METAPHYSICS, PRAGMATIC TRUTH AND THE UNDETERMINATION OF THEORIES

ALFREDO PEREIRA JR. AND STEVEN FRENCH

1. Introduction

The empirical underdetermination of theories by the data has recently attracted a great deal of critical discussion, to the point where it has come to occupy a pivotal position in the realist/empiricist debate as a whole. Our intention in this paper is twofold: firstly, to map out some of the more important lines of argument in this discussion and, secondly, to consider two possible realist responses, one involving the notion of "pragmatic truth" and the other the introduction of a priori or metaphysical beliefs.

The first philosophical interpretations of the problem go back to Duhem and Poincaré (see, for example, Worrall 1982, pp. 214ff.) and the use of the possibility of underdetermination to argue for the conventionalist position with regard to the philosophy of space and time was, famously, continued in the work of Reichenbach (1958). More recently the discussion has reached prominence once again through the analyses of Boyd (1973), Quine (1975) and Newton-Smith (1978, 1980) and has been placed by Ellis (1985) for example, at the centre of the debate between realism and empiricism, the latter being recently represented by van Fraassen's 'constructive empiricism' (van Fraassen 1980, 1985).

The apparent complexity (and also the longevity) of this discussion essentially derives from the dependence of the meaning of the term 'underdetermination' on the particular philosophy of science that is adopted. Thus in Reichenbach, for example, we find a characterization of the notion in logical empiricist terms, whereas Boyd's criticism springs from his 'holistic' conception of scientific theories. Quine's reworking of the problem was then intended to demonstrate that even for such theoretical 'holism' the possibility of underdetermination cannot be dismissed.

Even more recently this possibility has been given new guise in terms of the semantic or model-theoretic approach (van Fraassen 1980). The introduction of 'partial structures' into this approach (da Costa and French, forth-

Diálogos, **56** (1990) pp. 37-67.

coming *a*) then raises the possibility of a resolution of the problem in terms of a recently elaborated notion of 'pragmatic truth' (Mikenberg, da Costa and Chuaqui 1986), as we shall see below.

A further interesting aspect of this debate is the possibility of reintroducing metaphysical considerations into the philosophy of science (Watkins 1975). Such considerations may be regarded as forming a categorial metaphysical framework in which the theory concerned is embedded. Appeal to certain aspects of this framework may then allow us to resolve cases of so-called 'Quinean' underdetermination as we shall see below.

We begin, then, by considering the arguments of Boyd, Quine, Newton-Smith and van Fraassen, respectively, before examining the two possible resolutions of the problem briefly mentioned above.

2. *Boyd's argument against the possibility of underdetermination*

The holistic critique of the observational-theoretic distinction in scientific theories led to the assertion that the observational evidence in favour of a particular theory constitutes evidence for the *whole* of the theory, including its so-called 'non-observable' terms (that is, terms which refer to unobservable entities or to the causal relations between such entities). Against this realistic view, defended by Boyd (1973), the empiricists have raised the spectre of the underdetermination thesis: given any theory containing non-observable terms it is always possible to produce other theories which have exactly the same observational consequences but which are logically incompatible with the first (this logical incompatibility holding, obviously, between the non-observable terms). This thesis then supports the following principle (Boyd 1973, p. 2): "If two theories have exactly the same deductive observational consequences then any experimental evidence for or against one of them is evidence of the same force for or against the other"; thus there can be no experimental evidence in favour of a *unique* and consistent group of non-observational terms.

Boyd's argument against the underdetermination thesis attacks the empiricist principle above and not, directly, the thesis itself. This means he must admit that there *could* exist transitory cases of underdetermination in the history of science but that such cases are not representative of anything innate or 'congenital' which affects all scientific theories. His strategy is to show that we can, in fact, have *indirect* evidence in favour of one theory and against its competitors.

Thus he begins by noting that in order to deduce observational consequences from a theory we need, as well as the theory itself, a certain set of

auxiliary hypotheses AH. Three kinds of auxiliary hypotheses are identified and Boyd claims that for all three the empricist argument fails (Boyd 1973, p. 4):

(a) If AH is the set of all possible theories then, if theories T and T′ have, in conjunction with AH, the same observational consequences, T and T′ "must be exactly the same theory (their deductive closures must be identical)".

(b) If AH is the set of theories presently accepted, then a new theory AH′ might be such that from T ∧ AH′ certain observational consequences can be deduced which are not deducible form T′ ∧ AH′. According to Boyd, Reichenbach's argument can be subsumed under this case (1973, p. 5).

(c) If AH is the set of theories which may come to be accepted then we simply cannot know beforehand that T and T′ will produce the same testable consequences when combined with AH.

Boyd fails to take note of the fact that, in the first case above, if T and T′ were the same theory, this theory would be inconsistent, althought T′ and T are not, separately, necessarily inconsistent. It would perhaps be better to follow Horwich's line on this point (Horwich 1982, p. 65) and consider contradictory theories as incompatible and thus, a fortiori, as different theories. Secondly, it is worth pointing out that, with regard to case (b) and Reichenbach's argument as applied to conventionalism and spacetime theories, the latter's 'universal forces' would have to link up in some way with the set of auxiliary hypotheses AH′ in order for different consequences to be deduced. However, this is exactly what universal forces do not, and can never, do; they are in principle "beyond our reach" (we are grateful to Michel Ghins for this point).

