Popper, Rationality and The Possibility of Social Science

Danny FREDERICK

Received: 31.5.2011
Final version: 6.1.2012

BIBLID [0495-4548 (2013) 28: 76; pp. 61-75]

ABSTRACT: Social science employs teleological explanations which depend upon the rationality principle, according to which people exhibit instrumental rationality. Popper points out that people also exhibit critical rationality, the tendency to stand back from, and to question or criticise, their views. I explain how our critical rationality impugns the explanatory value of the rationality principle and thereby threatens the very possibility of social science. I discuss the relationship between instrumental and critical rationality and show how we can reconcile our critical rationality with the possibility of social science if we invoke Popper's conception of limited rationality and his indeterminism.

Keywords: human action; explanation; indeterminism; Popper; rationality; rationality principle; situational logic; social science.

RESUMEN: Las ciencias sociales emplean explicaciones teleológicas que dependen del principio de racionalidad, de acuerdo con el cual las personas exhiben un tipo de racionalidad instrumental. Popper señala que, además, las personas exhiben una racionalidad crítica, que es la tendencia a distanciarse, y a cuestionar o criticar, sus distintos puntos de vista. Aquí explico cómo nuestra racionalidad crítica impugna el valor explicativo del principio de racionalidad poniendo en peligro la propia posibilidad de las ciencias sociales. Asimismo, discuto la relación entre la racionalidad instrumental y crítica, y muestro cómo podemos reconciliar nuestra racionalidad crítica con la posibilidad de las ciencias sociales si invocamos la concepción de racionalidad limitada de Popper y su indeterminismo.

Palabras clave: acción humana; explicación; indeterminismo; Popper; racionalidad; principio de racionalidad; lógica situacional; ciencia social.

1. *Introduction*

The social sciences differ from the natural sciences in that they explain events in terms of purposive human action. The possibility of such explanation seems to depend upon the rationality principle, according to which people act in ways they see as suitable to their purposes. This, in turn, suggests something akin to a law of behaviour, according to which each person does what *anyone* one would do in the same situation, given the same purposes and the same information. This, further, suggests that people's idiosyncratic psychologies are not needed and, indeed, are unwelcome, in social-scientific explanations, which should instead proceed in terms of the logic of the situation.

An illuminating and cogent defence of this view is offered in Popper 1994a, which I summarise in section 2. In section 3, I show how Popper's account of critical rationality, according to which we are rationally entitled to stand back from, and to question or criticise, our views, threatens to rob the rationality principle of explanatory value and thus to render doubtful the very possibility of social science. In section 4, I explain how we can resolve this tension if we draw on other aspects of Popper's philosophy, in particular, his conception of limited rationality and his indeterministic metaphysics. This enables us to see how the actions of rational agents can be explained and

predicted and thus how social science is possible. I conclude the discussion in section 5.

2. Thesis: Popper on Social Science

In this section I expound briefly Popper's account of social science in terms of the rationality principle and situational logic. Popper's exposition and discussion is sometimes vague, ambiguous and misleadingly expressed; but I think that, taken in the context of his broader philosophy, his words can be given a clear and consistent interpretation. I do not discuss alternative interpretations or defend my own, since my aim here is not primarily exegesis.

The fundamental problem of the social sciences is to explain events in terms of human actions and social situations. A social situation involves not only physical things but also institutions and traditions, such as rules, organisations, signs and meanings, of which the agent takes account in deciding how to act. The situation also involves human agents who have aims and information relevant to achieving those aims (1994a, 166-67, 169). The rationality principle states that each agent acts adequately or appropriately to her situation as she sees it, given her aims (1994a, 178-79, 181, 183n19). In what follows I will use 'suitably' as shorthand for 'adequately or appropriately.' The rationality principle can be analysed as a conjunction of two propositions. The first is the instrumental principle, which says that a *rational* agent acts suitably to her situation as she sees it, given her aims. The second is the proposition that all humans are rational agents in the sense specified by the instrumental principle. Thus, social-scientific explanation is teleological: it explains social events as the outcome of human intentional actions, that is, actions directed toward an end (1994a, 181, note 1). Yet the outcomes produced, including institutions and traditions, are often unintended, being the by-products of actions that had other aims (1994a, 170).

The concern of the *historical* social sciences is to explain one or a smallish number of *singular* events. Such explanation is a deduction of a description of a singular event from a description of a concrete social situation in conjunction with the rationality principle. The concern of the *theoretical* social sciences is almost always to explain *types* of events. Such explanation is a deduction of a description of a type of event from a description of a typical social situation (a model) in conjunction with the rationality principle. In either historical or theoretical social science, explanation is achieved by treating the agent as a *typical* human agent with specified aims and information: her action is explained by being shown to be what anyone would have done in the same situation, given her aims and information. Thus, actions are explained in terms of the logic of the situation: what occurs is a consequence of each agent working out what is implicit in her situation (1994a, 162-70).

