PART I

Theory & Method

1 Theory and Method in Economics and Psychology

Levels and Depth of Explanation

Denis Hilton

1.1 Introduction

Entering the twenty-first century, economics increasingly looked to psychology and neuroscience in order to revise its assumptions about how people process information and make choices. New subdisciplines of economics sprang up such as behavioural economics, psychological economics, cognitive economics, and neuroeconomics which drew extensively on findings in psychology and neuroscience. In addition, the new field of experimental economics enabled economists to test theories about human choice behaviour directly in controlled conditions. Economists increasingly abandoned the assumption that people are fully rational in their choices, and became increasingly interested in constructing and testing models which incorporate realistic assumptions about human thought and behaviour.

This chapter will take a psychologist's point of view on these developments. Despite the fact that economists and psychologists are often addressing the same fundamental question – why humans make the choices they do – there is often puzzlement about the other side's theory and methodology. These kinds of misunderstanding may not only reflect lack of knowledge of the background and aims of the other's discipline, but also of the historical context of the social sciences.

I will try to explain the aims and methods of economics in a way that will make their theoretical orientations seem more understandable to psychologists. I will first try to chart the various ways that economists are seeking to use psychology in their work. I will then suggest that economics – even of the new psychological kind – is still strongly rooted in a metatheoretical perspective that many psychologists would recognize as behaviourist. I will accordingly address two major questions about theoretical frameworks and scientific explanation. The first is to do with the issue of levels of explanation, and whether and how psychological level explanations can inform economic questions, and in turn whether and how neuroscience can inform economics. The second has to with depth of explanation, which contrasts psychology's more realist approach to constructing and testing sociocognitive processes in scientific explanations of behaviour to the more instrumental perspective of economics. I argue that this realist approach not only can lead to improved explanation of the facts, but

to greater opportunities for modification of economic behaviour at both the individual and societal level.

This chapter therefore pays attention to differences in how theories are constructed and evaluated in the two disciplines. In my view, this perspective can shed light on the recent debate about methods in economics and psychology, which has focused on experimentation (e.g., Croson, 2005; Hertwig and Ortmann, 2001; 2003; Sugden, 2005b). I will suggest, for example, that much of the heat in this debate has been generated by differences in theoretical orientations (cognitivist and realist in the case of psychology, behaviourist and instrumentalist in the case of economics). While written from a psychologist's point of view, this chapter may still be useful to economists. Through throwing a different light on their assumptions and practices, it may help economists to reflect on their own professional identity and epistemological approach, in much the same way that spending a period abroad teaches us about our own national identities, through observing what raises foreigners' eyebrows.

1.2 Rational Behaviourism in Economics

At the beginning of the twentieth century, economics and psychology (like philosophy and linguistics) were confronted with the success of the natural sciences. Both attempted to adopt the methods of the natural sciences through taking behaviourist approaches (measuring observables, such as behaviours rather than thoughts and feelings in psychology, prices and market share rather than experienced value, etc., in economics). While psychology primarily adopted the experimental methods of the natural sciences, economists retained the eighteenth-century model of *homo economicus* as motivated by rational self-interest, but adopted mathematical formalisms to render their theories more precise and 'scientific'. In addition, they adopted the revealed preference axiom which assumes that preferences will be revealed in objectively measurable phenomena such as prices and market share (Lewin, 1996).

In this section, I identify some behaviourist characteristics of economics. This can be done by measuring economics up against Lyons' (1977) checklist for identifying behaviourist theories, namely, a rejection of internal states of the organism as scientific explanations of behaviour; a tendency to see no essential difference between the behaviour of humans and other animals; an emphasis on the importance of learning and reinforcement (positive and negative) in explaining behaviour; and a penchant for instrumentalist (i.e., predictive) rather than realist (i.e., explanatory) theories of behaviour.

