

- 34 *Ibid.*, p. 66.
- 35 *Scepticism*, pp. 47 ff.
- 36 And very possibly also, as Zinkernagel claims, in certain conditions for description (op. cit., see Note 7). But one must distinguish two senses, or applications, of 'unreasonable' here. Unless the arguments for the law of contradiction, or of that of Zinkernagel's conditions for description, are easily surveyable, or may still be unreasonable of the sceptic to accept the conclusions without reservation. The more complex an argument the greater the risk of uncritical conceptual commitment and also of failure of what Hume calls 'that faculty which judges' (David Hume, *A Treatise of Human Nature*, ed. and intro. by D. G. C. Macnabb, Meridian Books, Cleveland and New York 1962, p. 233).
- 37 However, according to Sextus, equanimity or mental quietude is a *consequence* of suspension of judgment, the latter in turn *due* to an inability to find such peace of mind when searching for the definite truth - in view of the conflicting opinions which the developing sceptic has been exposed to (cf. Sextus, op. cit., pp. 19-21). So one might argue, against this psychological objection, that the sceptic's vast learning is the result of his earlier, vigorous but unquiet period of quest, and that as a mature sceptic he is no longer his old inquisitive self. But then he sounds less like a 'seeker' (cf. Note 9) than a seeker in retirement, an ex-seeker, or if a seeker, perhaps one whose old enthusiasm returns only in the urgent quest for counterarguments when an overbalance in favour of a proposition threatens to extort a decision from him, or when he needs counterarguments to give a semblance of rationality to what is perhaps in effect a pathological inability to decide. The sceptic and his defender may feel a little alarmed now, with so few strokes of the pen, a description of the healthy- and open-minded sceptic (cf. Naess's 'Psychological and Social Aspects of Pyrrhonian Scepticism', *Inquiry*, Vol. 9 [1966] and Ch. III of *Scepticism*) can become a portrait of an anxiety-ridden neurotic. So perhaps there is an even more delicate balance which the sceptic must preserve, where the crucial factor is his relation to his peace of mind. When he ceases to accept it as a welcome but not deliberately sustained effect of suspension of judgment, and begins to regard it as so important that he's ready to create the conditions for suspension artificially, then his mind can no longer be described as open.
- 38 See Ayer, op. cit., p. 41; cf. p. 75.
- 39 I am very grateful to Carolyn Black for critical comments which I hope have led to improvements in the presentation of a number of points in this final version.

## Popper's Metaphysical Research Program for the Human Sciences<sup>1</sup>

Noretta Koertge  
Indiana University

Popper has provided a model for the scientific explanation of human actions and a metaphysical theory of man which can guide scientific research. In this paper I discuss the problems of the empirical content and nomicity of the Rationality Principle and extend the method of situational analysis to the problem of explaining beliefs. The domain of applicability of the Rationality Principle is bounded on one side by cases in which behavior is determined by processes which can not be influenced by criticism and on the other side by the phenomenon of substantive creativity. However, a large part of human activity lies within its scope.

### I. Introduction

People have always had an ambivalent attitude towards scientific inquiry. We are drawn into scientific explorations through the curiosity drive we share with other animals and the sheer joy of discovery and understanding. At the same time, the pursuit of knowledge is accompanied by feelings of fear - or at least, caution. With knowledge comes power and early man was convinced that the gods were not entirely happy with his ability to produce at will the fire and iron which had previously come only as gifts from the heavens. Anxiety about the vengeance of the gods disappeared through criticism of the theological theory involved. But our fears of scientific inquiry today cannot be so easily dispelled. We know that knowledge of the necessary or sufficient conditions for a phenomenon is a major first step in learning how to control it. The discovery of a practical method of removing a necessary condition will allow us to prevent the phenomenon in question. If we can learn to instantiate the sufficient conditions, we can create the phenomenon on demand. The distance from scientific understanding to engineering capability varies from case to case, but it is independent of the social value of the outcome.

Our ambivalence is especially deep and fundamental in the case of scientific inquiry into the activity of human beings. Nothing is more fascinating than the problem of understanding why people do the things

they do. No other domain of phenomena offers a bigger intellectual challenge to the scientist looking for uniformities, for invariances, for explanatory laws. All people behave both wisely and stupidly, selfishly and generously, creatively and blindly. They are often predictable – but only for a time and it seems impossible to predict when they will stop being predictable!

And yet in no other domain is the prospect of scientific success more frightening. We want desperately to understand our own species and to use this understanding to prevent needless suffering and humiliation. Yet the very scientific knowledge which would provide understanding seems to threaten our most fundamental values – those of freedom and dignity. Every modern metaphysical theory of man provides some resolution of this conflict between our desire for a science of humans and our desire for human autonomy. Existentialism celebrates man's ability to act, to choose, to affirm what is valuable. Existential man is even free to act absurdly without any motives at all. (In *Les Caves du Vatican* Gide describes Lafcadio's crime as an *acte gratuit*.) On this theory man is certainly autonomous, but he is also fundamentally beyond explanation. We may condemn the actions of Camus's *Étranger* or admire those of the doctor in *La Peste* but we cannot hope to give a scientific explanation of them.

Behaviorism provides an extreme response in the other direction. Behavioristic man is a soft machine, a clockwork orange. As such, he is eminently explicable, but beyond freedom and dignity. Yet one need not fear behavioral modification techniques, we are told, because positive reinforcement is more efficient than punishment. So we will find the experience of having our behavior shaped a very rewarding one.

Popper's theory of man denies that there is a contradiction between human science and human freedom. Popper says that we can give a scientific explanation of human activities by analyzing people's problem-situations, their aims, their theories about what options are open to them, and their evaluations of those options. We may influence their behavior by providing them with additional relevant information and by criticizing mistaken elements in their own assessments of their situations. Yet this in no way impairs or limits their freedom. They must appraise and evaluate our information and criticism – they must decide whether to act on it.

It is people's ability to react appropriately to their situations and to respond to criticism which makes them rational and autonomous. It is

also what makes their behavior accessible to scientific inquiry. Thus there is no fundamental conflict between a science of human behavior and the values of human freedom and autonomy.

Each metaphysical theory of man suggests not only a program of scientific research but also a political theory. Existentialism is especially compatible with anarchism; behaviorism lends support to the idea of a dictatorship of benevolent engineers; Popper's conception suggests a liberal democracy, one which maximizes the opportunities for criticism. It is because I find Popper's political theory so attractive that I am determined to subject his scientific research program to critical scrutiny. Students of ideology have warned us that the moral or political attractiveness of an idea may interfere with our appraisal of the strength of its claim to be true.

I will describe in detail both his model for explaining human behavior by means of situational analysis and the Rationality Principle and his heuristic program for discovering such explanations. I then raise critical questions about both the model and the research program:

(1) *Adequacy*. Are situational explanations adequate scientific explanations? Do they satisfy all of the requirements of the covering law model, especially that of nomicity?

(2) *Scope*. Are all human activities open in principle to explanation in terms of the Rationality Principle? What about creative activities such as the invention of new scientific theories? What about the development of tastes?