Popper rejects 'psychologism.' *First*, he uses 'information' and 'knowledge' interchangeably, but neither should be equated with beliefs: the agent's knowledge may be conjectural, accepted rather than believed, and the agent might even believe it to be false, though good enough for practical purposes. For example, in engineering and navigation, formulae known to be false are used in deciding how to act because they are good approximations that are easy to handle (Popper 1972b, 56; see also Bratman

1999 for other kinds of example). *Second*, the aims and information invoked to explain the action need not be at the forefront of the agent's mind when she is acting: they may instead be in the background or they may be implied by aims and information that are in the forefront of her mind (1994a, 167-68, 182 note 6). *Third*, the agent's aims and information are not discovered by psychological methods but are rather imputed to her as aims and information that *anyone* would have in her situation (1994a, 167-68). *Fourth*, when we are constructing a model, that is, a simplification of reality, the aims and information that we attribute to the agent might not have psychological reality for any agent: our explanatory aim may be served by an approximation to the truth, in a way similar to that in which idealised models of the solar system are used in conjunction with Newton's theory to explain the approximate motions of the planets (1994a, 162-66, 170, 182 note 6).

Popper gives an example of a man who is in a hurry to cross a road. We may explain his somewhat erratic movements across the road by referring to the moving and parked cars and other traffic, the rule of the road, police regulations, traffic signals, zebra crossings and other social institutions as well as the man's knowledge of these things, and his aim of crossing the road quickly and safely (1994a, 166-68).

The rationality principle, Popper says, is almost empty (1994a, 169), because people usually do act suitably to their situation; but it is clearly false because people often do not so act (1994a, 172, 177-78). For example, flustered drivers trying to get out of a traffic jam or desperately trying to park when there is no parking space to be found are not acting suitably to their situation (1994a, 172), as they would acknowledge if they reflected about their aims and circumstances. But social scientists should adopt the rationality principle as a methodological rule to search for teleological explanations; and this amounts to treating the principle as if it were true when testing a social explanation (1994a, 171). For, to the extent that we reject the rationality principle, human action ceases to be explainable in teleological terms (1994a, 178). If we denied that the rationality principle ever had application, social science would become impossible: any explanation of social events would then have to be in terms of physics, biology, psychology or some other non-social science, as we might explain the behaviour of the flustered drivers in terms of an evolved psychological mechanism (1994a, 172). And if we reject the rationality principle only whenever our favoured explanatory model fails a test, thus dismissing the unexplained actions as irrational, this would make our model unfalsifiable: such rejection of the rationality principle would be an ad hoc, or unscientific, way of saving the model from refutation. Our model would be rendered empirically vacuous, for it would be held to explain the social events for which it can provide an explanation, while all other social events would be declared incapable of teleological explanation. However, we would learn more, and explain more, by retaining the rationality principle and trying to develop an alternative model (perhaps an independently testable revision of the old model) to explain the social events that the old model cannot (1994a, 177-78).

We can note parenthetically that Popper's methodological treatment of the rationality principle is analogous to his recommendation (1972a, 61-62; 1982a, 149) that in the natural sciences the principle of universal causation should be regarded as a meth-

odological rule to seek causal explanations, even though (so Popper thought) indeterminism is true (1982a, 1982b). In science our aim is to explain as much, and as fully, as we can.

Popper suggests two ways in which the rationality principle can be false. People may fail to act suitably to their situation either because they are behaving irrationally (like the flustered drivers), or because they misunderstand the situation they are in (1994a, 172). That Popper should consider this second type of case is curious, since the agents will satisfy the rationality principle if they act suitably to the situation *as they see it*. Their actions are mistaken, rather than irrational, because their view of the situation is mistaken. (Popper offers a brief discussion relevant to this matter, at 1994a, 183-84, note 19, but the distinctions he makes there seem to confuse things further. I neglect that passage here.)

It has been objected that a model employing the rationality principle can be used to explain a specific action only in 'single exit' situations (Latsis 1972, 211), that is, situations in which, as the agent sees it, there is just one action that is suitable to her aims. But social situations often present agents with a number of options none of which seems more suitable than any of the others. Thus, there are cases of indifference in which the agent is faced with a choice between options that seem equally good, or equally bad, such as a choice between rival investments which pay the same rate of return (and are indistinguishable in all other relevant respects), or between equally competent plumbers who quote the same price for a job. There may also be cases of indeterminacy. For example, I may have a choice between three investments that offer the same rate of return, but one will help protect the environment, the second will help some very poor people out of poverty, while the third will fund research into medicines that will relieve suffering and save lives. How do I decide between these three diverse values? It does not help much to suppose that my priority is the environment, for it may be that the impact of the first investment on the environment is likely to be relatively small while the impact of the other two on poverty and health, respectively, is likely to be larger. How much alleviation of poverty or of ill-health should I give up for how much amelioration of the environment? It seems unlikely there will be a determinate answer (see the essays in Chang 1998). There are often cases of ignorance in which there is no way of knowing some of the vital consequences of choosing between the options being considered, leaving us unable to say which is suitable. For example, in a choice between investments we often cannot know which, if any, of the companies will go bust. Something similar applies to a choice between marriage partners. These cases contrast with indifference and indeterminacy because the agent may suspect that some of the options are unsuitable, though she has no way of finding out which options they are. So, unlike cases of indifference and indeterminacy, we cannot say that the options are equally suitable from the agent's point of view (the agent may protest that she does not know whether any of them is suitable); though it may simplify theory to treat these cases *as if* they involved equally suitable options from the agent's point of view (such simplification will be assumed in what follows). Further, in some situations the agent may have no idea what the relevant options are, in which case she might just try something at random.