Boyd then goes on to admit that an empiricist may defend a weakened version of the above principle, claiming that, on the basis of the set of theories currently accepted, no possible evidence which does not result in these theories being modified in any way, could favor one of the underdetermined theories, T or T′, over the other. The argument which he raises against such an attempt is based on the principle that new theories should be fundamentally similar to their predecessors with regard to the causal relations between theoretical entities (1973, p. 8). Given this, no possible evidence could ever come to favor, within a group of underdetermined theories, one or more which contain causal relations different from those proposed by accepted theories. Thus, those theories T which postulate non-observational terms similar to those postulated by accepted theories are effectively favored, in advance, over those theories T′ which postulate other such terms.

His argument in general is therefore an argument against the existence of underdetermination in an irreducible and definitive form and is based on the

assertion that the causal relations postulated by theories allow of a certain degree of inter-theoretic judgment, that which produces "collateral information" (Boyd 1973, p. 9), which in turn functions as indirect evidence for or against a given theory. In consequence of this, we can classify underdetermined theories according to their degree of plausibility and eliminate the less, in favour of the more, plausible.

3. Quine's reformulation of the problem

In his well-known work of 1975 Quine (1975) effectively reformulated the problem in terms of the underdetermination of integrated theoretical "blocks" or "world systems", with regard to the totality of observable events. His aim was to demonstrate the possibility of underdetermination between world systems, for which an analysis of the relations between those particular theories which are components of each system would be without value. This paper also represents an attempt to clarify the fundamental notions of "observable", "relation between theory and evidence" and "theory" itself, on which a full explication of the concept of underdetermination depends.

We generally understand by "empirical equivalence" equivalence at the observational level, but what constitute the criteria of observability is of course problematic. Quine's proposal is that observable expressions be those that can be learned ostensively (1975, p. 316). The relation between theory and evidence is thus considered to be a relation between sentences in a given language: on the one side we have the theory, generally considered in terms of a deductively closed set of generating axioms, and on the other we have the observation sentences, regarded as existentially quantified Hempelian type observation reports.

The relation between the two is thought of by Quine —as by Boyd, in the context of the hypothetico-deductive approach— as being that of a relation of deduction (or not) of the latter from the former. However, as well as taking account of auxiliary hypotheses —which are contained within the "world system"— Quine also needs two further conditions in order that an observation sentence be deducible from a theory:

(1) That the observation sentences be specified in terms of a system of spatio-temporal coordinates. Each observation sentence expressible in the language can be associated with each combination of coordinates, generating a series of "pegged observational sentences" (Quine 1975, pp. 316-317) which are true or false according as to whether that which supports them occurs or not, respectively, within the region of space-time with which they are associated.

(2) That the pegged observational sentences (POS) be conditionalized by others already verified (VPOS) which describe the boundary conditions of the situation under consideration (Quine 1975, p. 317). Thus what theories really deduce are "observation conditionals" of the form VPOS → POS. In other words, we can say that from a theory (world system), in conjunction with the VPOS's, we can deduce certain POS's which refer to events which may not have been observed. We use the theories to establish the true values of the POS's from that of the VPOS's.

A "theory formulation" is basically a sentence composed of the conjunction of all the axioms of a theory (if necessary this concept can be put in a more "liberal" form; Quine 1975, p. 321). Quine holds that two theory formulations which are empirically equivalent and which may become logically equivalent through some suitable rearrangement of their predicates, are expressions of one and the same theory. This then allows him to define "theory" in the following terms: the theory expressed by a particular theory formulation is the class of all theory formulations equivalent to it and which may be transformed into its logical equivalents by means of a reconstrual of the predicates (Ibid.).

The Quinean version of the underdetermination thesis then claims that, for whatever theory formulation, it is always possible to generate (at least) another with which it is empirically equivalent but logically incompatible and to which it cannot be made logically equivalent by a suitable reconstrual of the predicates (p. 322). The reason for this is that there does not exist a one-one relation between theoretical terms and empirical categories.

This thesis does *not* apply to theories which may be reduced to the conjunction of observation conditionals which they imply—not even those which encompass an infinite number of observation conditionals in universally quantified form. However, the underdetermination thesis *does* apply to those theory formulations which contain irreducible non-observational terms. Quine holds that we frequently need to resort to precisely such formulations because many POS's refer to inaccessible spatio-temporal situations and they represent the only way we have to determine (*via* deducibility), together with a finite set of axioms, the infinite set of observation conditionals which we consider to be valid. In order that finite formulations can determine an infinite set of POS's, it is necessary to introduce non-observable terms which effectively "loosen" the former (Quine 1975, p. 324). An infinite conjunction of observation conditionals constitutes a "tight" theory, for which the problem of underdetermination does not exist. However, our theories (world-systems) in general cannot be reduced to a conjunction of this kind, unless we employ certain

reformulatory techniques such as Craig's, which run into well-known difficulties.

Quine thus defends the view that, because of these characteristics of our theories, it is possible for them to be underdetermined in relation to the totality of spatio-temporal situations. His conception of the problem excludes in advance the existence of new discriminatory evidence, since this is defined in relation to all possible POS's. However, this position generates the following epistemological problem: *how* can we know that T and T′ have the same empirical consequences, if we do not know *all* their empirical consequences? Syntactically it is only possible to determine that each observation conditional is either deduced from both or from neither, if we have available some mechanical decision-making process. However, for complex theories (world systems) it is possible that one may be unable to decide if an observation conditional is implied or not by T or by T′, "world systems" being theories which explain every possible experimental fact. In this case, we cannot affirm, for certain, that T and T′ are empirically equivalent with regard to all possible experience. Consequently, *if* there exist, in the history of science, such "Quinean" cases of underdetermination, we would never be able to ***know*** if they really are genuine examples of this kind of underdetermination or whether they represent merely transitory cases, which may be resolved one way or the other as science progresses.