1.2.2 Rational Behaviourism

The 'revealed preference' assumption embodies what might be called a 'rational behaviourist' approach, as it combines the rational self-interest model

of choice with a behaviourist approach to measurement. To know an agent's preferences, one does not have to ask her; one simply observes what choices she makes. In addition, there are some important symmetries between economists' models of the rational calculation of self-interest and psychologists' models of how S-R associations are formed: both assume that with experience, the organism (or agent) will learn the costs and benefits associated with actions that it (she) takes. Both approaches emphasise the importance of learning and incentives for understanding behaviour.

The rational self-interest model effectively makes the same predictions as Thorndike's (1911) Law of Effect that links reward to response – any response that leads to a reward is likely to be repeated. In the language of 'rational behaviourism', this can be expressed as the price effect in economics – an activity can be encouraged by raising the price paid for its performance. Neoclassical economics – like behaviourist psychology – assumes that human behaviour will be explained by situational costs and benefits (gain-loss matrices, reward-punishment schedules). Economics assumed that no curiosity need be expressed about the intervening cognitive processes that led from stimulus to response (it was assumed that gain-loss matrices would be calculated correctly by rational choice processes), while radical behaviourism in psychology argued that no attention should be paid to intervening cognitive processes, as they were unobservable.

Although it may seem paradoxical that economists' model of *homo economicus* drawn from the eighteenth-century Enlightenment should yield essentially the same method of analysing behaviour as early-twentieth-century psychologists' model of *Rattus Norvegicus*, that is indeed what happened. Whereas nineteenth-century economists were interested in psychology and used it to explain economic phenomena (Bruni and Sugden, 2007), from the early-twentieth century onwards the revealed preference approach held that values and preferences of decision makers are to be inferred from observing them make choices under varying conditions. Economists, like behaviourist psychologists, abhorred finding out people's values and preferences just by asking them disinterested questions. Indeed, a striking illustration of the convergence of economics and behaviourism comes from the use of revealed preference theory to infer demand curves for animal preferences (Kagel et al., 1981).

1.2.2 Experimental Procedures in Economics and Psychology: The Debate over Learning and Incentives

Tellingly, the first line of defence that economists put up against the implications of psychologists' experimental findings on bias and error (e.g., Kahneman and Tversky, 1979) was that they were based on questionnaire research with no payoffs for correct responses. Experimental economists were sceptical about whether these findings would be reproduced with experienced, financially motivated participants who fully understood their task and who had opportunities

to learn. This position is a variant of the *tabula rasa* theory that underpins behaviourism, in assuming that cognitive biases are 'malleable' rather than being inherently structured, and thus can be eliminated by appropriate market conditions. And indeed markets sometimes do eliminate biases: for example, List (2004) found that the procedure devised by Kahneman et al. (1986) to demonstrate the endowment effect (involving exchanging mugs for chocolate bars) works with normal consumers but not with expert dealers in picture cards of sports stars whose behaviour approximates to that predicted by the standard neoclassical model.

Nevertheless, the major result of economists' research on human bias and error has been to demonstrate how often biases in judgment and choice observed in experimental settings generalise to both real-world economic settings (Camerer, 2000; Hilton, 2001) and to experimental settings with financial incentives (Camerer and Hogarth, 1999; Hertwig and Ortmann, 2001). These reviews show that while incentives reduce (but do not eliminate) bias and error in approximately 55 per cent of cases, they have no effect in approximately 30 per cent of cases, and have paradoxical effects (i.e., actually *increase* bias and error) in approximately 15 per cent of cases. On the whole, then, these results vindicate the cognitive programme of research of psychology, since they show that the way cognitive processes frame how people perceive reality will indeed affect their behaviour independently of learning and incentives.