However, although in multiple-exit situations there can be no teleological explanation for which specific action was taken, it is at least a *partial* explanation of why someone acted as she did that her action was one of a limited set of alternative actions each of which appeared suitable, or not unsuitable, to the situation as she saw it. In such circumstances we may even be able to offer a *disjunctive* or *vague* prediction that the agent will perform one out of a restricted set of actions. Although Popper neglects such partial explanations in his discussion of theoretical social science, he explicitly allows partial explanations in his discussion of historical events, such as creative acts in the history of science (1973b, 170-80). But, as partial explanations, and their corresponding disjunctive or vague predictions, cannot explain or predict *which* specific action the agent will take, they are much weaker than their single-exit counterparts.

Another objection is that some parts of social science offer non-teleological explanations, either of action or of characteristics that influence action. Thus, individual, social or evolutionary psychology may be able to explain types of irrational action in social situations or the ways in which people are moulded by the institutions and traditions in which they grow up and live. Such explanations will be non-teleological, in terms of 'mechanisms;' yet they are part of the explanation of social events (Gorton 2006, 99-110). However, Popper concedes all this, as we saw, partly, five paragraphs back (the explanation of irrational action in terms of an evolved mechanism) and as we will note below (the discussion of limited rationality in section 4 and the references given there). He could also point out that to affirm that such explanations are part of the explanation of social events is not to affirm that they are part of social science. For, while the rational agent is brought into being by a natural process, is constantly impinged upon by the operation of natural laws, and has many effects upon others and on the world by the mediation of natural laws, it is only to the extent that she acts suitably to an aim in the light of information that she becomes an object for social science. There are two points here about the relationship between social and natural science. First, they are complementary: social-scientific explanations take the operation of natural laws for granted; for example, bidding at an auction only works because light and sound are transmitted in the usual ways. Second, they are in competition: for a given range social phenomena, mechanistic explanations are an *alternative* to social-scientific ones.

Finally, we should briefly note two points. *First*, our bodies do not always do what we intend them to do, due to temporary paralysis or some other physical malfunction. Such failure to act as we deem suitable is not a failure of rationality. *Second*, where an agent has multiple aims, acting rationally may mean acting unsuitably to some of these aims for the sake of acting suitably to more important ones: what is irrational is acting unsuitably to our overall, or overriding, aims. In the light of these two qualifications, the rationality principle can be stated more strictly thus: an agent will *try* (or *exercise her volition*) to act in a way that appears to her suitable to her *overall* aims in the situation as she sees it. For simplicity, I will usually leave these qualifications implicit.

3. Antithesis: Popper on Rationality

The rationality principle affirms that each person is rational in the sense that he always acts suitably to his situation as he sees it, given his aims. Popper says that such *instrumental rationality* is a minimum principle of rationality. He distinguishes it from *critical rationality*, which is the attitude of readiness to correct one's beliefs, which in its intellectually most highly developed form is the readiness to correct one's theories through critical discussion. Thus, even a madman can be understood, so long as he complies with the rationality principle and thus acts suitably to his aims and his interpretation of his situation. What makes him irrational or mad is that he lacks a readiness to correct his aims or views in the light of discussion or of events (1994a, 179-81).

However, that human agents have the capacity for critical rationality seems to render teleological explanation via the rationality principle, and thus social science, impossible. This is not because of any conflict between critical rationality and instrumental rationality, or between critical rationality and the rationality principle. It is rather because, for a critically rational agent, the assumption that he is instrumentally rational no longer permits us to explain or predict how he will act. Put simply, the problem is that even if, as the agent sees it, there is a single action that is suitable to his situation, his refusal to take that action need not be instrumentally irrational. For, if he recognises his fallibility, then he recognises that things might not be as they seem to him, and thus that his current representation of the situation may be false. So, he may question whether the single action that seems suitable to his situation really is so. He might try to criticise and improve his present theory about what he should do, or he might try to create a new and better theory; or, if he lacks the time for such creative endeavours, he might try to refute his present theory by acting in a way which conflicts with it. The agent's capacity for critical rationality seems to render teleological explanation of his actions impossible. Popper did not seem to be cognisant of this tension between critical rationality and *the explanatory value* of the rationality principle. I will explain the tension in detail with respect to single-exit situations: the obvious extension to multiple-exit situations will be made briefly.