4. The realist's dilemma and possible ways out of it

In a detailed analysis of the consequences of the above line of reasoning for the realist thesis, Newton-Smith (1978, 1980) introduced an important distinction between *strong* and *weak* forms of underdetermination, the former holding that *all* theories are underdetermined by the data, the latter claiming merely that there *can be* cases of this "phenomenon". Whereas Quine is taken as arguing for the strong thesis, Newton-Smith himself inclines towards the weak form admitting only that cases of underdetermination *may* occur in the history of science. To support this view he gives the following two examples: (1) underdetermination between a theory which claims that "time is linear and history cyclic" and another which holds that "time is closed"; (2) underdetermination between Newtonian mechanics, with its assumption that time is continuous and a rival theory which includes all aspects of the former except the continuity of time, claiming, for example, that time is dense but not continuous. Newton-Smith further argues that these present genuine cases of underdetermination.

In a short reply to these arguments, Quine (1979) denied that he had stated that *all* theories are underdetermined by the data and claimed that

Newton-Smith's first example above, as in the case of his own concerning Poincaré (Quine 1975, p. 322), does not express a true case of underdetermination but rather of two formulations of the same theory.

In a further critique of Newton-Smith, Bergstrom (1984) claimed that we may have theories which are "tight" in Quinean terms and that strict underdetermination may obtain between such theories. Newton-Smith's second example is not, for Bergstrom, a genuine case of underdetermination because both theories contain gratuitous extensions of the "core" of Newtonian mechanics which do not affect their explanatory power. However, Bergstrom himself fails to note Quine's argument for the "loose" character of our theories.

Why does the possibility of such examples present a dilemma for the realist? The answer is deceptively simple: because this possibility together with what Newton-Smith calls the "ontological ingredient" of realism (the Law of the Excluded Middle plus the Correspondence Theory of Truth—Newton-Smith 1980, pp. 230-233) entails the unacceptable conclusion (at least for non-Hegelians) that *reality* is inconsistent (for a defense of exactly this conclusion see Priest 1987, and for criticism and discussion, da Costa and French 1988, forthcoming *b*). To avoid this consequence the realist must adopt one of four strategies:

4.1. To argue against the possibility of underdetermination in general.

4.2. To accept the existence of cases of underdetermination but preserve a form of realism by weakening its "epistemological ingredient" ("the claim that we can have warranted beliefs—at least in principle—concerning the truth-values of our theories"—Newton Smith 1980, p. 230).

4,3. To change the "ontological ingredient' of realism.

4.4. To accept the existence of cases of underdetermination while arguing that these may be resolved by the use of supplementary criteria of theory choice.

Of course, in philosophy as in life, you don't get something for nothing, and adopting one of the above may cause further problems for the realist. In the rest of this section we will consider various arguments presented by realists and antirealists alike, concerning each of the above possibilities.

As regards the first, one could say, firstly, that theories which are apparently underdetermined by the data are only apparently contradictory or, at least, that it cannot be proved that they are not synonymous (Worrall 1982, p. 223) or, further, that "there could be nothing to prevent our attributing incompatibility to equivocation." (Dummett 1973, p. 543). Secondly, one could consider as empirically non-equivalent two theories which attribute different truth values to a set of counterfactual propositions (Newton-Smith

1980, p. 71). However these arguments can be criticized on the following grounds: with regard to the first, the syntax of the theories is passed over, as if it were merely an accidental property of them; concerning the second, apart from treating as real certain possibilities which are merely logical, it involves a procedure which, as well as being very problematic philosophically speaking, only produces a resolution of certain cases of underdetermination, whereas the realist needs some kind of solution for all such cases.

The second possible strategy is the one proposed by Newton-Smith. Weakening the "epistemological ingredient" of realism comes in two forms: the "arrogance responce", preferred by Newton-Smith himself and the "ignorance response", adopted by Bergstrom, both of which involve the assumption of the theories' decidability.

The former consists in claiming that the world is simply indeterminate with respect to underdetermined theories and that "if we cannot know about something, there is nothing to know about." (Newton-Smith 1980, p. 234). This involves a restriction in the use of the Principle of the Excluded Middle (and therefore a change in our logic) as we shall see below. The ignorance response, on the other hand, claim that there exist facts for which we do not have evidence or, more exactly, that "those propositions responsible for underdetermination are either true or false but with regard to these propositions we could not possibly have evidence concerning their truth value. As such, this response involves embracing the possibility of *inaccessible facts.*" (Newton-Smith 1980, p. 233) This is the move favored by Bergstrom, based on two arguments: "In the first place, the supposition that some *specific* inaccessible fact exists may explain, or help explain, the empirical content of the theory which postulates its existence. [...] Secondly, the *general* supposition that inaccessible theoretical facts can explain something [...] even if we cannot know what they are." (Bergstrom 1984, p. 357). However, it is difficult to see how this could be acceptable to the scientific realist, involving, as it does, giving up the position that science provides all the epistemic access that is possible to the objective "world-structure".

As regards 4.3, two kinds of changes are possible: restricting the range of application of the Principle of the Excluded Middle or changing the theory of truth employed (that is, abandoning the correspondence theory, with the inclusion of certain pragmatic elements realistically interpreted).

The first possibility is that proposed by Newton-Smith (1980, pp. 233-235). Not that he advocates abandoning the Principle altogether, rather he believes that one should restrict its application in the case of empirically undecidable propositions: "If the underdetermination of theory by data is a rel-

atively rare phenomenon this will not mean a very extensive restriction." (Ibid., p. 235).