1.2.3 The Costs and Benefits of the Behaviourist Stance: Instrumentalism Versus Realism in Explanation

A common puzzle for psychologists is to understand how economists can hold on to the rational self-interest model of human choice and behaviour. After all, it has been so clearly discredited by innumerable experimental studies. The clue to this puzzle may be that psychologists' aims in developing models of choice are different from those of economists. Whereas psychologists aim to develop *realist* models of cognitive processes that are accurate and testable descriptions of how the human mind works, economists' aims in building models are often much more pragmatic. In particular, economists seek models of choice that will be useful in helping them understand and explain economic questions at the collective level, concerning market behaviour, prices, laws, institutions and so on. So if many economists hold on to the rational choice model despite its evident falsity as an exact descriptive model of individual human choice, they may adhere to it for other reasons, such as its simplicity and elegance in generating more-or-less correct implications for understanding these higher-level economic questions. This *instrumentalist* position means that these economists are likely to give up the standard expected utility (SEU) model of choice only if a more psychologically plausible model enables them to make significantly better predictions about economic questions such as market share and prices, or help them write more effective contracts and legislation (cf. Sugden, 2005b).

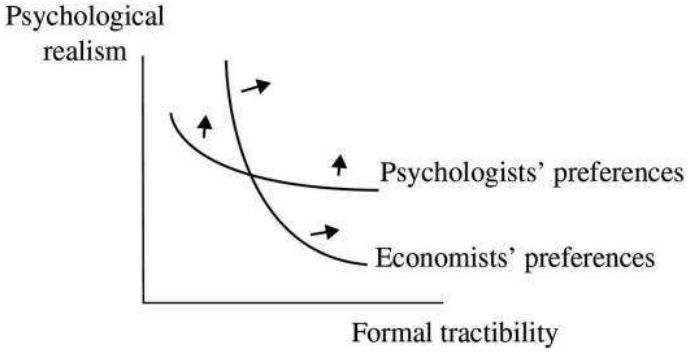

Figure 1.1. *Indifference curves for theory preferences among economists and psychologists (from Rabin, 2002).*

Rabin (2002) cleverly models these different orientations of economists and psychologists with indifference curves (see Figure 1.1). Economists' preference for models is strongly influenced by their formal properties (elegance, tractability), and less so by their realism and descriptive accuracy. For psychologists, the inverse is true.

1.3 Levels and Kinds of Explanation

We have seen that despite the diversity of ways (behavioural, psychological, cognitive, neuro- and experimental) in which economists have integrated psychological insights into economic theorising, each of these subdisciplines possesses significant assumptions that differentiate them from experimental psychology. However, I will argue that economists' use of psychology and neuroscience is likely to profit from careful attention to two questions.

The first, which I term the 'levels of explanation' question, concerns distinguishing between economic (social), psychological and neuroscientific levels of explanation, and the ways in which these levels of explanation interrelate. I will argue for a hierarchical model of scientific questions and explanation, where explanations are constrained but not determined by explanations at a lower level. I first illustrate this generative model of explanation by showing how properties of collectives (e.g., markets) may emerge from (and be explained by) lower-level properties of their components (agents). In turn, I consider how properties of agents may emerge from (and be explained by) properties of their brains. However, I conclude that economics cannot be reduced directly to neuroscience: an intermediate level of psychological explanation is necessary.

Having established the need for an autonomous psychological level of explanation, I then turn to the second 'depth of explanation' question. Here I address the need at the cognitive (psychological) level of explanation to develop theories that are explanatory in nature, and which thus offer greater

power for prediction and control than 'shallow' behaviourist style explanation, which I will show to still be often favoured in psychologically inspired economics.

1.3.1 Levels of Scientific Question: Economics, Psychology and Neuroscience

Economics, psychology and neuroscience have different objects of study and pose their scientific questions at different levels. Economics typically studies the behaviour of collective phenomena such as firms, markets and prices, and asks questions like: how can markets achieve efficient allocation of goods? What causes markets to fail? Psychology typically studies the individual; individuals' personalities; the way they perceive, remember and make judgments and choices; and asks questions like: what causes depression in an individual? Why do people make irrational choices? Neuroscience typically studies brains, mapping brain anatomy and circuits, and relating this to functional processes such as vision, hearing and decision-making. So far, neuroscience seems to focus more on mapping what parts of the brain or brain circuits are associated with different psychological functions (e.g., short- vs. long-term memory; emotional vs. 'rational' decision-making) than in answering theoretically driven questions (see Table 1.1).

