108, 370-392.

Rosenberg, A. (1992), *Economics – Mathematical Politics or Science of Diminishing Returns?* Chicago: University of Chicago Press.

Rosenberg, A. (1995), *Philosophy of Social Science*, 2nd ed., Boulder, CO: Westview Press.

Rosenberg, A. (1998), 'Folk Psychology', in J. Davis, W. Hands and Mirowski. P (eds.), *The Handbook of Economic Methodology*, 195-7, Cheltenham: Edward Elgar.

Rosenberg, A. (2001), 'The Metaphysics of Microeconomics' in Mäki, U. (ed.) *The Economic World View: Studies in the Ontology of Economics*, Cambridge: Cambridge University Press.

Ross, L. and Nisbett, R. (1991), *The Person and the Situation*, Philadelphia, PA: Temple University Press.

Runde, J. (1988), 'Subjectivism, Psychology and the Austrians', in Earl, P. (ed.) *Psychological Economics: Developments, Tensions and Prospects*, 101-120, Boston: Kluwer Academic Press.

Runde, J. (2001), 'Chances and Choices: Notes on Probability and Belief in Economic Theory', in Mäki, U. (ed.) *The Economic World View: Studies in the Ontology of Economics*, Cambridge: Cambridge University Press.

Runde, J. (2002), 'Filling in the Background', *Journal of Economic Methodology*, Vol. 9, No. 1, 11-30.

Russell, T. (1997), 'The Rationality Hypothesis in Economics: From Wall Street to Main Street', *Journal of Economic Methodology*, Vol. 4, 83-100.

Russell, T. and Thaler, R. (1985), 'The Relevance of Quasi Rationality in Competitive Markets', *American Economic Review*, Vol. 75, No. 5, 1071-82.

Ryle, G. (1949), *The Concept of Mind*, London: Hutchinson.

Samuelson, P. (1937), 'A Note on the Measurement of Utility', *Review of Economic Studies*, Vol. 4, No .2, 155-61.

Samuelson, P. (1938), 'The Empirical Implications of Utility Analysis', *Econometrica*, Vol. 6, No. 4, 344-356.

Samuelson, P. (1947), *Foundations of Economic Analysis*, Cambridge, MA: Harvard University Press.

Samuelson, P. (1948), 'Consumption Theory in Terms of Revealed Preference', *Economica*, Vol. 15, 243-53.

Savage, L. (1954) *The Foundations of Statistics*, New York: John Wiley and Sons.

Schmeidler, D. (1989), 'Subjective Probability and Expected Utility Without Additivity', *Econometrica*, Vol. 57, No. 3, 571-587.

Searle, J. (1979), *Expression and Meaning: Studies in the Theory of Speech Acts*, Cambridge: Cambridge University Press.

Searle, J. (1983), *Intentionality: An Essay in the Philosophy of the Mind*, Cambridge: Cambridge University Press.

Searle, J. (1990), 'Collective Intentions and Actions', P. Cohen, J. Morgan and M. Pollack (eds.) in *Intentions in Communication*, Cambridge, Mass.: MIT Press.

Searle, J. (1991), 'Intentionalistic Explanations In The Social Sciences', *Philosophy of the Social Sciences*, Vol. 23, No. 3, 332-344.

Searle, J. (1994a), *The Rediscovery of the Mind*, Cambridge, Mass.: MIT Press.

Searle, J. (1994b), 'The Connection Principle and the Ontology of the Unconscious: A Reply to Fodor and Lepore', *Philosophy and Phenomenological Research*, Vol. 54, No. 4, 847-855.

Searle, J. (1995), *The Construction of Social Reality*, Middlesex: Allen Lane, The Penguin Press.

Searle, J. (1997), *The Mystery of Consciousness*, London: Granta Books.

Searle, J. (1999), *Mind, Language and Society*, London: Weidenfeld & Nicolson.

Searle, J. (2001), *Rationality in Action*, Cambridge, Mass.: MIT Press.

Searle, J. (2002), *Consciousness and Language*, Cambridge: Cambridge University Press.