Modifying one's theory of truth is perhaps even more risky for a realist since one does not know up to what point realism can still be maintained without such a theory. Ellis, for example, believes that realism is only able to offer an adequate solution to the problem of underdetermination if a pragmatic or coherence theory is adopted. He himself prefers a variant of the latter in the form of "internal realism". "For the internal realist what is true, if anything, *is* just what is ultimately justifiable... truth is a kind of limit notion of reasonable belief... It is what we should believe if our knowledge were perfected, if it were based on total evidence, was internally coherent and was theoretically integrated in the best possible way." (Ellis 1985, p. 68).

However, some form of pragmatic truth seems to be the natural way for the realist to establish the truth value of propositions referring to unobservables, while the correspondence theory is maintained for that which refers to observable phenomena. In Ellis' words: "scientific realism can be combined with a pragmatic theory of truth: and given such a theory of truth, all of the criteria we use for the evaluation of theories, including the so-called pragmatic ones, can be seen as being relevant to their truth or falsity." (Ellis 1985, p. 41).

We shall return to this point below when we consider the notion of pragmatic truth proposed by Mikenberg *et al.*, and the introduction of "partial structures" into the semantic approach. However, it is worth noting here that any realist view which supports a theory of "approximate" truth is capable of accommodating the underdetermination thesis in its "transitory" form, which says that theories may be underdetermined by all ***known*** data. In such cases the realist can say that the theories are all "approximately true" and that further investigation will reveal which is "more" approximately true than the others. Of course, the standard objection to such a tactic is that no one has yet succeeded in giving a viable formalization of this notion of "approximate truth". It is our belief that Mikenberg's characterization might be a suitable candidate for just such a formalization (see French, forthcoming *a*).

Finally, we come to the fourth strategy above, involving the introduction of supplementary criteria of theory choice. In this case, the existence of underdetermination is accepted (at least in its "transitory" form) and certain criteria are looked for (in some cases through contesting the "Quinean" form) which allow of a rational choice between underdetermined theories.

There exist at least three kinds of argument here, concerning simplicity, the distinction between the set of empirical consequences of a theory and the overall domain of evidence (Ellis 1985), and inductive closure (Glymour

1985). Before considering such arguments, however, it is worth noting at least two *antirealist* criteria of theory choice in underdetermined situations.

The first is simply choice by convention, as presented by Reichenbach for example. As Horwich (1982, p. 61) explains, "it rests upon the idea that particular sentences of a theory are rendered true or false, solely as a consequence of convention; once we have adopted T, we thereby implicitly stipulate the truth of its bridge statements and determine a priori the falsity of the elements of T_2 which conflict with these statements." Since the choice is arbitrary, one obviously cannot (and should not according to the defenders of this view) interpret the truth of a theory in terms of its correspondence with reality. The second criterion, originally presented by Carnap (1956) involves the so-called "pragmatic" approach (cf. van Fraassen 1980) which consists in choosing between two underdetermined theories by means of their heuristic power, without any commitment to a particular interpretation. "For an observer *X*, 'accepting' the postulates of T means [...] not only taking T as an uninterpreted calculus but using T [...] to guide one's predictions of future observable events." (van Fraassen 1980, p. 227). Thus, as in the conventionalist view, those propositions which refer to unobservables do not possess truth values determined by the existence of their referents, contrary to the realist position.

The pro-realist argument that the simplest theory should be preferred is based on the questionable assumption that reality is itself simple, and collapses into ambiguity (there does not seem to exist any agreed formalization of, nor even definite criteria for, such a notion)—motives which explain why this position, although generally popularized, finds few defenders these days.

Ellis, on the other hand, seeks a solution in the distinction between the set of empirical consequences of a theory and that which he calls its "field of evidence". The latter is taken to be larger than the former and is "open", since it may be enlarged through the development, often unexpected, of other theories: "It may be evidence which can be seen as relevant to the theory is question *only* because some new linked theoretical developments might occur." (Ellis 1985, p. 65). However, he admits that this evidence is not a consequence of the theory alone. Thus he falls into the pre-Quinean situation, as discussed by Boyd above, of favouring one theory over another by way of new auxiliary hipotheses. But Quine has demonstrated the possibility of underdetermination for whatever "group" of theories; that is, a theory T together with new theories NT, may differentiate its field of evidence with regard to its rival T′, but then, if we accept Quine's reasoning above, there will be a further theory T″, logically incompatible with and different from, T, which, in conjunction with NT, will cover exactly the same field of evidence as T plus NT.

As regards inductive closure, Glymour has argued that we have reasons, of an inductive nature, for believing more strongly in a theory than in its set of empirical consequences. The argument is essentially based on Hempel's well-known affirmation that a theory establishes inductive connections between pieces of evidence, connections which the class of empirical consequences is not by itself capable of establishing (cf. Glymour 1980, pp. 161-163). Glymour's famous bootstrapping strategy then allows us to say that, given a certain finite set of evidence claims, this set tests a theory better than it does the set of empirical consequences of the latter. The reason underlying this is that the set of empirical consequences does not possess a well-defined logical structure, in contrast with the theory. Subsequently Glymour has put forward the following principle of theory comparison (Glymour 1985, p. 109): "*Ceteris paribus*, it T and Q are theories and for every established pair of regularities H, K, such that Q explains H as a result of K, but there exists established regularities, L, J, such that T explains J as a result of L but Q does not explain J as the result of any other established regularity, T is preferable to Q."