Of course, the picture given in Table 1.1 is something of an oversimplification, as it is easy to see points of overlap. For example, while economists following in the tradition of Hayek see economies as 'brains' that transmit information (cf. Forsyth et al., 1992), and the social psychology of group processes examines how information flows around groups in a way that is close in spirit to this kind of economic analysis (e.g., Latané and Nowak, 1991). But the important point is to see that different levels of scientific question exist, and to understand how questions at one level can be answered by facts, assumptions or hypotheses that draw on another level.

1.3.2 Generative Explanation: How Economic Phenomena Can Emerge from Psychological Characteristics

A cornerstone of classical economics has been to explain a seemingly paradoxical result: how socially desirable equilibrium states can emerge from a society of uniquely self-interested actors (Smith, 1776; Turgot, 1761). This can be explained in the following way. If there is a lack of grain at Toulouse due to crop failure but a surplus at Limoges where the harvest was good, prices will rise in Toulouse and drop in Limoges. This will have two advantages. First, the high prices in Toulouse will discourage people there from hoarding grain because it will be too expensive, thus encouraging grain to be distributed to more people. And second, the higher prices in Toulouse will attract grain from Limoges through encouraging an entrepreneur to buy low in Limoges, and

Table 1.1 *Three levels of scientific question and associated scientific disciplines*

Domains of study (level of question)	Level of explanation	Kinds of science
Market behaviour *Small group decision-making*	Human collectives	Economics Social psychology of groups
Cognitive processes in judgment and decision-making *Individual values and preferences*	Human individuals	Cognitive psychology Social cognition Artificial intelligence
Brain and body function	Brain, body parts, systems	Neuroscience

accept the cost (and risk) of shipping grain to Toulouse in order to sell high there. These classic market mechanisms make some hypotheses about human nature: namely, that some people are smart enough to spot opportunities, and greedy enough to take the risks necessary to exploit them.

In science as in common sense, questions typically arise when we need to know how something surprising or undesirable has come about (Weiner, 1985). We can see that Smith's famous example of the "hidden hand" that redistributes goods to where they are needed is so powerful because it shows how surprising properties (e.g., a socially desirable distribution mechanism) can emerge from a self-organizing system based on actors whose sole motive is rational self-interest. The point of the example is that market success occurs *because* (not *in spite*) of a psychological property of its agents – their self-interest. Smith's equilibrium model is a paradigm case of generative explanation (Harré, 1988) which shows how an effect (social welfare in the distribution of goods) can be produced by a mechanism (prices) built of components that possess specified (but unexplained) causal powers (selfish agents). The kind of explanation involved here thus seems to consist in building a model which *can* generate the effects of interest (cf. Sugden, 2000).

1.3.3 Questions about Deviations: The Logic of Contrastive Explanation

Following Smith, neoclassical economics established efficient market theory as an ideal mechanism for the distribution of goods. Subsequent theories in economics have taken *homo economicus* and efficient market theory as a foil, and have sought to explain deviations from this 'norm' (e.g., Sugden, 2005b). Such 'contrastive explanation' (Hesslow, 1983; 1988; Lipton, 1991) may seek to identify as a cause the property in the deviation model is different to standard SEU theory that 'makes the difference' between the effect observed (a market

anomaly) and what should be expected according to efficient market theory. The kind of explanation involved here seems to involve counterfactual reasoning about what will happen if we *modify* a certain characteristic of the SEU/efficient market model. Deviation models can focus either on market failures or on psychological 'irrationality' for explanations of market anomalies.