(3) *Fertility*. How useful is Popper's approach to the working social scientist? What avenues of research does it suggest? How could his research program be strengthened?

I have tried to give an accurate representation of the views found in Popper's later writings, especially those in 'Of Clouds and Clocks' (reprinted in Popper, 1972) and 'La Rationalité et le Statut du Principe de Rationalité' (1967). Nevertheless, this paper is interpretative and critical, not primarily exegetical – so much so that Popper, following the example of Marx, may very well wish to declare, 'Moi, je ne suis pas Popperien!'

## II. Popper's Account of Situational Explanations

Popper's remarks concerning explanations based on the Rationality Principle (RP) are puzzling. On the one hand he says that the RP is an animating principle, 'une loi d'animation' (1967, p. 144), which plays a role

in situational analyses similar to that of Newton's Laws in the explanation of motions within the solar system. (The description of the agent's situation corresponds to the description of initial and boundary conditions in explanations found in natural science.) However, he also stresses that the *RP* is almost empty (1967, p. 144) and false (1967, p. 145). He says that it does not play the role of an empirical, explanatory theory and is not a testable hypothesis (1967, p. 144).

Furthermore, whereas Popper exhorts the physical scientist to look for true theories of high empirical content by ruthlessly weeding out false ones through severe testing, Popper's methodological advice to social scientists is never to abandon the *RP*. Instead, they should always revise their models of the agent's situation.

If there are in fact no candidates for better or deeper laws of human activity and if one always revises one's model of the situation in a content-increasing, testable way, there is nothing disreputable about Popper's methodological advice to the social scientist. Yet one wonders why the Rationality Principle should enjoy this privileged status: Is it simply because there are so many practical difficulties in subjecting it to experimental tests? Are there other ways of submitting the *RP* to criticism?

Neither need we be too shocked by the claim that the *RP* is literally false. (Popper thinks it doesn't hold for the case where a person repeatedly tries to park a car in a space which he or she knows is too small.)<sup>2</sup> It may be a true statistical generalization or it may serve as an ideal law (Popper sometimes calls it a 'zero principle') which can then be augmented by a theory of when deviations from the simple ideal case are to be expected. But again, one would like to know exactly what the program is for improving the *RP* account of human activities.

One would also like to be clearer both about the content of the *RP* and about its role in explanations. In order to have before us an explicit explanatory schema, I propose the following as an informal model which is suggested by the discussion so far.<sup>3</sup>

1. *Description of the Situation:* Agent *A* was in a situation of type *C*.
2. *Analysis of the Situation:* In a situation of type *C*, the appropriate thing to do is *X*.
3. *Rationality Principle:* Agents always act appropriately to their situations.
4. *Explanandum:* (Therefore) *A* did *X*.

The above schema is restricted to simple cases in which there is a single best response to the situation. It would be quite easy to adapt it to cover cases in which there are various options of equal appropriateness. It would have to be modified in order to handle cases in which the agent is physically impeded from carrying out the action which he calculates to be the most appropriate. We would also have to augment the schema in order to explain the unintended consequences of actions.

Perhaps the most immediately problematic feature of the above schema comes from the words 'appropriate' and 'rational'. How much normative or approbative force are they intended to carry? Exactly what is the content of the *RP*? What is its domain? What precise claims are made about the domain? In short, what sort of behavior is ruled out by the *RP*? As Watkins remarks, 'Popper's own expositions of the *RP* are "appropriately vague"' (1970, p. 172). The above formulation of the *RP* is based on the following quotation:

En effet, le principe de rationalité me semble certainement faux, même dans sa formulation la plus large, celle adoptée ici, qui peut s'énoncer de la manière suivante: 'Les individus agissent toujours d'une manière adaptée à la situation où ils se trouvent'. (1967, p. 145)

The exegetical problem is increased because his views on the *RP* (or at least his emphases) seem to have changed somewhat over time. In his earlier writings on the subject Popper limited the scope of the principle to the actions of sane people. Thus, in *The Open Society and Its Enemies* he spoke of '... tacitly [assuming], as a first approximation, the trivial general law that sane persons as a rule act more or less rationally...'. (1962, Vol. 2, p. 265), where by 'rational' he meant '... in accordance with the logic of [the] situation...'. (1962, Vol. 2, p. 97). Later on, Popper stressed that the principle could also be applied to the actions and beliefs of a madman (1967, pp. 147-8). Watkins (1970) suggests that it is precisely in cases of seemingly crazy, botched-up actions that the *RP* is most illuminating.

As he explicitly widened the domain of applicability of the principle to include madmen, Popper also weakened his claims about the kinds of actions which agents could be expected to perform. Where he had earlier spoken of actions as being 'rational' or 'appropriate', he now tended to characterize them as 'adequate', or 'adapted', or 'in accordance with' the situation.<sup>4</sup> These terms may indicate that fewer restrictions are to be placed on the kind of fit between situation and action than those sug-

gested by the term 'appropriate'. But perhaps the major reason for this change in terminology was his increasing emphasis on the fact that the situation which was central in the explanation was not the agent's objective physical-physiological-psychological situation. Rather it was the agent's *theory* of his situation — the situation as the agent saw it: '... leur actes . . . peuvent être considérés (au moins avec un bon degré d'approximation) comme adaptées à la situation telle qu'ils la voient' (1967, p. 147. Italics in the original).

Of course, given Popper's characterization of knowledge as objective, the agent's theory of his situation can also be considered as an objective component of reality. Nevertheless, people's *ideas* about their situations are often difficult to investigate experimentally. In general, the bigger the discrepancy between the agent's theory of his situation and his actual situation, the less easy it will be to apply the *RP* to practical cases. The more conjectural (or far-fetched) our reconstruction of the situation in which the agent found himself, the more important it is to seek independent evidence for *our* theory about the agent's theory of his situation.

In his recent work, Popper has suggested that the method of situational analysis be extended into the domain of animal behavior and evolutionary development (1972). But this unified treatment of human and animal activities has been bought at a considerable price. Although animal behavior can be purposive and thermostats may be viewed as teleological systems, human actions are distinctive because of their susceptibility to change by argument.<sup>5</sup> Since one can prevent or modify an action through the input of information, one infers that the agent's *theory* of the situation exerts some influence on his action. Knee-jerk reflexes and the behavior of amoebae are not changed by arguments and so one infers that they are to be explained in different terms.

Thus there are practical scientific reasons for distinguishing between the theories, problems, and solutions which figure in the explanations of human actions and those which we may ascribe to the lower animals.<sup>6</sup> It should also be noted that unless there are fairly marked differences between the correct explanatory accounts of animal and human behavior, much of the plausibility of Popper's theory of the open society disappears.

### III. The Empirical Content of the Rationality Principle

I will now give a more precise formulation of the *RP* and specify its empirical content. In doing this I will exploit analogies with Newton's Laws as much as possible. I trust the disanalogies will be blatantly obvious.