Van Fraassen has contested this line of reasoning on the basis of the argument that since a well-established logic of induction does not exist, one cannot guarantee that there will be confirmation that is both "objective" ("a relation solely between theory and the total body of evidence, independent of the context of evaluation"—van Fraassen 1985, p. 277) and "unique" ("The propositions to be believed on the basis of the evidence are a determinate and logically consistent set [...] and rationality requires that, given total evidence, one believes all and only those propositions."—Ibid.). Therefore, from the fact of a theory being "empirically adequate" one cannot deduce a correspondence between those propositions which refer to unobservables and a unique, objective (unobservable) reality (van Fraassen 1985, pp. 276-280, 295-296). (It is worth noting than van Fraassen's opinion of inductive logic might not be borne out by recent developments; see da Costa and French, 1989).

5. Underdetermination within the semantic approach

With the stagnation and eventual collapse of the old empiricist programme which effectively underpinned the underdetermination argument, many realists came to believe that the problem had lost much of its force. That this is not so has been strikingly emphasized with the reappearance of empiricism in model-theoretic guise. Van Fraassen's powerful espousal of the tenets of this view within the semantic approach to scientific theories (van Fraassen 1980) marks a fundamentally important new

chapter in the development of the realist/empiricist debate as a whole. However, we believe, with Giere (1985) for example, that the introduction of this approach also gives new hope to the realist, in particular with regard to the present context.

The basics of "constructive empiricism" are by now very well-known and we shall only summarize them here. It begins with the fundamental principle that experience is the only legitimate source for our factual opinions and can give us information only about what is both observable and actual (van Fraassen 1980, pp. 3, 12; 1985, p. 253). It is important to note that van Fraassen neither denies the existence of unobservable entities, nor dismisses statements involving them or the question of their existence as meaningless. Rather he adopts a strictly agnostic attitude in arguing that such statements should be neither believed nor disbelieved.

On the basis of the above principle, van Fraassen erects his antirealist position that the aim of science is to give us theories which are empirically adequate in the sense of agreeing with all observable events and nothing else. This then leads him to make a radical break between acceptance of a theory and belief that it is true. For the constructive empiricist, acceptance involves only the belief that the theory is empirically adequate; that is, that it is true at the level of the empirical consequences only. Two further innovations are then introduced which, in the first case, allow him to avoid the internal problems which afflicted the older forms of empiricism, and in the second, provide him with a shield against certain well known realist criticisms of these older positions.

Thus, instead of giving a linguistic formulation of scientific theories, van Fraassen adopts the "semantic" view, which considers a theory in terms of a description of its set of models, regarded as the set of structures it makes available for modelling its domain (van Fraassen 1980, pp. 41-69). Certain parts of these models, the "empirical substructures" (p. 64) are then specified as candidates for the direct representation of observable phenomena. Such structures as can be described in observation reports are called "appearances" and hence a theory is empirically adequate if it has some model such that all appearances are isomorphic to empirical substructures of that model.

The second innovation allows him to counter the criticism that empiricism in general cannot account for the prevalence in science of certain methodological principles, such as simplicity, explanatory power, etc. Van Fraassen's response is to take on board these principles but to regard them as *pragmatic* virtues only (1980, pp. 87-96; 1985, pp. 276-281). Thus, to take the case of explanation, for example, he argues (1983, pp. 276-281) that insofar as its virtues go beyond mere description, they may provide reasons for

accepting a theory but not for believing that it is true. The basis of this argument is that to be more explanatory the theory must be more informative but to contain more information is to have more ways of being false and hence to be, at least, no more likely to be true. Explanation, therefore, is merely a pragmatic virtue as opposed to the empirical ones of adequacy and strength.

Van Fraassen, like the good empiricist that he is, fully accepts the underdetermination thesis. However, he gives it a new, agnostic, twist in arguing that it is not that there is no truth of the matter concerning the difference between empirically equivalent theories, as the older empiricists claimed, but that we simply cannot know what this is, and this is because the best we can do is to construct theories which are empirically adequate. The model-theoretic representation of underdetermination hinges on the view that the theoretical model is regarded as merely a mathematical representation of the observable substructure and only the latter "corresponds" (and here van Fraassen seems to adopt the correspondence theory of truth) to certain features of reality. Thus, it is possible, according to this position, for there to exist many different but empirically equivalent embeddings of an empirical substructure into models of the theory (regarded as equivalent because of their *empirical* indistinguishability).

The importance of underdetermination in this semantic form for the constructive empiricist is that it guarantees that there is always a plurality of models available and that the choice between them cannot be made on the basis of empirical evidence. Such choice, or so the argument runs, can only be made on *pragmatic* grounds.

"Constructive empiricism" has, of course, been much discussed since it first entered the scene (see for example, the collection of essays in Churchland and Hooker 1985). In the present context we wish only to mention the following two points.

The first concerns the argument above regarding explanation, information and belief that the theory is true. Van Fraassen accepts that theories contain more information than the set of their empirical consequences but wants to argue that such information, by increasing the falsifiability of theory, should effect a decrease in our belief that the theory is true. However, if this further information is of an empirical character (although, given the above, it is difficult to see how it could be) then in the context of underdetermination, one theory is simply empirically stronger than the other. If, on the other hand, this information is not empirical, then how can it increase the probability of the theory being false if, according to van Fraassen, that which is not strictly empirical in his sense cannot be assigned a truth value?

The second criticism is more complicated but also goes to the heart of the underdetermination issue in model-theoretic terms (see French 1988). The crucial notion here is that of "embeddability" characterized by van Fraassen thus: a scientific theory T determines a set of mathematical structures B such that the content of the theory can be expressed as the assertion that the physical system A is embeddable in at least one of these structures (van Fraassen 1980, pp. 64-69; 1985, pp. 270-276). Mathematically speaking an embedding is simply an injective function which is a "faithful" homomorphism so that every fact in A is represented in B. In these terms the difference between the realist and the constructive empiricist is that the latter claims that an empirical substructure can be *embedded* into some model of the theory, in precisely this sense that there exists a faithful injective homomorphism from the substructure into the model. It is then possible, as we said above, for there to exist many different such embeddings of the substructure into empirically equivalent models of the theory. For the realist, on the other hand, the only mapping between the empirical substructure S and the model M is the canonic inclusion (which sends each element of the former to itself regarded as a member of the latter), since it is claimed that S is literally a submodel of M. In this case, all elements of the model correspond to physical reality, clearly a rather severe restriction to impose.