Sen, A. (1993), 'Internal Consistency of Choice', *Econometrica*, Vol. 63, No. 3, 495-521.

Sent, E-M. (1997), 'Sargent versus Simon: Bounded Rationality Unbound', *Cambridge Journal of Economics*, Vol. 21, 323-38.

Shackle, G. (1949), *Expectations in Economics*, Cambridge: Cambridge University Press.

Shackle, G. (1955), 'Businessmen on business decisions', *Scottish Journal of Political Economy*, Vol. 2.

Shackle, G. (1969), *Decision, Order and Time in Human Affairs*, Cambridge: Cambridge University Press.

Shackle, G. (1970), *Expectation, Enterprise and Profit*, London: George Allen Unwin.

Shackle, G. (1972), *Epistemics and Economics: A Critique of Economic Doctrines*, Cambridge: Cambridge University Press.

Shackle, G. (1983), 'A Student's Pilgrimage', in S. Frowen (ed.) *Business, Time and Thought. Selected Papers of G.L.S. Shackle*, 230-9, London: Macmillan.

Shafir, E., Diamond, P. and Tversky, A., (1997), 'Money Illusion', *The Quarterly Journal of Economics*, Vol. 112, No. 2, 341-74.

Shefrin, H. (2000), *Beyond Greed and Fear*, Boston: Harvard Business School Press.

Shefrin, H. and Statman, M. (1985), 'The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence', *Journal of Finance*, Vol. 40, No. 3, 777-790.

Shefrin, H and Thaler, R. (1988), 'The Behavioral Life Cycle Hypothesis', *Economic Inquiry*, Vol. 26, No. 4, 609-643.

Shiller, R. (2000), *Irrational Exuberance*, Princeton: Princeton University Press.

Shiller, R. (2003), *The New Financial Order: Risk in the 21st Century*, Princeton: Princeton University Press.

Shleifer, A. (2000), *Inefficient Markets*, Oxford: Oxford University Press.

Simon, H. (1947), *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*, New York: Macmillan.

Simon, H. (1955), 'A behavioral model of rational choice', *Quarterly Journal of Economics*, 69: 99-118.

Simon, H. (1956), 'Rational Choice and the Structure of the Environment,' *Psychological Review*, Vol. 63, 129-38.

Simon, H. (1959), 'Theories of Decision-Making in Economic and Behavioral Sciences', *American Economic Review*, 49: 253-83.

Simon, H. (1964), 'Information Processing in Computer and Man', reprinted in Simon (1992), 76-101.

Simon, H. (1966), 'Thinking by Computers', reprinted in Simon (1992), 55-75.

Simon, H. (1977), 'Rational Decision-making in Business Organizations', *Nobel Memorial Lecture*, 8 December 1977.

Simon, H. (1992), *Economics, Bounded Rationality and the Cognitive Revolution*, Aldershot: Edward Elgar.

Slovic, P., Fischhoff, B., and Lichtenstein, S. (1982), 'Facts Versus Fears: Understanding Perceived Risk,' in Kahneman, D., Slovic, P., and Tversky, A. (ed.) *Judgment under uncertainty: Heuristics and biases*, Cambridge, England: Cambridge University Press.

Smart, J. (1959), 'Sensations and Brain Processes', *Philosophical Review*, Vol. 68, 141-156.

Smith, B. (1997), 'The Connectionist Mind: A Study of Hayekian Psychology', in Frowen, S. (ed.), *Hayek: Economist and Social Philosopher*, 9-29, Basingstoke: Macmillan Press.

Steedman, I. (1989), *From Exploitation to Altruism*, Cambridge: Polity Press.

Steedman, I. (2000), 'On Some Concepts of Rationality in Economics', in P. Earl and S. Frowen (eds), E*conomics as an Art of Thought: Essays in memory of G.L.S Shackle*, London and New York: Routledge.

Stich, S. (1983), *From Folk Psychology to Cognitive Science: The Case Against Belief*, Cambridge, MA: MIT Press.

Stigler, G. (1961), 'The Economics of Information', *Journal of Political Economy*, Vol. 69, 213-25.