Newton's Second Law ( $F = ma$ ) is a biconditional — it says (roughly) that every acceleration is caused by a force and that any net force acting on a body causes it to accelerate. It also describes the precise quantitative relationship between the force and the acceleration. Given an acceleration to explain, the first clause of the Second Law tells us the magnitude of the force which must have been acting but it does not tell us anything about the source of the force. For this we need a supplementary theory of forces. At the beginning only a few kinds of force were known, e.g. those from springs and gravitational masses, although Newton suspected that there were also repulsive forces, short-range cohesive forces, forces which caused fermentations, etc. The law which figures in the actual explanation of the acceleration is the second clause. Once one has independent evidence for the existence of the required force, one can then explain why the acceleration occurred.

I believe that the *RP* also consists of two clauses. The first (*RP-1*) says: Every action (by a person) is a rational response to some problem-situation. This claim tells us what to look for when we are trying to explain an action. In order to test it we require a supplementary theory which tells us how to detect problem-situations.<sup>7</sup> Once we have found the requisite problem-situation, we explain the given action by using the second clause (*RP-2*): Every person in a problem-situation responds rationally to it.

Of course, both clauses of the *RP* might become tautologous should we make 'being intelligible' or 'being reasonable' part of the definition of 'action', as some philosophers of mind are wont to do. So let us conceive of action as an ethologist would, i.e. the concept must be broad enough to include the purposive, exploratory behavior of a hungry animal, but narrow enough to rule out rigid reflex movements.

Under 'problem-situation' I include not only the person's problem but also the entire set of solutions which the person considers and his evaluations of them. By saying that a response is rational I mean (i) that it was arrived at through a methodical appraisal of the set of possible solutions; (ii) that a description of both the problem-situation and the appraisal process could in principle be verbalized by the actor; (iii) that the person acted as he did because of the appraisal process (i.e. if a better alternative had been presented to him he would have taken it).

Let us look at these requirements a little more closely. In requiring that the agent potentially be able to articulate the process (requirement [ii]) we allow for unconscious motives, but rule out as rational (or reasoned) the

appropriate responses made by an amoeba. The fact that the process can be verbalized helps explain why semantic input can influence behavior. The third requirement forces us to look for the real reasons for an action — we cannot rest satisfied with a plausible rationalization. It also ensures that when we describe an agent as having used a certain appraisal rule, we in fact attribute a disposition to him, i.e. we imply that the same method of appraisal would be applied should the set of solutions be expanded. It is not an accidental feature of the agent's response to the situation.

Requirement (i) ensures that some decision procedure be used. It might seem that if we describe a response as rational we might also wish to require that the appraisal procedure which is used be one of the 'respectable' ones, such as maximizing expected utility, minimizing maximum possible loss, or using a decision scheme involving gambler-indifference curves, such as that described by Shackle.<sup>8</sup> What if an agent deliberately set out to *minimize* expected utility? Would the resulting action count as a rational one?

As far as I know, writers on Popper's *RP* have not explicitly discussed this question. Popper simply says that situational logic is the method of economic analysis (1962, p. 97). Watkins points out that we may use different decision procedures in different situations (1970, p. 207). For purposes of simplicity of exposition I will not try to specify the set of minimal requirements which admissible decision rules should obey.<sup>9</sup> Here the *RP* approach needs to be supplemented with an empirical theory of how people make decisions. Kohlberg's studies on the development of moral reasoning in children show clearly that young people's methods of analyzing and resolving dilemmas change over time in a predictable way (1971).

The *RP* approach also needs to be supplemented with a theory of mechanical errors. Billiard balls never make mistakes when accelerating in a given force field. However, two comparison shoppers in a supermarket may be in the same situation, use the same decision procedure, and yet act differently simply because one made a mistake in calculating the unit price. However, the attribution of any mismatch between our reconstruction of the agent's situation and the action which occurred to an error (or 'glitch' as the space scientists call an unexplained malfunction) will be *ad hoc* unless we have a good theory of error to justify this move.

Similar methodological care must be taken when invoking unconscious motives, repressed data, selective misperceptions of the situation, highly unorthodox utilities, etc. In each instance we must provide independent

evidence for our description of the agent's situation. This is best provided by a supplementary theory of the phenomenon in question, e.g. a theory of unconscious desires.

I will now present a more detailed model for situational explanations. It turns out that it is convenient to present the account in two stages, so *RP-2*, the second clause of the Rationality Principle, must be broken up further. Roughly what happens is that we explain the agent's rational deliberation first and then explain his rational action.

1. *Description of Problem-Situation:* *A* thought he was in problem-situation of type *C*.
2. *Dispositional Law:* For all such problem-situations *A* would use appraisal-rule *R*.
3. *Analysis of the Situation:* The result of appraising *C* using *R* is *X*.
4. *Description of Agent's Competence:* *A* did not make a mistake in appraising *R* to *C*.
5. *Rational Appraisal Principle:* All agents appraise their situations in a rational manner.
6. *Explanandum-1:* (Therefore) *A* concluded *X* was the rational thing to do.
7. *Rationality Principle:* People always act on the outcome of their rational appraisals.
8. *Explanandum-2:* (Therefore) *A* did *X*.

Thus we see that the complete *RP* account not only relates the action to a deliberation but also permits us to predict the result of that deliberation from initial conditions.

Let us now appraise the adequacy of the above explanatory model. Are there any important practical, methodological, or philosophical differences between it and the covering law explanations typical in the natural sciences? I will discuss each of the premises in turn.

Premises 1 and 4, the description of the agent's situation and competence, correspond to the description of initial conditions in more traditional scientific explanations. Although the entities referred to are different, we require that statements about the agent's situation and his appraisal of it not be *ad hoc*. There must be independent evidence for each component of the description. All this is exactly as in physics or any other science. There are certainly practical problems in testing claims

about how the agent made the decision, but there seem to be no qualitatively different epistemological problems.

Premise 2 attributes a disposition to the agent and thus incorporates one of the laws which figures in the explanation. (For a discussion of the role of dispositions in explanations, see Hempel, 1965, pp. 457-63.) This law could be strengthened if we had a general theory of decision making.

Since premise 3, the analysis of the situation, adds no empirical information to the *explanans* it could be omitted without loss. (When Popper speaks of 'situational logic' it may be this step which he has in mind.)<sup>10</sup> We might compare it to the mathematical operation which is needed to derive the predicted path of a light ray from the description of a piece of optical apparatus and Fermat's Principle of Least Time. In actual examples of situational explanations we often do not present enough information to determine the exact nature of the analysis. It would often be very difficult to describe in detail the process of weighing the options open to the agent, especially when there is no straightforward way of assigning numbers to the various factors involved.

Perhaps the difficulty of giving a comprehensive description and analysis of human decision problems was one reason why *verstehen* theorists thought that explanations of human behavior could not be written down completely, but only grasped through an empathetic faculty. I would certainly not want to deny the enormous practical difficulties of writing these things down — this is one reason why I would not like to see juries replaced by computers which are fed the facts of the case. However, I see no reason in principle why premise 3 could not be spelled out quite explicitly.