In the context of a discussion of the relationist versus absolutist theories of space-time, Mundy (1986*a*) has recently argued that the above function must satisfy a certain "uniqueness' condition (there cannot exist more than one such function, up to an automorphism), otherwise the relationist theory cannot have all the consequences of the absolutist. Translating this into the realist/empiricist context (French 1989*a*) one can argue that given the open acceptance of underdetermination in the model-theoretic form above, the constructive empiricist must clearly reject this uniqueness condition, since it restricts the set of possible embeddings to include only one.

However, Mundy has further argued (1986*b*) that this condition is both necessary and sufficient for what he calls "expressive equivalence" between the absolutist and relationist theories and therefore, by generalisation, between the realist and empiricist interpretations as a whole. By this he means, in our terms, that all the properties of the empirical substructure definable within the model as a whole are also independently definable within the empirical substructure itself. Now this seems to be exactly what van Fraassen wants, since the empirical substructures are, on his view, independent, in a certain sense, of the rest of the theory as far as truth is concerned and their properties refer to the representation of observable phenomena only, having nothing to do with the theoretical terms in the larger model in which

they may be embedded. But then the constructive empiricist is caught on the horns of a dilemma: he must abandon either "expressive equivalence" or empirical underdetermination. Dropping the first means giving up, in some form or other, the general desire for constructive empiricism to "do" all that realism can do but without having to attribute truth to anything but the empirical substructures. Throwing out the possibility of underdetermination seems to be even more difficult to swallow, given its position as a cornerstone of the whole empiricist edifice (for further details of this argument, see French 1988).

Leaving aside these criticisms, it is, of course, also possible to adopt a realist stance within the model-theoretic conception (see for example, Giere 1985 and Friedman 1983; for criticism of the latter's "realism-through-unification" view, see Hiskes 1986 and French 1988). However the underdetermination problem has not been adequately treated from this point of view. In particular, very little attention has been paid to the possibility of constructing a semantic version of the notion of "approximate" truth which, as we noted above, offers the most obvious realist solution to this problem (although Giere, at least, does seem to be groping towards some such characterization; 1985, pp. 79-80). In what follows we shall argue that one, particularly plausible, way of capturing the above notion is through the introduction of "partial structures" into the model-theoretic programme.

The latter idea was first introduced by Mikenberg, Chuaqui and da Costa (1986) in an attempt to formalize the concept of "pragmatic truth". This formalization had a double motivation: first of all, to try to capture in strictly logical terms certain pragmatist views on meaning, etc.; and secondly, to formalize the intuitive view that strictly "refuted" theories are, in certain cases, not entirely abandoned, the most obvious example being Newtonian mechanics, but retain a certain "pragmatic" value. It is not our intention to discuss these motivations here (for an exploration of a possible "Peircean" attitude toward the realist-empiricist debate in general, see French 1989*a*), except to note that a theory is pragmatically true, now and forever, because it was at one time approximately true (thus, these two terms will be taken to be synonymous for our purposes). This formalization has since been incorporated into a general research programme regarding "pragmatic" probability (da Costa 1986), inductive logic (da Costa, forthcoming; da Costa and French 1989) and the introduction of partial structures into the model theoretic approach in general (da Costa and French, forthcoming *a*; French, forthcoming *a*). Without going into all the technical details (see Mikenberg *et al.*, op. cit.) the notion of "pragmatic" or "quasi-" truth can be defined in the following way.

We begin with a given domain of knowledge Δ, modelled by a structure of the form: $\mathcal{M} = \langle A, R_i, P \rangle$ $(i \in I)$, where A is the set of individuals (observable and unobservable) of Δ, P is a set of "privileged" sentences, expressing what is "directly known', about Δ, in the language $\mathcal{L}$ in which talk about $\mathcal{M}$ can be expressed, and (and this is the crucial point) R_i $(i \in I)$, is a family of *partial* relations defined on A, where I is an appropriate index set. The R_i are "partial' in the sense that any relation R of arity n is not necessarily defined for all n_i-tuples of elements of A.

We now let $\mathcal{K}$ be a *total* structure, whose relations of arity n are defined for all n-tuples of elements of its universe and we suppose that $\mathcal{K}$ is also interpreted in $\mathcal{L}$. Then $\mathcal{K}$ is said to be $\mathcal{M}$-normal if (da Costa and Chuaqui, forthcoming):

(i) the universe of $\mathcal{K}$ is A;

(ii) the relations of $\mathcal{K}$ extend the corresponding partial relations of $\mathcal{M}$;

(iii) if c is an individual constant of $\mathcal{L}$ then c is interpreted by the same element in both $\mathcal{M}$ and $\mathcal{K}$;

(iv) if $\alpha \in P$, then $\mathcal{K} \models \alpha$.

We are now able to say that a sentence α of $\mathcal{L}$ is pragmatically true in the partial structure $\mathcal{M}$ according to $\mathcal{K}$, if $\mathcal{M}$ is a partial structure in the above sense, $\mathcal{K}$ is an $\mathcal{M}$-normal structure and α is true in $\mathcal{K}$ according to the usual Tarskian definition.