We know from the problem of induction that, no matter how well-corroborated a theory may be, it might still be refuted in the very next moment. Indeed, the history of science exhibits one highly corroborated theory after another being rejected for something better, as Galileo's theory of terrestrial motion and Kepler's theory of celestial motion were replaced by Newton's mechanics, and as Newton's mechanics was replaced by relativity theory (Popper 1983, 75, 139-45). Even observation statements involve theoretical interpretation and so are similarly liable to be refuted. For example, 'Here is a glass of water' implies that the receptacle will exhibit the law-like behaviour of glass and that its contents will behave in the law-like way that water does; and these implications may be inconsistent with other observation statements (Popper 1972a, 94-95). Indeed, Urey's discovery of heavy hydrogen in 1931 showed that what scientists had previously taken to be water was in fact a mixture of two physically different substances (Popper 1966, 374-75). The irremovable possibility of error means that it is always open to us, rationally, to question currently accepted theories and facts and, if

we have the time and other resources, to attempt to criticise or test them and to try to replace them with better ones.

This holds, in particular, for our representations of the facts in practical decision-making. These representations of the facts include statements about *how things stand* and statements of the *options for change* available to us and their possible consequences. In many cases there might not be the time or other resources needed to try to think up and test alternative statements about how things stand or to think up alternative, perhaps entirely novel, options and criticise them in view of their possible consequences. But there is almost always time to doubt the currently accepted representation of the facts, and then it is open to us to try to refute it by attempting to achieve our aim in a way that the representation implies is impossible: we can try to refute it by acting against it. This is akin to what scientists do when they test a theory by means of an experiment: they create controlled conditions in which they try to engender an event that the theory forbids. If they succeed, the theory is refuted (though not necessarily rejected); if they fail, the theory may be corroborated.

It might be objected that it would be rational to try to refute a particular view only if one notices a specific problem with it, such as a seeming self-contradiction or a conflict with some other accepted view; and even then, it would be irrational to act contrary to the view if it is accepted, by oneself or by people generally, as the best that is currently available. But this is not so. Since it is rational to acknowledge that the view may be false, it cannot be maintained that it must be irrational to try to refute it. And even if the view is better than any available alternative, if it may be false, one cannot rationally be required to act in conformity with it. It may help to see this if we consider two types of example.

First, there are cases in which performing the action which seemed to the agent to be suitable to his situation led to disaster, even when the agent was acting on a view that was generally accepted. For example, there have been mass ergot poisonings of people who acted on the accepted view that bread nourishes, and many people in totalitarian societies were betrayed by their best and most-trusted friends (Popper 1983, 63). Other examples are the drivers in the last few years who came to grief when they acted on the proposition, that was generally shared, that Toyota cars are safe; and the deaths from variant Creutzfeldt-Jakob disease of people who acted on the assumption, generally accepted, that it is safe to eat beef bought from the supermarket.

Second, there are cases where an agent has achieved success by acting against the generally accepted view. On the basis of the most advanced optical theory (his own), Kepler thought it impossible that a telescope could be constructed; but Galileo proceeded to construct one by trial and error, relatively successfully, and afterwards received Kepler's praise (Feyerabend 1975, 105, footnote 21). Edison's production of his electric light refuted the unanimous scientific opinion that such a light was impossible (Kuhn 1977, 238). Some authors whose first novel was rejected by innumerable publishers and agents successfully refuted the expert consensus by self-publishing their novel which then became a top seller. Similarly, more than twenty scientists have won a Nobel Prize for work that was rejected for publication by many journals (Nature Editorial 2003, 645): persisting in their work meant acting contrary to the general-

ly accepted view of its worth. In all these cases, people acted in contradiction to the generally accepted view because they regarded that view as false and they set out to refute it. But, given that they were successful, it behoves us all to put our own entrenched views to the test by sometimes setting out to do what those views proclaim to be impossible.

It might be objected that there are some ways of acting against one's accepted view, or the generally accepted view, or failing to act in conformity to it, that are irrational because they risk substantial adverse consequences for the agent or for other people, as with someone who leaves a building via a tenth-floor window because, for all we can know, he *might* just float to the ground gently, or a military commander who orders his troops to face oncoming fire because, for all we can know, the bullets *might* just pass through them harmlessly. We could accept such cases as irrational without jeopardy to the point that there is always an alternative action available to a rational agent. For, it seems likely that there will always be some available alternative that, according to currently accepted theories, does not seriously threaten substantial adverse consequences.

A more robust response to the objection is also possible. For, the laws of nature which are posited in our theories may reflect a particular kind of stability which, in the cosmic scale of things, is just a blip; and, for all we can know, that blip may come to an end in the next moment. Thus, all our best theories may have been tested against a state of the world that is only temporary and almost over, so they may all be about to be refuted. How, then, can it be more rational to act as if these theories are true than to act as if they are not? Of course, given our ingrained expectations and fears, it will be more *natural* to do the former than the latter; but that is a different matter. Although I fully endorse this robust response to the objection, the argument of this paper does not require it, so I will not press it here.