Let us now turn to the role of the Rationality Principle itself. The above premises already show that *A* has made a rational appraisal of his problem-situation. All that remains is to assert that the agent actually acted on the result of this appraisal. Some linguistic philosophers have claimed that this link between intention and action is a conceptual one. I think that philosophers of science have argued convincingly against this position.<sup>11</sup> But however this may be, the most interesting and informative part of the explanation lies elsewhere — in the description of the process of appraisal itself.

To explain an action using the *RP* does not imply that the agent's beliefs are reasonable nor even that his method of making decisions is a good one. However, it does presuppose that the agent appraised his situation in a systematic way, that the result of the appraisal procedure is

in principle open to change should the set of available solutions change, and that there is a match between the appraisal and action. Thus the *RP* is far from being almost empty and its content can be further increased by building supplementary theories of error, decision making, and belief formation.

We may tend to think of the *RP* as uninformative simply because it is so familiar. Popper has pointed out that if a theory is successful over a long period of time we may come to view it as nothing but a set of implicit definitions or conventions (1963, p. 240). In most cases all of our inquiry is focused on finding the correct initial conditions. For example, during the recent Watergate crisis a White House spokesman made the following solemn announcement: 'I can assure you that the President will respond in a way that is appropriate.' Everyone laughed. We already knew that he would respond 'appropriately' — but *what* would he consider the situation to be — and *how* would he decide what was appropriate?

So the *RP* has a respectable amount of empirical content — although we would certainly like to increase it — but is the generalization law-like? Although nomicity is not the clearest of philosophical notions, perhaps all of us would agree that Newton's Laws are taken to assert some sort of strong linkage between antecedent and consequent. Does the *RP* have the same sort of nomic force? We can ask this question without specifying exactly what analysis the concept of nomicity should receive. Neither do we need to raise the question of the truth of the *RP*. Newton's Laws are still law-like even though false.

Before tackling this question, I would like to deal briefly with two issues which sometimes get confused with the problem of nomicity. First, there is the practical problem of open systems. Predictions about the future positions of planets may go wrong, not because the laws of physics are not really laws or because the planets have free will, but simply because a new heavy comet comes barreling through our solar system. Likewise predictions about what an agent will do may be extremely unreliable because of the impossibility of sealing off a human being from new ideas — he may create a solution to his problem which no one ever dreamt of before. But this possibility does not dictate an answer to the question of whether there is a law-like connection between his creative appraisal of the situation and his action. It may well be that problem-situations are open in principle in a way that physical systems are not, but this is not now relevant either.

Secondly, it should be noted that although I have written the *RP* as a universal generalization, to reformulate it as a statistical generalization would have no effect on the question of nomicity. Salmon has argued that the statistical generalizations which can be used to support explanations should list all of the factors relevant to the *explanandum* event (1971). This means that  $p(C,A) = r$  cannot be used in an explanation unless it is true that given any further physical conditions  $Q$ ,  $p(C,A \cdot Q) = p(C,A)$ .<sup>12</sup> The fact that the fundamental laws of physics may be statistical provides no solace whatsoever to the defender of free will. Statistical determination is just as inexorable as classical determinism. The question we are concerned with is whether it is possible in principle to specify decision situations such that given these conditions, the *propensity* for a given action to occur is determined.

Is the relationship between situation and action as described by the *RP* a nomic one? Philosophers' intuitions on this issue do not agree. Dray's answer (on behalf of Collingwood) is unequivocally, 'No' (1964, p. 14). One reconstructs the agent's situation to find out which action was appropriate. If the agent happened to perform this action, we can then say we understand why he did it, but according to Dray, no laws are involved; this is the crucial difference between historical and scientific explanations. Hempel, on the other hand, argues that explanations of rational actions involve psychological dispositions and fulfill all the requirements of the covering law model, including nomicity (1965, pp. 472-3). Others might wish to say that all of this talk about reasons is just to describe epiphenomena - the real causes are bio-chemical ones. Popper argues that there are three distinct, but interacting, ontological domains - those of physical objects, mental entities, and lastly the world of theories, arguments, and problems. This theory is not argued for in a dogmatic aprioristic way. Like all Popper's metaphysical theories, it is put forward to solve certain problems connected with science. Therefore a growth in scientific knowledge might result in a modification of his ontological view. In (1974, p. 1054) he discusses the possibility of a coalescence of worlds 1 and 2. He seems to think that the separate ontological status of world 3, the world of ideas, is more secure.

Popper speaks of the possibility that the physical world 1 is 'causally changed or causally influenced' by the (mental) world 2 (1974, p. 1055) and he often describes the relationship between theories and deeds as one of 'plastic control':

Mozart and Beethoven are, partly, controlled by their 'taste', their system of musical evaluation. Yet this system is not cast iron but rather plastic. It responds to new ideas, and it can be modified by new trials and errors - perhaps even by an accidental mistake . . . (1972, p. 254)

I take this to be saying that the connection between the system of musical evaluation and the actions of the composers is a nomic one (although obviously other factors are also relevant). There is control but we call it plastic, not because it is ineffective but because the controlling theory is open to change through criticism.

But if we take the *RP* to be nomic doesn't this position constitute a new threat for the defenders of free will and responsibility? Isn't claiming that actions are controlled by ideas just as inconsistent with the concept of a free agent as saying that they are determined by physical or psychological forces? I think not. This is not the occasion on which to attempt a re-analysis of the concept of freedom, but perhaps the following example will be suggestive.

Suppose as I return from carrying out the garbage one summer afternoon I discover that the screen door has slammed and locked. Just then the phone starts to ring and I remember that I am expecting an extremely important call. I am just on the point of kicking out the screen (and crawling through) when my best friend appears and does one of the following:

- (a) Grabs my legs so I can't move;
- (b) Says that she hates rough, unladylike behavior and will never see me again if I kick down the door;
- (c) Produces a key and unlocks the screen door;
- (d) Tells me that the side door is open.

In each of these four cases my friend will have a profound influence on my behavior - I will not carry out the action which I had planned. However, in the first two cases (note that the first intervention is physical, the second linguistic) we would describe her influence on my behavior as coercive or manipulative. We might say that she is interfering with or restricting my freedom. In the second two cases we would say that she is being helpful, broadening my horizon of possibilities, increasing my freedom of choice, etc.

Whether something which impinges on me is a threat to my freedom or not seems to have nothing to do with whether it *causes* me to act dif-

ferently or not and everything to do with whether it makes it easier or more difficult for me to realize my aims.

From this analysis I conclude that explanations in terms of the Rationality Principle satisfy all of the traditional formal and epistemological requirements for scientific explanations. We now turn to an exposition and appraisal of other aspects of Popper's approach to the understanding of human activity.

#### IV. Explaining People's Ideas

Even if one finds *RP* explanations of actions unproblematic, it must be admitted that they do not provide the answers to all of our questions about human activity. This is illustrated by the following explanation sketches. They can easily be filled out so as to satisfy the requirements for explanations in terms of the Rationality Principle; yet they seem exceedingly unsatisfying.

(1) *Question*: Why is that hungry girl eating a lard and catsup sandwich (when peanut butter and jam are readily available)?