Stiglitz, J. (1987), 'The Causes and Consequences of the Dependence of Quality on Price', *Journal of Economic Literature*, Vol. 25, 1-48.

Stiglitz, J. (1992), 'Methodological Issues and the New Keynesian Economics, in Vercelli, A. and Dimitri, N. (eds.) *Macroeconomics: A Survey of Research Strategies*, 38-86, Oxford: Oxford University Press.

Stiglitz, J. (1994), *Whither Socialism?,* Massachusetts: MIT Press.

Stiglitz, J. (2000), 'The Contributions of the Economics of Information to Twentieth Century Economics,' *Quarterly Journal of Economics*, Vol. 115, No. 4, 1441-1478.

Steedman, I. (1989), *From Exploitation to Altruism*, Cambridge: Polity Press.

Thaler, R. (1980), 'Toward a Positive Theory of Consumer Choice', *Journal of Economic Behaviour and Organisation*, Vol. 1, No. 1, 39-60.

Thaler, R. (1999), 'Mental Accounting Matters', *Journal of Behavioural Decision Making*, Vol. 12, No. 3,183-206.

Thaler, R. (2000), 'From Homo Economicus to Homo Sapiens', *Journal of Economic Perspectives*, Vol. 14, No. 1, 133-41.

Thaler, R. and Johnson, E. (1991), 'Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice', in Thaler, R. (ed.) *Quasi-Rational Economics*, 48-73, New York: Russell Sage Foundation.

Thaler, R., Tversky, A., Kahneman, D., and Schwartz, A. (1997), 'The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test', *Quarterly Journal of Economics*, Vol. 112, No. 2, 647-661.

Tuomela, R. (1991), 'We Will Do It: An Analysis of Group Intentions', *Philosophy and Phenomenological Research*, Vol. 51, 249-77.

Tuomela, R. (1995), *The Importance of Us: A Philosophical Study of Basic Social Notions*, Stanford: Stanford University Press.

Tversky, A. and Kahneman, D. (1981), 'The framing of decisions and the psychology of choice', *Science*, Vol. 211, No. 4481, 453-458.

Tversky, A. and Kahneman, D. (1982), 'Availability: A heuristic for judging frequency and probability,' in Kahneman, D., Slovic, P., and Tversky, A. (ed.) *Judgment under uncertainty: Heuristics and biases*, Cambridge, England: Cambridge University Press.

Tversky, A. and Kahneman, D. (1986), 'Rational Choice and the Framing of Decisions', *Journal of Business*, Vol. 59, No. 4, 5251-78.

Tversky, A. and Simonson, I. (1993), 'Context-dependent preferences', *Management Science*, Vol. 39, 1179-1189.

Vanberg, V. (1994), *Rules & Choice in Economics*, London: Routledge.

Van Gulick, R. (1995), 'Why the Connection Argument Doesn't Work', *Philosophy and Phenomenological Research* Vol. LV, No. 1, 201-7.

Varian, H. (1992), *Microeconomic Analysis*, New York and London: W. W. Norton.

Varian, H. (2002), *Intermediate Microeconomics: A Modern Approach*, New York and London: W. W. Norton.

Veblen, T. (1909), 'The Limitations of Marginal Utility', *Journal of Political Economy*, 17: 235-45.

von Neumann, J. and Morgenstern, O. (1947), *Theory of Games and Economic Behaviour*, Princeton: Princeton University Press.

Warneryd, K-E. (2001), *Stock-Market Psychology*, Cheltenham: Edward Elgar.

Weimer, W. (1979), *Notes on the Methodology of Scientific Research*, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Weimer, W. (1982), 'Hayek's Approach to the Problems of Complex Phenomena: An Introduction to the Theoretical Psychology of *The Sensory Order*', in W. Weimer and D. Palermo (eds), *Cognition and the Symbolic Processes*, 241-85, Vol. 2, Hillsdale, N.J.: Lawrence Erlbaum Associates.

Winter, S. (1964), 'Economic "Natural Selection" and the Theory of the Firm', *Yale Economic Essays*, Vol. 4, No. 1, 225-72.