It is not only our command of the facts that is fallible. Our command of our values is also. When we are considering what we should do, we consider not only the facts about how things stand and the options for changing them, but also our aims and relevant values and norms. We may consider not only the normative propositions we accept, but also what normative propositions we *should* accept; and we can make mistakes about both questions. This is evidently the case for big decisions, such as buying a house, where the range of values is large, diverse and often conflicting, covering the different things we find desirable not only in a property but also in neighbours, local amenities, transport links, and convenience for work, family and friends. There is a very real risk of overlooking an important value and of being mistaken about which things we regard as important or most important, as many of us will know from unhappy experience. But even in simple decisions it is easy to forget or overlook some relevant desirable features. For example, when purchasing a light bulb I may decide that I value high wattage and low cost, and make my choice accordingly; but when I get home I realise that I wanted a screw fitting rather than a bayonet fitting (I might have known this all along but just forgot about it when making my purchase). The agent is therefore rationally entitled to question whether the values that seem to him relevant to his decision really are all and only the relevant ones.

Even in the unlikely event that we have assembled all the relevant factual and normative statements, and got them all right, we may still be mistaken in the conclusion we draw from them about what we should do. This is plainly the case with regard to complicated decisions involving complex situations and a wide range of options and values, where our premises and conclusion are not put into logical form (so we do not have a mechanical check for validity). Since every proposition entails an infinity of consequences, we can easily fail to notice some significant and relevant implications and thereby come to a conclusion which does not follow from our premises (Watkins 1970, 208). But it is also the case with regard to simple decisions in which the conclusion is derived from just a few premises by means of an argument which is formally valid in standard logic. For, we may doubt whether the standard inference rules actually hold in this case. This seems so no matter what standard rule of inference is in question, since every significant rule of inference has been impugned by competent logicians as a means to resolving serious logical problems (Priest and Thomason 2007, 96-98). Therefore, the action that seems to the agent to be suitable to his situation is not necessarily one that is implied by his relevant descriptive and normative views: it may be one he mistakenly infers to be suitable, either because he makes a slip in employing rules of inference he accepts as sound, or because the rules of inference he accepts are not sound. And this possibility of error means that the agent is rationally entitled to question whether his inference is valid, no matter how evident it seems to him. It is eminently rational to question and (if we have time) criticise received opinions, including opinions about the validity of arguments.

Thus, no matter how far and how conscientiously an agent's investigations have proceeded, even where everything he believes or accepts points in favour of a particular course of action, he is rationally entitled to question whether that action really is suitable to his situation. And if he does question this, he may act in accordance with his current views (given that he has no better) or he may try to refute his current views by acting against them (given that they may be relevantly false). Whichever choice he makes he will not be in violation of the rationality principle, since either way he will be acting suitably to his situation of doubt. Even if he acts in a way which, if successful, would refute what he regards as the best-corroborated appraisal of the situation, he need not be choosing an option he regards as, in any sense, sub-optimal; indeed his overriding concern may be to do what is best in the situation. As Popper puts it (1972b, 56):

> [I]t is perfectly reasonable to *act* on the assumption that [the future] will, in many respects, be like the past, and that well-tested laws will continue to hold (since we can have no better assumption to act upon); but it is also reasonable to believe that such a course of action will lead us at times into severe trouble, since some of the laws on which we now heavily rely may easily prove unreliable.

There are many places (including 1972a, 282; 1973a, 21-22; 1978, 103-104; 1983, 61-62, 65) in which Popper insists that it is rational to act on the theory that has stood up best to criticism so far. If that is understood to mean that acting contrary to the best-corroborated theory is irrational, it would appear to contradict his fallibilism as well as the quotation just given. However, although he does not say it clearly or consistently enough, Popper generally seems to be saying that acting on the best-corroborated

theory is rational in the sense that it is rationally *permitted*, just as it is rationally permitted to act in a way that would refute the best-corroborated theory. We have a choice between equally rational options. The fact of the matter is that *whenever* we act we are making a guess.

Therefore, even where an agent is in a situation which, on the view of it he finds most plausible, is single-exit, if he holds that view open to question or criticism it is not instrumentally irrational for him to try to refute that view by choosing an alternative action (or no action); in which case he has more than one suitable option, so the situation will not appear to him to be single-exit. But the rationality principle can be used to explain or predict a *specific* action only in a situation the agent takes to be single-exit. The same argument can be run through for *partial* explanations and disjunctive or vague predictions. For, the rationality principle can be used to generate such explanations or predictions only where there is a circumscribed set of possible actions that appear suitable to the agent. But since a critically rational agent may doubt the descriptive and evaluative views, and the inferences, that generate the set of suitable options for action, it would not be instrumentally irrational for him to perform an action that is not in the set. It may seem that the rationality principle is thereby rendered impotent in the explanation of social events, thus making social science, and the teleological explanation of human action in general, impossible. That this is not so will be shown in section 4.