*Answer*: Because she *likes* lard and catsup sandwiches (better than peanut butter, etc.).

(2) *Question*: Why is that boy who was sent out to make ice putting hot water in the ice cube tray (when cold water is readily available)?

*Answer*: Because he believes hot water freezes faster than cold.

We should note first of all that the above answers are not uninformative. They tell us, for example, that the girl is not eating the strange sandwich in order to win a bet and that the boy's behavior is not designed to anger his mother. (Recall that the description of the situation is supposed to include *all* of the factors relevant to the choice.) However, these examples indicate that often the real point of an inquiry is to find out why people have beliefs and preferences which we consider to be very strange, either because they seem unfounded or else because they are so startlingly novel or creative.<sup>13</sup>

For this reason we need to supplement our model for the explanation of actions with a model for the explanation of propositional attitudes. If we wish to explain people's beliefs, ideas, or theories the typical explanatory question will be: Why does Jones consider *p* to be *A*? The content of

*p* can be anything one likes, a factual claim, a statement about values and duties, a mathematical theorem. *A* will then describe Jones's assessment of *p*. For example, Jones may consider *p* to be true, false but of high truth content, interesting, irrelevant, heuristically valuable, pragmatically useful, simple in form, etc.

We can break this basic explanatory question about Jones's idea that *p* is *A* into two subquestions:

(1) How did Jones come to entertain *p* at all?

(2) Why did Jones decide *p* has property *A*?

The first is the problem of explaining the *origins* of theories and the second is the problem of explaining the *appraisals* of theories.

How do we explain why Jones came to entertain *p* in the first place? In by far the majority of cases, the answer to this question poses no fundamental problems since most of our ideas are transmitted to us by our culture. There are interesting questions about the retrieval and selection of information from our memory, but the question of how I came to consider the *possibility* that Columbus was the first to discover America, or that pledging allegiance to the flag is a good thing, or that science grows by induction is easily answered — people told me these things in a loud, clear voice. All the rest is the story of (negative) appraisal. But can we explain how Jones came to entertain *p* if Jones is the first to do so? Can one explain the origins of novel ideas?

Before answering this question it will be useful to distinguish three different kinds of historical novelty and briefly discuss how each might be given explanatory treatment. We will then ask which kinds of novelty occur in the domain of ideas.

What I shall call 'combinatorial novelty' is the most familiar case. An object is *combinatorially novel* at time *t* if it never existed before *t* but can be brought into being by combining old entities (i.e. ones which existed before *t*), and if all of its properties can be deduced from laws describing the old entities and their relationships. An example might be the new transuranium element, Einsteinium. (I am assuming that *Es* is not produced naturally in the stars, but was first made in accelerators.) Probably each individual animal which is not an identical twin is combinatorially novel. There are no fundamental problems in explaining instances of combinatorial novelty. One simply describes the process by which the old components came to be juxtaposed and derives the properties of the new complex from the laws governing the components.





Thus if a new idea consists of a new arrangement of familiar elements we may hope to explain it as the result of the juxtaposition of a pre-selected set of conceptual elements followed by the selection of those combinations which look at all promising. I have no idea if this is a correct description of any of our actual mental processes, but it certainly seems to be a possible mechanism.

It is more difficult to characterize the stronger sense of novelty which has figured in biological debates, but I propose the following.<sup>14</sup> An object is *emergent* at time *t* if it never existed before, but can be brought into being by combining old elements, and if at least some of its properties *cannot* be deduced from the laws needed to describe the entities and relationships between them which existed before *t*.

To illustrate this concept I will use a somewhat artificial example. Optical activity was first discovered in the nineteenth century. Suppose that before that time the rotation of light by a solution of 'right-handed' molecules such as tartaric acid had never occurred. In this case dextrorotation would be an historically novel phenomenon. Suppose (although this was not actually the case) that one had a theory of chemicals which completely explained all of the observable properties of those materials which had been actualized before time *t*. (Since one can explain the solubility, color, density, combustibility and a multitude of other properties of a molecule without saying whether it is right-handed or left-handed, this is not an intrinsically implausible state of affairs.) Relative to such a theory, the phenomenon of optical rotation would be emergent. This kind of novelty certainly does not rule out the possibility of explanation. We simply must expand our theory of the structure of molecules to cover the newly realized property of optical activity. The typical defender of biological emergence will argue that although highly organized biological organisms can be synthesized from simpler building blocks and although there are laws which relate the properties of the complex to the properties of the simples, it is not the case that the laws which adequately describe the behavior of simple organisms can be extended in a routine way to cover the behavior of the complex case.

It could also be argued that the aesthetic effect of Bach's Two and Three Part Inventions is emergent. Even though an invention consists of a rather simple combination of elements which can be thoroughly understood in isolation (the themes might even be based on folk tunes), the aesthetic experience of hearing the theme played against itself in counterpoint seems to be one of emergent novelty. Nevertheless, it does not seem

impossible in principle to draw up some rules for the construction of good counterpoint. Certainly patterns in Bach's work have been detected.

Much of the novelty in the domain of ideas would seem to be of the emergent variety. Agassi (1968) describes philosophical advances in which solutions and problems are combined in ways that lead to surprising results. And Koestler's theory of bissoassociation (1964) seems to be an attempt to describe the process by which new creative ideas emerge from the juxtaposition of familiar frameworks.

There is a third, stronger notion of novelty which has generally not been defended with much philosophical enthusiasm. To use Aristotle's terminology, this is the case where 'something comes-to-be in the unqualified sense' (*De gen. et corr.*, 317b), i.e. the case where there is no way even in principle of producing the new entity using the old ingredients alone. Perhaps we should call this *spontaneous substantial novelty*. (The 'substance' is new and it comes into existence in a spontaneous fashion.) It would appear that the generation of objects which are novel in this sense would be beyond the reach of scientific explanation.

Are new scientific theories ever novel in this third, strongest, sense? I suppose some might wish to argue for spontaneous substantial novelty even in the case of a new theory which is formulated in an already familiar language. But the case for radical novelty looks strongest of the new theory is formulated in a new language, i.e. employs new concepts.

But are concepts ever novel in this very strong sense? To a large extent, it was their failure to reduce theoretical terms to observational ones which led to the breakdown of the logical positivists' program. The Chomsky reductionist program of analyzing concepts in terms of semantic markers, some of which may be quite abstract in character, seems to be more promising. Still language did not always exist, so it must have emerged in some way. And the history of science describes many incidents in which people acted as if concepts such as force, field, and valency were genuinely novel and incomprehensible, not just unacceptable or unnecessary.

Although Popper does not discuss exactly this issue (he tends to avoid questions of meaning) it seems likely that he believes that there is spontaneous substantial novelty in the domain of ideas. The number of qualitatively different primary entities in world 3 increases over time as they are created by people who are attempting to solve problems. He stresses that his world 3 differs from the domain of eternal Platonic forms just because the inhabitants emerge from the historical process. And in his attacks on historicism he has argued that today one cannot predict the

content of the knowledge which will be available tomorrow. Here we should not confuse the issues of whether tomorrow's *language* is unpredictable and whether the set of propositions *accepted* tomorrow, but written in the same old language, is unpredictable. As it stands, the phrase 'new knowledge' is ambiguous, but combined with his rejection of the possibility of a logic of discovery, this seems to be another indication that he thinks ideas are novel in a strong sense of the word.