If T^* is the set of all (total) models of a theory T, then we say that T^* is "pragmatically' or "quasi-" true in $\mathcal{M}$ if and only if some structure of T^* is $\mathcal{M}$-normal. In other words, a theory, characterised by a set of models, is regarded as pragmatically true in a simple pragmatic (or partial) structure when some of its models are $\mathcal{M}$-normal.

This, we would claim, furnishes an adequate characterization of the notion of "approximate" truth useful to the realist and can be used as the basis for the construction of a "pragmatic" realism (French, forthcoming *a*). However, our interest here is simply to point out that this approach effectively resolves the problem of underdetermination (in its "transitory" form at least) by simply accepting the phenomenon! Not in the constructive empiricist sense of permitting many different embeddings of the same empirical substructure into models of the theory but by changing the nature of the latter through the introduction of "partial" structures and thus weakening the naive realist requirement (clearly unacceptable as it stands) that physical reality should be exactly reflected in the model concerned. Partial structures do *not* capture everything in the domain concerned, reflecting the twin facts that we are not omniscient and theories cannot, given the observed frequency of theory-change in the history of science, be taken to be absolutely and per-

manently true. As Dorling (1972) has said, it is (naive) philosophers, rather than actual scientists, who insist on asserting the "literal", that is, absolute, universal and unrestricted, truth of scientific theories.

Thus, strictly contradictory propositions or, more generally, incompatible theories may be simultaneously quasi- or approximately true. This may happen within the same partial structure or in the same domain Δ, either because the propositions are quasi-true in the same partial structure employed to cope with certain problems related to Δ, or because they are quasi-true in two different partial structures which constitute two different kinds of models of Δ. Intuitively speaking, we can easily accommodate underdetermination if we accept that two theories can both be "quasi-" or "approximately' true at any given moment, together with the hope (springing from consideration from the history and actual practice of science itself) that future developments with regard to further empirical confirmation, changes in the domain (its extension, for example), etc., will reveal which one is "closer" to the truth. Of course there is much more that needs to be said here—the nature of the domains modelled, their relationship with the elements of the model, the all-important question of the existence claims with regard to the unobservable elements of A, etc. (see French, forthcoming *a*)—but we hope to have indicated at least, a possible realist way out of the problem under discussion (althought it should be emphasized that it involves a move away from "naive" realism to a more sophisticated form).

6. Underdetermination and Metaphysics

There still remains, of course, the problem of underdetermination in its "Quinean" form, that is, the possibility of theories being underdetermined by *all possible* evidence. In this case the above response is clearly of no avail. Is it possible, then, to elaborate rational grounds for choosing between theories which are underdetermined in this manner? Answering in the affirmative, our intention in this section is to elaborate such grounds in terms of possible philosophical problems associated with the general conceptual schemes in which the theories concerned are embedded, such schemes being regarded as corrigible. This will lead us deep into that nexus of problems surrounding the relationship(s) between logic, metaphysics and science, which, of course, we cannot hope to deal with in its entirety here. However, we will, at least, try to map out those aspects of these problems which bear on the present discussion and point the way to some possible solutions.

To begin with, let us consider some further examples of underdetermination in the Quinean sense.

The first was discovered by Glymour (1977; see also Malament 1977) in the context of general relativistic space-times. Given that in such a space-time all data used in a physical calculation must come from the causal past of the calculator, it is possible to construct non-isomorphic models of General Relativity which are "observationally indistinguishable" in the sense that no observer could ever tell if she were in one or the other, even if she were immortal (for a discussion of this example in an antirealist context, see Torretti 1987, pp. 175-177).

Our second example proceeds from the fact that any theory written in terms of particles can be rewritten in field-theoretic terms and vice versa (Redhead 1983). This is a result first suggested by Isaac Newton (1982) and depends, in the form given by Redhead, on the representation of particles in terms of dichotomic "Yes/No" fields, where the predicate Yes/No associated with a given space-time point indicates that the "particle" is present/not-present, respectively, at that point. Thus it can be argued that although there is an ontological difference between the particle and field-theoretic representations, there is no observable difference in the (strong) sense that no experiment could ever decide which is "correct" (French 1985, pp. 331-336).

In the context of this second example, in particular, our response to the "Quinean" form of the underdetermination thesis is that considerations of particle individuality may help decide in favour of the field-theoretic and against the particle-theoretic view, such considerations forming part of a categorial metaphysical framework which is corrigible. Before we come to expand on this answer, however, we must consider what is meant by a "categorial metaphysical framework" and the role of metaphysics in science (French 1987).

Dismissed by the positivists as at best "narcotic" stimulations to scientific inquiry, metaphysical considerations have been assigned a central role in recent discussions of scientific progress (Buchdahl 1970, 1980; Harman 1982; Hesse 1978; McMullin 1978; Watkins 1975). Watkins in particular has located a "metaphysical component" at the heart of physical theories and has forcefully argued that such components play a fundamentally important role in theory evaluation. The existence of such a role is clearly demonstrated in scientific reductions, for example, which, he claims, can only be understood in terms of a significant change in the metaphysical component concerned. On the basis of such considerations he arrives at the double-headed conclusion that the role of metaphysical speculation is "to prepare the way for the most important scientific advances of all" (Watkins 1975, p. 113) and that the history of science can be written in terms of changes in "world-views".

By "metaphysical component" here, Watkins means the ontological/theoretical component of a theory obtained, or rather, identified, by taking the

Ramsey-equivalent of the theory concerned and then seeing how this "Ramsified" version must be expanded to give us back the original form, "the excess content of the theory over and above its empirical content being regarded as its metaphysical content" (Watkins 1975, p. 96). Underdetermination arises, in these terms, when one has two theories whose Ramsey-sentences are logically equivalent but which are logically inconsistent with one another at the "metaphysical" level (ibid., p. 102).