4. Synthesis: Popper on Limited Rationality and Indeterminism

We began section 3 by noticing that a mad person could exemplify instrumental rationality even though she fails to exemplify critical rationality. Somewhat similarly, we can conceive of someone who complies fully with the instrumental principle, always performing actions she judges suitable to her overall aims, but who never stands back from her aims, beliefs and values and subjects them to critical appraisal. We could, if we knew enough about her aims and information, use the rationality principle to explain her actions in situations that appear to her to be single-exit and to give a partial explanation of her actions in situations that appear to her to be multiple-exit (with a circumscribed set of suitable options). But her lack of critical rationality would make her more like an automaton than a rational human being (Popper 1966, 380-81; 1973a, 24-27; 1982a, 81-85; Popper and Eccles 1983, 75-81, 109-11, 144-46). A rational human being will exemplify both critical and instrumental rationality.

One might take Popper's description of instrumental rationality as a 'minimum principle' to entail that, if a person exemplifies critical rationality, he must *always* exemplify instrumental rationality. However, that would be a mistake. People who exhibit critical rationality may sometimes fail to satisfy the principle of instrumental rationality and thereby act irrationally. Examples are well-known and their irrationality is commonly accepted. They include the following types: an action prompted by emotion which the agent would affirm to be detrimental to his overall aims if only he paused to think about it (Popper's frustrated drivers); akratic actions in which the agent does something that he judges to be unsuitable to his overall aims, or worse, even though he struggles against doing it (these are a staple of literature, for example,

Tolstoy 1911); and froward actions, in which an agent may act in a way he judges unsuitable simply because he can.

It may therefore seem that the two types of rationality, critical and instrumental, are independent of each other. But I think that is not so. We can express the single-exit version of the instrumental principle as the following *rational requirement*.

> (*) Instrumental rationality requires of an agent that, if she deems a particular action the only one that is suitable to her overall aims, then she tries (or wills) to perform that action.

A rational requirement expresses a relationship that must hold between a person's mental states if that person is to be rational (Broome 2005). In (*) the relevant mental states are a deeming and a trying (or a willing). This principle is a rational requirement because to act in a way one regards as unsuitable to one's overall aims is to intend self-frustration. Now suppose that one of our overall aims is to obtain as much knowledge of the world as we can, and suppose that we acknowledge that we are thoroughly fallible. Given that aim and that information we may infer the conclusion that we should make it our policy to hold *all* our views open to criticism, so that we may be able to identify errors and eliminate them, thereby improving our knowledge. It follows from (*) that instrumental rationality requires that, if we deem this policy as the only one of a set of rival policies that is suitable to our overall aims, then we should adopt that policy in preference to any of those rivals, that is, we should be critically rational with respect to all our views. However, this derivation of full critical rationality from instrumental rationality depended upon the agent accepting a specific aim (to acquire knowledge wherever possible) and specific information (we are always fallible) and drawing a specific inference (therefore we should hold all our views open to criticism). People who dispute the inference, or who reject the aim (perhaps their overall aims include holding fast to some specific beliefs or theories), or who deny the information (perhaps they think some of us are infallible on some issues), can be instrumentally rational without being fully critically rational. Still, they can be expected to exercise their capacity for critical rationality selectively; and they will do this in compliance with the instrumental principle where they aim to improve their knowledge in an area where they think they might be mistaken and they do not doubt the inference that they should therefore hold their relevant views open to question. The rationality of critical rationality can therefore be explained in terms of instrumental rationality; and perhaps this gives the sense in which instrumental rationality is a minimum principle of rationality.

As we noticed in section 3, however, an agent may comply with (*) either by trying to perform the action he deems the only one suitable, or by ceasing to deem it the only one suitable by questioning his current views. That is what seems to rob the rationality principle of explanatory value. However, in practice, there will be many cases in which an agent views his situation as single-exit and in which he does not question his view of the situation, where he in consequence may report that he 'has no choice.' This is not just because most people are unfamiliar with, or unconvinced by, the arguments for fallibilism. It is also because our rationality is limited. Most of our views are inherited biologically or culturally so, even for fallibilists, questioning a view is in

most cases subsequent to having the view, if it happens at all (Popper 1972b, 49-52; 1972c; 1994c, 134-39). Indeed, since belief seems to be our default condition and doubt requires effort, even if an agent has a policy of holding *all* his views open to question, he might never have got around to doubting the views involved in his situational appraisal; and even if he has in the past questioned (some of) those views, he might not be doubting them at the time he takes his action. Although the non-doubting agent is not *rationally required* to perform the action he views as the only one suitable (since he still has the option of doubting whether it is the only one suitable), he will perform that action if he does not doubt and he acts rationally. There will therefore be many cases for which the rationality principle can be straightforwardly employed to give an explanation, or prediction, of a specific action, namely, where the agent views his situation as single-exit and does not question that view. The same argument shows that there will also be many multiple-exit cases for which the rationality principle can be employed straightforwardly to give a partial explanation or a disjunctive or vague prediction.

This recognition of our limited rationality therefore permits some scope for explaining human action. But the scope is restricted to situations where the agent does not reflect critically on her view of her situation; and it can be used to generate explanations or predictions of *specific* actions only where, *in addition*, the agent's view of her situation identifies just one action as suitable. However, we can give the rationality principle, and thus social science, a much wider range of explanatory and predictive application if we invoke another element of Popper's philosophy, namely, his indeterminism.