If ideas are sometimes cases of spontaneous substantial novelty and if it is impossible in principle to give a set of sufficient conditions for their occurrence, what happens to our program for explaining human activity? It certainly means that we cannot have a complete theory of the origins of ideas. But this in turn means that it is impossible in principle either to know in advance or to control the theoretical content of people's situations, although we may explain them after the fact. And since ideas affect physical behavior, this means that spontaneity in world 3 may affect world 1. Thus if Popper is correct, we see that any laws of physical entities which ignore their interactions with minds and/or ideas are fundamentally incomplete. The openness of the world of ideas 'infects' (or 'liberates', depending on your point of view) the physical world.

Even if there is spontaneous generation in world 3 we need have no fears of a population explosion. As any historian who has played the precursor game will testify, the amount of genuine novelty in the world of ideas is quite limited and the size of the typical quantum of creativity is very small. New ideas have a way of bearing strong family resemblances or analogies to old ones. Historians deal with the problem of explaining the origin of new ideas just by analyzing what at first appear to be giant creative leaps into a series of small innovations.

Although we cannot explain the origins of a radically novel idea in detail, we may hope to be able to specify situations in which it is likely that a novel event of a general type would occur. The case of dolphins which can be conditioned to invent novel tricks is instructive. One can predict that after such training they will perform some new stunt when the bell rings; one can also put physical and 'mental' limitations on the type of new trick which is possible for dolphins but one cannot anticipate the exact form of the trick. (I do not mean to suggest in this example that dolphins' tricks are examples of substantial spontaneous novelty.)

Let us now turn briefly to the problem of explaining a person's appraisal of an idea once it occurs to him. One proceeds in a fashion roughly parallel to the method of explaining actions described above. One de-

scribes all of the components of Jones's situation which she considers relevant to his epistemological appraisal of *p*. These will typically include the reports of experiences and experiments available to her, other theories which Jones holds, etc. An account must also be given of what sort of appraisal procedure Jones is using. For most people, appraising a proposition as a basis for action is quite different from appraising it as a suitable item for cocktail conversation or as a candidate for scientific testing. Our description of the epistemological situation will include an account of how critical or credulous Jones is, just as in explaining actions we need to describe the extent to which Jones likes gambling, whether she is a maximizer, minimaxer, satisficer, or what. There is no reason to think that all people weigh things the same way in all situations.<sup>15</sup>

One then shows that given Jones's epistemological problem-situation and her appraisal rule, the outcome to be expected is *A*. By adding the *RP* one can then explain the fact that Jones actually did give proposition *p* appraisal *A*. Popper's account of Galileo's attitude towards his theory of the tides provides an interesting example of this kind of explanation (1972, pp. 170-6). Historians of science often make informal use of this approach.

The complete explanation of beliefs or other propositional attitudes requires two steps. First one gives an account of the origins of the idea. In the case of novel ideas there are serious limitations on the feasibility of providing an explanatory account. Then one describes how the idea was appraised.

## V. The Limits of Rationality

We have seen that novel ideas cannot be explained in terms of the *RP* approach. I want to explore another possible limitation — the problem of irrational fears and desires.

Suppose we are asked to explain the odd behavior of someone suffering from claustrophobia. To explain his action we cite his belief that being in a closed area will cause him great distress. We may then go on to explain his belief by reporting on various instances in which he truly did feel great distress in such a situation. Perhaps he finds it hard to breathe, his heartbeat becomes irregular, etc. Yet both we and the victim agree that it is not the closed space *per se* which causes the violent reaction. It is his fear of it. Should he be moved into a closet while he was asleep, nothing

would happen. Thus the explanations given so far leave the really puzzling question untouched: Why does the agent have this inappropriate or irrational fear?

But can a fear be said to be irrational except in a metaphorical sense? Should the *RP* have anything at all to say about emotional states? We are certainly now approaching the border of the *RP*'s domain. However, I think that we have reason to believe that sometimes emotional reactions, such as fears, can be changed through pure argument. Both we and the claustrophobe are puzzled at his reaction because, even though we rehearse together the good reasons why his fear is disproportionate, in this case the fear does not disappear. As Popper has pointed out (1967), psychoanalysis presupposes that emotions are affected by theories. Once these theories are made conscious and criticized, inappropriate feelings are supposed to disappear. Behavior modification techniques, on the other hand, assume that the most direct way to change emotional responses is through conditioning — no intellectualization is necessary.

I tend to think that an adequate theory of emotional states will probably have to include both rational and arational elements. To overcome one's natural fear of falling down the ski hill and the accompanying tendency to lean into the slope, perhaps knowledge of the physics of edging and the observation of other people's success when they lean out is not enough (although it is certainly helpful). Perhaps the fear does not really go away until one has directly experienced the fact that leaning out on the hill is not dangerous.

If actions and beliefs are influenced by emotions and if emotions are at least in part immune to rational argument and criticism, this places another boundary on the application of the *RP* to the problem of understanding human behavior.

## VI. The Metaphysical Research Program

Although Popper is best known (and rightly so) for his emphasis on the role of bold empirical conjectures and severe experimental testing in the growth of science, throughout his career he has recognized that bold, but unfalsifiable, metaphysical theories have also been important in the history of science.<sup>16</sup> Atomism, for example, obviously had a profound influence on the growth of science, although according to Popper it only became testable in 1905.<sup>17</sup>

But if scientific theories always 'say much more than we can test'

(1963, p. 266), how are we to evaluate their metaphysical components? This problem deserves more careful attention from philosophers of science,<sup>18</sup> but I find three major criteria in the writings of Popperians.

First, as Popper explains in 'On the Status of Science and of Metaphysics' (1963), we can criticize a metaphysical theory for not providing an adequate answer to an interesting problem. Secondly, if a metaphysical theory conflicts with a highly corroborated scientific theory, that may provide good reason for abandoning it.<sup>19</sup> For example, Cartesian metaphysics was in a sense refuted by the success of Newton's theory of force. Thirdly, we can evaluate a metaphysical theory in terms of its heuristic power and fertility for science.<sup>20</sup>

I have argued that the program of explaining human actions and beliefs in terms of the *RP* does have testable elements and I have indicated how the empirical content of the theory could be increased through the construction of auxiliary theories of error, perception, decision making, etc. But much of the present content of the *RP* is clearly metaphysical. According to the theory of the nature of man which lies behind situational explanations, man's actions are controlled by his theory of the situation and his decision procedure. Man's beliefs are controlled by the ideas and information available to him and by his epistemological appraisal procedure. It is because actions and theories are *controlled* that we may hope to understand them — we may even be able to discover the method in madness. (Voltaire said that madness was to have erroneous perceptions and to reason correctly from them. But in some cases one may also need to look for erroneous patterns of reasoning.) It is because the theories and appraisal procedures are *open* to correction and improvement that man can become rational in the strong, prescriptive sense of the word.