1.3.3.1 *Within-Level Explanation: Explaining Market Anomalies through Market Failure*

In a classical form of explanation in economics, questions will be resolved by explanations which stay at the same level of analysis. For example, a market anomaly will be explained by a market failure (e.g., failure of buyers and sellers to have access to the same information about the quality of the product). For example, Akerlof's (1970) analysis of 'lemons' shows that in markets of rational actors where there is asymmetric information (e.g., sellers of used cars know more about the quality of the car than potential buyers), the price of used cars will drop dramatically as buyers assume that sellers will only offer low-quality cars (lemons) leading rational buyers to refuse to pay high prices. So sellers will in turn not be motivated to sell high-quality cars, and only low-quality cars will be offered for sale. This leads to 'adverse selection' in the market, as bad products drive out good products, leading the example to be presented as showing how 'asymmetric information can result in market failure' (Pindyck and Rubinfield, 1998, p 620; see also Sugden, 2000, for a more detailed exposition of Akerlof's argument). Here the explanation for market failure is found in terms of market characteristics (distribution of information in the market), rather than in terms of the irrationalities of individuals.

Similarly, typical explanations in psychology stay 'in-house' at the same level. For example, decisions to choose a certain diet can be predicted by a weighted combination of our beliefs that the diet will work, the desirability of being thinner for us, the desirability of being thinner for important others, and our desire to please those others (Ajzen and Fishbein, 1980). However, with the advent of neurosciences, much thought has gone into how psychological-level explanations (in terms of beliefs, desires, rules of inference, etc.) can be constrained by what is known about the components of the system (serial or parallel computer, brain) on which these algorithms run (Marr, 1982). Consequently, interesting properties at one level of explanation can emerge from the properties of the components at a lower level.

1.3.3.2 *Cross-Level Explanation: Explaining Market Failure through Psychological Characteristics*

In cross-level explanation, the underlying psychological properties of market agents can also be used to explain market failure. Efficient market theory assumes that under the right market conditions (e.g., informational symmetry between buyers and sellers), markets composed of fully rational agents will allow goods to be exchanged to the benefit of both parties. However, change

the nature of the component parts of the system and new properties will emerge in the system itself. For example, if we assume – following Kahneman and Tversky's (1979) prospect theory – that losses loom larger than gains in the minds of market agents – then markets can lose liquidity because of the 'disposition effect' (Shefrin and Statman, 1985), as agents refuse to trade goods of equivalent value because the experienced loss has higher disutility than the experienced gain has utility for them.

This 'disposition effect' could explain anomalies such as that observed in the British housing market in the early 1990s when a dramatic fall in house prices meant that many owners would have to sell houses at a loss compared to what they had paid before the market collapsed. Of course, because of the market collapse, they could also buy other houses cheaper. But loss aversion would explain why many would continue to live for months and even years in a house in one city (say Manchester) while commuting four hours to another (such as Edinburgh) where they had taken a job, with all the attendant commuting costs and dislocation of personal and family life. The psychological hurt that would be caused by selling a house at a loss would outweigh the gain incurred by buying one in their new place of work.

1.4 Explanation across Two Levels: Is Neuroeconomics Possible?

A question that merits consideration is whether intelligible explanations can be achieved by going down *two* levels (e.g., from economics through psychology to neuroscience). Scientific explanation often proceeds by attributing a phenomenon to some disposition of another entity, which serves as the end-point in an explanatory chain. In turn, the explanation can be expanded by attributing this disposition to the disposition of another entity at a lower level, which is attributed with a unexplained 'causal power' to produce the effects it does. As Harré (1988 p 142 writes):

> the chemical behaviour of liquids, solids and gases is explained by the behaviour of unobservables, molecules and chemical atoms … But one might well ask for an explanation of the behaviour of chemical atoms, for example why do they chum up in the proportions they do? The next level of explanation simply repeats the pattern of the level above. Drawing on the behaviour of positively and negatively electrically charged bodies as a source-model, a further step is taken, in which electrically charged electrons and protons are invoked, the story being filled out with neutral neutrons. The electrical properties of these structures explain the differences in behaviour of chemical atoms.

As we shall see in this section, this kind of reductionism across levels of explanations seems to be a pattern of explanation envisaged in what has come to be called 'neuroeconomics'.