In Popper's view (1982b, 46-86,159-211; 1983, 358-60) we live in an indeterministic world, with an open future. From our limited point of view the world presents a largely deterministic appearance, which is why *prima-facie* deterministic theories can be used by us for prediction with reasonable success. But in reality every situation is *objectively* indeterminate in that there is always a range of possible outcomes each with its own probability value (the sum of the probabilities of all the possible outcomes being 1). The objective probabilities in question are probabilities relative to the total present state of the universe (Miller 1994, 182-90). Popper explained his indeterminism in terms of propensities rather than frequencies; but we need not take a position on that issue for our purposes here. What is essential, though, is that the probabilities in question are construed objectively (ontologically) rather than subjectively (epistemically).

Turning to human actions, we can conjecture that in every situation of action the agent has real options, it being more probable that he will select some of them rather than others, and that the agent, by acting, selects which one of these options to realise (see my 2010). The rationality principle can then be recast as a set of propositions about the objective probability of different types of action. As a first stab, these may include the following:

(a) when an agent views his situation as single-exit and does not pause to question his view of the situation, then, that he will try to take the single exit has a probability value close to 1 (but it will fall short of 1 because he might act irrationally);

(b) when an agent views his situation as single-exit but is conscious of his fallibility in this matter, then, that he will try to take the apparent single exit has a probability value close to 1, though perhaps lower than the probability value referred to in (a);

(c) when an agent is in a situation in which his descriptive and normative views would be taken by the vast majority of people to imply logically that just one particular action is suitable to his aims, then, that he will try to take that particular action has a probability value close to 1, though perhaps lower than the probability value referred to in (b).

The reason there may be a decrease in probability as we pass from (a) to (c) is that, apart from (a), each allows for an additional type of failure to take the apparent single exit compared to its predecessor. None of these propositions requires for its truth that the probability value referred to will be the same for different agents. Proposition (a) translates Popper's claim that the rationality principle is false but approximately true. Proposition (b) is a low-level psychological hypothesis to the effect that, while there is nothing rationally to choose between acting on the most plausible (or best-corroborated) representation of the situation and trying to refute it in some way, nevertheless, people will generally prefer to act on it. This allows us to treat the situations described in (b) as close to the situations described in (a). Proposition (c) allows us to assume that people can generally be expected to reason in a commonplace way (which, as behavioural economists and evolutionary psychologists may point out, need not be a logically valid way). This gives scope for the application of a version of Popper's logic of the situation, and it allows us to treat the situations described in (c) as close to the situations described in (b). A parallel trio of propositions can be enunciated for situations in which the agent views his situation as offering a restricted set of options each of which is suitable to his aims. Perhaps it is fanciful to suppose that we could, through research, obtain numerical values for the probabilities referred to in (a), (b) and (c); but we probably do not need them, for two reasons.

First, *in the historical social sciences*, we can assume, for the purpose of testing an explanation, that actions with a probability value close to 1 actually have a probability value of 1. This is analogous to Popper's demand that the rationality principle (in his version) be treated as if it were true when we test a social-scientific explanation. A historical explanation will then be treated as refuted if the action it implies is very highly probable did not occur. But we can permit the explanation to be saved from this refutation if it can be bolstered by an auxiliary hypothesis which conflicts with our assumption that the probability value of the predicted action was 1, *so long as this auxiliary hypothesis passes an independent test*. The auxiliary hypothesis could affirm that the agent acted irrationally (a type (a) reason for a probability value less than 1) or that he has a contrary disposition (a type (b) reason) or that he is a 'deviant' logician (a type (c) reason).

Second, *in the theoretical social sciences*, high probabilities are usually all we need to explain social events and social regularities. For, social events and regularities are usually explainable by the *preponderance* of similarly placed individuals who act in the same way. For example, except in special circumstances, a price rise for a particular good will

lead to lower demand, other things being equal. This holds even if a relatively small number of people do not lower their demand and even if some of these people increase theirs.

Finally, as (a), (b) and (c) are such low-level psychological hypotheses, the small concession to 'psychologism' that they involve is one to which Popper might not have demurred.

5. Conclusion

Social-scientific explanation is characteristically teleological, explaining social events in terms of purposive human actions and their unintended as well as intended consequences. Teleological explanation depends upon the rationality principle, according to which people act suitably to their situation as they see it. And social-scientific explanation is universal in the sense that each agent is taken to respond to the 'logic of the situation' and thus to do what anyone else would do in the same circumstances. Popper argued that while the rationality principle is false, it is approximately true and must be treated as if it were true in testing social-scientific explanations. However, he overstated the explanatory value of the principle and of situational logic, because he overlooked:

- the frequency of situations which appear to the agent to be multiple-exit;
- that a rational agent's recognition of her fallibility can turn an apparent single exit into an apparent multiple exit, and an apparent restricted multiple exit into an apparent open-ended multiple exit;
- that a rational agent's reasoning might not comply with the logic assumed by the social scientist.