Thus, in addition to setting forth a model for the explanation of human activities, Popper has provided systematic methodological and heuristic advice. He has laid out a research program in the Lakatosian sense (1970). The fundamental methodological maxim for his research program (what Lakatos would call the *negative heuristic* which protects the *hard core*) might be formulated as follows: Try to explain all actions and beliefs in terms of situational analysis and the Rationality Principle. If a given action or belief appears to be irrational always blame your model of the agent's situation, *not* the Rationality Principle.

Here the terms 'rational' and 'irrational' are being used in the restricted sense discussed above. We only assume there is a match between the rules used to make decisions and appraise ideas and the outcomes of

these procedures. But I think the stronger, laudatory sense of 'rational' also plays a role in Popper's research program. It provides what Lakatos calls the *positive heuristic*. When we are trying to reconstruct the process by which an agent reached a decision or appraised a theory, it is good strategy to assume that the agent's rationality is greater than the minimal sense implied by the *RP*.

For example, if there is a flaw in a mathematical proof, one is well-advised to look for a hidden lemma instead of conjecturing that the mathematician had made a mistake – or abandoned logic. It is not simply a question of which assumption is more likely to be correct; we are also guided by a consideration of what would be a more interesting conjecture if true.<sup>21</sup> Of course, either assumption about the mathematician's procedure must be submitted to test.

I think this heuristic policy is suggested by Popper's Principle of Transference, which he formulates as follows: "... what is true in logic is true in psychology" (1972, p. 6); "... what holds in logic must hold in genetics or in psychology..." (1972, p. 68). (Popper's own position regarding the status of this principle is not clear. When he first introduces it, he calls it an 'admittedly... somewhat daring conjecture in the psychology of cognition or of thought processes' [1972, p. 6]. Later in the same essay, he calls it a 'heuristic principle' [p. 24]. In a footnote elsewhere he remarks, in passing, that it is a 'fact' [p. 68].)

If we construe 'logic' to include not only the rules of valid deductive reasoning, but also the principles of good scientific method and wise practical decision making, then the Principle of Transference advises us to seek for the instantiation of these normative procedures in the actual behavior of human beings.

There is an optimistic appraisal of man implicit in Popper's philosophy of social science. But it lies in his heuristics, not in his assertive descriptive claims.

I have articulated and analyzed Popper's research program for the social sciences. The core of the program, the model for situational explanations, bears no formal differences from explanations in the natural sciences. The heuristic potential of this approach is provided by a meta-physical theory of man as a rational problem-solving animal. This again is no different from the situation in the natural sciences where theories about the nature of matter or life guide research. Popper's theory of man is certainly metaphysically attractive, but how fruitful is it for guiding

scientific research? Here the verdict is not yet in. Situational analysis has had considerable success in the areas of economics, history, anthropology, cognitive psychology, and in the study of mental illness (although in the latter case the biochemical approach also looks promising). Its final evaluation will depend on the outcome of on-going scientific research. Perhaps it will go the way of Aristotle's doctrine of final causes – a plausible, appealing theory, but in the end it did not make for good physics. However, we also remember that the more austere metaphysics of the Mechanical Philosophers did not succeed either – the idea of force which they thought to be occult is now central to our understanding of the world.

#### NOTES

- 1 In preparing this paper I have benefited greatly from conversations with Mr. Larry DeWitt. I also received many helpful comments from members of the Philosophy Seminar at the London School of Economics where I read a much earlier version.
- 2 Upon further analysis even this example might turn out to be a case of appropriate behavior. Is the person truly convinced that it is physically impossible to get the car in? In my experience, if you try hard enough, a VW will fit in an amazingly small gap!
- 3 This schema may be compared with that given by Hempel (1965, p. 471).
- 4 This is particularly noticeable throughout (1967) where he repeatedly speaks of 'the principle of adequacy'. In *Objective Knowledge* he says the following: "... we can try, conjecturally, to give an idealized reconstruction of the *problem situation* in which the agent found himself, and to that extent make the action "understandable" (or "rationally understandable"), that is to say, *adequate to his situation as he saw it*. This method of situational analysis may be described as an application of the *rationality principle*' (1972, p. 179. Italics in the original).
- 5 Feysabend has claimed that arguments often owe their efficacy to repetition, not semantic content, and there is some truth in this. There are also borderline cases in which the reasons for an action are unconscious and hence are not readily accessible to criticism.
- 6 Although I find the application of situational analysis to the problem of understanding human actions a fruitful one, I am not at all convinced of its value in evolutionary biology. I have no quarrel with the manner in which Lorenz ascribes purposive behavior to animals, because he gives independent behavioral criteria for his ascriptions. (Perhaps when we see a dog's ears laid back we know he wants to fight.) But it is not clear to me what explanatory power is gained by speaking of a species' solution to a survival problem.
- 7 Watkins noted that the *RP* includes two separate claims which he formulated as follows:

... to every practical conclusion drawn from a decision-scheme there corresponds, physical circumstances permitting, an appropriate action; and behind every action there is an appropriate practical conclusion drawn from a decision scheme. (1970, p. 209)

He concluded that both were unfalsifiable because of their 'all-some' logical form. However, I believe that enough restrictions can be placed on the class of admissible problem-situations to render both clauses testable.

8 For references and a discussion of Shackle's work, see Watkins (1970).

9 Thus according to my schema as it stands it would be possible in principle to subsume under the *RP* the behavior of a person who minimized expected utility in a consistent, methodical way. Of course, we would certainly want to go on to explain why such a decision policy ever came to be adopted – and how the agent has survived!

10 Perhaps this also explains one of Popper's rather puzzling early accounts of how historians give explanations by describing the logic of the situation. In *The Poverty of Historicism* he quotes the following explanation with approval (I have added the italics.) 'Tolstoy, for example... describes how it was *not* decision but "necessity" which made the Russian army yield Moscow without a fight and withdraw to places where it could find food' (1961, p. 149).

11 As Dray (who does *not* think the *RP* is a law) puts it, 'It does not follow from the fact that an agent has "compelling" reasons to act, and knows it, that he will in fact do what they require...' (1964, p. 14). Hempel points out that 'grounds for believing it would have been rational for *A* to do *x* [do not constitute] grounds for believing that *A* did in fact do *x*' (1965, p. 471). And Watkins argues that it is logically possible for an agent to deliberate and reach a conclusion, but then not act on it because he had changed his mind for no reason (1970, p. 173).

12 I am not here concerned with the problems encountered by Von Mises and others in specifying the exact restrictions on admissible *Q*'s.

13 The reader may be interested to know that both of the examples describe real-life cases. The girl was an Estonian refugee who lived on badly distributed Red Cross supplies after World War II. For a time they lived on lard and catsup and she grew to like the combination. The story that hot water freezes faster than cold goes back at least to the time of Bacon. Recently scientists have found that in some special cases, hot water *will* freeze faster than an equal quantity of cold simply because the hot water evaporates and so there is less to freeze.

14 Cf. the definition given by Mehl and Sellars (1956).

15 My forthcoming paper, 'On Explaining Beliefs', addresses the question of how desires affect beliefs.

16 See, for example, the preface and Section 85 of (1959) and the various comments on atomism (1963). Agassi sees metaphysics as playing an even more central role in the growth of science than does Popper. See his essay, 'The Nature of Scientific Problems and Their Roots in Metaphysics' (1964). It should be noted that Popper's demarcation principle loses some of its importance once we recognize the large role of non-testable, hence 'non-scientific', elements in science.

17 See Magee (1971).

18 To mention one problem – it is often said that a good metaphysical theory should provide a framework which will accommodate a wide variety of future theories, cp. Agassi (1964), p. 207. Yet it would seem that if a metaphysical

theory is to have much heuristic power it should rule out a wide variety of possible scientific theories. On the face of it, these two desiderata are in conflict. 19 See Wisdom (1963).

20 See Agassi (1964) and Lakatos (1970).

21 Studies show that many examples of subjects' reasoning which had been thought to be fallacious were really enthymemes – often the implicit premises were ones the experimenters had not thought of. For the procedure of looking for hidden lemmas, see Lakatos (1963-64). Since rationality is one of our highest values, this policy of always giving the subject whose activity is to be explained the benefit of the doubt, as it were, also has desirable moral overtones.

#### BIBLIOGRAPHY

- Agassi, J., 'The Nature of Scientific Problems and Their Roots in Metaphysics', in M. Bunge (Ed.), *The Critical Approach to Science and Philosophy – In Honor of Karl R. Popper*, Free Press, Glencoe 1964.
- Agassi, J., 'The Novelty of Popper's Philosophy of Science', *International Philosophical Quarterly*, Vol. 8 (1968), pp. 442-63.
- Dray, W. H., *Philosophy of History*, Prentice-Hall, New Jersey 1964.
- Hempel, C. G., *Aspects of Scientific Explanation*, Free Press, New York 1965.
- Koertge, N., 'On Popper's Philosophy of Social Science', in K. F. Schaffner and R. S. Cohen (Eds.), *PSA 1972*, Vol. XX: Boston Studies in the Philosophy of Science, Reidel, Dordrecht 1974, pp. 195-207.
- Koertge, N., 'On Explaining Beliefs' (forthcoming in *Critical*).
- Koestler, A., *The Act of Creation*, Hutchinson, London 1964.
- Kohlberg, L., 'Stages of Moral Development as a Basis for Moral Education', in C. M. Beck, B. S. Crittenden and E. V. Sullivan (Eds.), *Moral Education*, University of Toronto Press, Toronto 1971, pp. 23-92.
- Lakatos, I., 'Proofs and Refutations', *British Journal for the Philosophy of Science*, Vol. 14 (1963-64), pp. 1-25, 120-39, 221-43, 296-342.
- Lakatos, I., 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos and A. Musgrave (Eds.), *Criticism and the Growth of Knowledge*, Cambridge University Press, London 1970, pp. 91-196.
- Magee, B., 'Conversations with Karl Popper', in *Modern British Philosophy*, Secker and Warburg, London 1971.
- Mehl, P. and Sellars, W., 'The Concept of Emergence', in H. Feigl and M. Scriven (Eds.), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, Vol. I, Minnesota Studies in the Philosophy of Science, University of Minnesota Press, 1956, pp. 239-52.
- Popper, K. R., *The Logic of Scientific Discovery* (English ed.), Hutchinson, London 1959.
- Popper, K. R., *The Poverty of Historicism* (3rd ed.), Routledge & Kegan Paul, London 1961.
- Popper, K. R., *The Open Society and Its Enemies* (4th ed.), Routledge & Kegan Paul, London 1962.
- Popper, K. R., *Conjectures and Refutations*, Routledge & Kegan Paul, London 1963.
- Popper, K. R., 'La Rationalité et le Statut du Principe de Rationalité', in E. M. Classen (Ed.), *Les Fondements Philosophiques des Systèmes Économiques*, Payot, Paris 1967, pp. 142-50.
- Popper, K. R., *Objective Knowledge*, Oxford University Press, London 1972.

- Popper, K. R., 'Replies to My Critics', in P. A. Schilpp (Ed.), *The Philosophy of Karl Popper*, Vol. XIV, *The Library of Living Philosophers*, Open Court, La Salle 1974.
- Salmon, W. C. (Ed.), *Statistical Explanation and Statistical Relevance*, University of Pittsburgh Press, Pittsburgh 1971.
- Watkins, J. W. N., 'Imperfect Rationality', in R. Borger and F. Cioffi (Eds.), *Explanation in the Behavioral Sciences*, Cambridge University Press, London 1970, pp. 167-217.
- Wisdom, J. O., 'The Refutability of "Irreducible Laws"', *British Journal for the Philosophy of Science*, Vol. 13 (1963), pp. 303-6.

## Discussion:

# Danto's Rejection of Immanent Causation

Thomas E. Wren

Loyola University of Chicago

Against Danto's recent argument that the causation internal to basic actions is not a special, immanent causation, it is objected that (i) he introduces a notion of truncated action that involves a fallacious use of the Equals-subtracted-from-equals axiom, (ii) his version of the Identity Thesis turns upon a misleading notion of co-referentiality, and (iii) he falls into what, by his own theory of meaning, amounts to a category mistake concerning intentions as causes within actions. Hence Danto's arguments do not warrant his materialist claim that causation is a univocal concept.

## 1. The Question

For well over a decade Arthur Danto has been proposing, defending, and revising his now well-known theory of basic actions. In a recent book, which he says 'for worse or better, supersedes whatever I have written on this subject',<sup>1</sup> he continues to argue that there are some actions which, although parts of larger action-compounds, are not themselves reducible to smaller, component actions. But to make this familiar point all the more convincing, as well as to limn it in greater detail than before, Danto now treats, with none of his earlier qualifications, basic actions as instances of 'transseunt' causation. This harder line apparently came as some surprise to Danto himself,<sup>2</sup> but while his openness to the unforeseen is certainly admirable, his argumentation fails to support his important claim that causation is a univocal concept, applicable only as a transeunt ligature between two events.

This is not to challenge his portrayal of *compound* (non-basic) actions as transeunt causal liaisons. In an article published in these pages shortly before his book on action appeared, he nicely refined his account of these liaisons.<sup>3</sup> It is now quite clear that he is not, e.g., saying that a basic action such as lifting one's arm - represented as *mDb* (a man *m* does *b*), with *b* representing the arm-rise itself - causes a non-basic action such as moving a stone - represented as *mDa*, with *a* representing the stone's rolling out of its socket. Rather, the arm-event *b* causes - transeuntly - the stone-event *a*. Similarly, a prior event such as the falling of a stone (*f*) could, when properly perceived by an agent, cause - again, transeuntly - his lifting his arm: i.e. *f* could cause the entire basic action complex *mDb* (but not *b* taken in isolation). And so on. These important examples show transeunt causation as it precedes or