The first point highlights a limitation of situational logic, which means that we often have to rest content with partial explanations. However, the second and third points can largely be overcome. Popper's account of our limited rationality permits significant scope for the explanatory use of the rationality principle despite our capacity for critical rationality, and his indeterminism permits a reformulation of the rationality principle which assigns probability values close to 1 to an agent's tendencies to act on tried-and-tested views, to draw the inferences that the vast majority of people would do in a similar situation, and to avoid action she recognises as self-frustrating. The testability of explanations in theoretical social science should not be affected by the introduction of these probabilities, because social regularities require only a mass, rather than a universal, uniformity. The testability of explanations in historical social science can be achieved by treating probabilities close to1 as equal to 1, except in special cases where there is a hypothesis to the contrary that has passed an independent test.

An agent exemplifies instrumental rationality when she acts suitably to her situation as she sees it, given her aims. It is doubtful that anyone is fully instrumentally rational. An agent exemplifies critical rationality when she has the capacity and the tendency to hold her views open to criticism. Although one may hold a policy of holding all one's views open to criticism, few people are thus fully critically rational. One could

conceivably be consistently instrumentally rational without being at all critically rational. Yet critical rationality is instrumentally rational for an agent insofar as she aims to acquire knowledge, accepts her fallibility and infers the natural conclusion.

REFERENCES

Bratman, Michael. 1999. Practical Reason and Acceptance in a Context. In his *Faces of Intention*, 15-34. Cambridge: Cambridge University Press.

Broome, John. 2005. Does Rationality Give Us Reasons? *Philosophical Issues* 15: 321–37.

Chang, Ruth (ed.). 1998. *Incommensurability, Incomparability and Practical Reason.* Cambridge, MA.: Harvard University Press.

Feyerabend, Paul. 1975. *Against Method.* London: New Left Books.

Frederick, Danny. 2010. Popper and Free Will. *Studia Philosophica Estonica* 3.1: 21-38.

Gorton, William. 2006. *Karl Popper and the Social Sciences.* New York: State University of New York Press.

Kuhn, Thomas. 1977. The Essential Tension. In his *The Essential Tension*, 225-39. Chicago: University of Chicago Press.

Latsis, Spiro. 1972. Situational Determinism in Economics. *British Journal for the Philosophy of Science* 23:207-45.

Miller, David. 1994. *Critical Rationalism.* Chicago: Open Court.

Nature Editorial. 2003. Coping with Peer Rejection. *Nature* 425/6959: 645.

Popper, Karl. 1966. *The Open Society and its Enemies*, Volume 2, fifth (revised) edition. London: Routledge and Kegan Paul Ltd.

—. 1972a. *The Logic of Scientific Discovery*, sixth impression (revised). London: Hutchinson.

—. 1972b. Science: Conjectures and Refutations. In his 1972d, 33-65.

—. 1972c. Towards a Rational Theory of Tradition. In his 1972d, 120-135.

—. 1972d. *Conjectures and Refutations*, fourth (revised) edition, London: Routledge and Kegan Paul.

—. 1973a. Conjectural Knowledge: My Solution of the Problem of Induction. In his 1973c, 1-31.

—. 1973b. On the Theory of the Objective Mind. In his 1973c, 153-90.

—. 1973c. *Objective Knowledge*, corrected edition. Oxford: Oxford University Press.

—. 1978. *Unended Quest.* Glasgow: Collins.

—. 1982a. *The Open Universe.* London: Routledge.

—. 1982b. *Quantum Theory and the Schism in Physics.* Totowa, NJ: Rowman and Littlefield.

—. 1983. *Realism and the Aim of Science.* London: Routledge.

—. 1994a. Models, Instruments, and Truth. In his 1994b, 154-84.

—. 1994b. *The Myth of the Framework.* London: Routledge.

—. 1994c. *Knowledge and the Body-Mind Problem.* London: Routledge.

Popper, Karl. and John Eccles. 1983. *The Self and its Brain.* London: Routledge and Kegan Paul.

Priest, Graham and Neil Thomason. 2007. 60% Proof. *Australasian Journal of Logic* 5: 89-100.

Tolstoy, Leo. 1911. 2003. *The Devil.* In his *The Devil and Other Short Stories,* translated by Louise and Aylmer Maude, revised by Richard Gustafson, 191-236. Oxford: Oxford University Press.

Watkins, J. W. N. 1970. Imperfect Rationality. In *Explanation in the Behavioural Sciences*, ed. Robert Borger and Frank Cioffi, 167-217. Cambridge: Cambridge University Press.

DANNY FREDERICK has degrees from the London School of Economics and Birkbeck College (London). He is currently doing independent research in philosophy and has a webpage at: http://independent.academia.edu/DannyFrederick. His main research interests are in rationality, reasoning, free will, freedom and their interconnections.

ADDRESS: Slate House, Hunstan Lane, Old Leake, Lincolnshire, PE22 9RG, UK.
Email: dannyfrederick77@gmail.com