There are two points which can be made here. The first is that Watkins himself, being concerned with other issues, gives us very little to go on as regards resolving such cases of underdetermination. One possibility would be to adopt Horwich's "global conventionalism". According to this view, and without going into all the details, we may, in such cases of underdetermination, reject a priori all but one of the theories concerned since "our adoption of a whole theory formulation will constrain the referents of its terms in such a way that the alternatives will violate the requirements of our reference-fixing practice and can therefore be rejected a priori" (Horwich 1982, p. 63).

Thus, in the case of the infamous linear vs. closed time example, he shows that if we adopt the latter theory, with its associated "metaphysical component" we may reject the former on a priori grounds to do with how we fix the referents of theoretical terms. However, adoption of such a "whole theory formulation" is purely and simply a matter of *convention* (hence the name) and so this approach fails to answer the question as to *why* we should choose that particular formulation in the first place. The point is that "solving" the problem of underdetermination is not just a question of showing that if we decide on one of the alternatives involved we can automatically reject the others. Some warrant must be given for making that particular choice in the first place.

Returning to the particle vs. field theory example, we are therefore still left with the problem of providing a reason for preferring one representation over the other. To see how a possible answer might be obtained through a consideration of the role of metaphysics in science, we must go back to Watkins' account and a possible ambiguity in his discussion (this is the second of our two points mentioned above).

As we have seen, in explicating the notion of "metaphysical component", Watkins effectively equates it with the theoretical and "non-empirical" elements of a theory, via his considerations of the "Ramsification" process. If this is all he means by metaphysics, then his claims as to its importance in theory change in general amount to nothing more than the realist's position concerning the importance of theoretical terms with regard to scientific progress over all (reduction, unification, etc.). If this is the case, then his discussion obviously becomes a good deal less interesting.

However, earlier in the paper he seems to invoke a different sense of metaphysical when he cites Leibniz's use of the principle of continuity to argue against aspects of both Newton's and Descartes' physics (Watkins 1975, pp. 94-95). According to Watkins, this principle is metaphysical in the sense of being both non-analytic and empirically irrefutable, but it is difficult to see how it could be regarded as part of the ontological or theoretical component of a theory, distilled out by a process of Ramsification. Rather it finds a place within Harman's characterisation of metaphysics in terms of 'the attempt to justify the conceptual rationale of a scientific theory by appeal to regulative maxims such as the law of causality or to criteria of simplicity, analogy or continuity; as well as to attempts to justify the intelligibility of a theory by an explication of the meaning of concepts of matter and force' (Harman 1982, p. 3). In these terms, of course, the problem of identifying, in a consistent manner, the metaphysical component remains, and apart from one or two interesting suggestions Watkins' approach does little to resolve it.

Leaving aside this question, Harman's "regularitive maxims" themselves can obviously be set down within the context of a "categorial metaphysical framework" (Quinton 1973, pp. 235-251), regarded as being composed of the following (Körner 1969, pp. 192-195):

(i) a categorization of the universe in terms of genera of entities;

(ii) a set of constitutive attributes, such that the belonging of an entity to a category logically implies the applicability of the attribute to the entity;

(iii) a principle of individuality, such that the applicability of the principle to an entity of the category implies, and is implied by, the entity's being an individual entity of the category;

(iv) an underlying logic in which the categorial framework is embedded.

As an example, Körner gives Kant's account of the structure of external phenomena, which form the category of entities (Körner 1969, p. 193). Their constitutive attributes are the Kantian Categories, the Principle of Individuality is their position in (classical) space and time and the underlying logic is classical logic. (The example given above of the Principle of Continuity is actually a little more complicated than might first appear, since the Principle is better understood as a logical consequence of principles associated with a categorial framework together with certain other assumptions (Körner 1969, p. 209); Kant, for instance, deduced it from his constitutive Principle of Causality and the assumption of the divisibility of time).

Much of the ire directed against such frameworks and metaphysics in general derives from what is perceived as their dogmatic character (the

"classic" example being the supposed failure of Kantian metaphysics to accommodate the overthrow of the Newtonian world-view). The obvious response, of course, is to insist that the categorial frameworks should be regarded as *corrigible*, in the sense that they may enter into conflict with experiences or external considerations, that this conflict may be resolved by abandoning or modifying this structure and that such a resolution of the conflict may be supported by rational argument (Körner 1969, p. 196-211).

Perhaps the most obvious example of the corrigibility thesis occurs with respect to the set of constitutive principles. Causality for example, is seen as entering into conflict with that form of quantum mechanics covered by the term "Copenhagen School". That the other parts of our metaphysical frameworks may also be corrigible is however, little recognized, at least within the philosophy of science. The categorization in terms of certain entities is obviously corrigible, being connected with our fundamental ontology in the sense of how we "slice up" the world. But so are the Principles of Individuality which are employed (French 1985, 1989*b*, 1989*c*; forthcoming *b*; French and Redhead 1988) and even the logic in which the framework as a whole is embedded (the development of quantum and "paraconsistent" logics being two cases in point).

The exact nature of the corrigibility of our categorial metaphysical frameworks and that of the interrelationships between them, our scientific theories and the evidence supporting the latter, is, of course, a profound and important problem which we cannot deal with in this entirety here. Without doubt such structures are abandoned only very reluctantly; it is only after a particular theory demonstrates certain severe problems, both experimental and otherwise, which hamper its further development, that one might feel tempted to alter the metaphysical structure in which the theory is embedded, causality again being the most obvious example (cf. Worrall 1988, pp. 270-271). Furthermore, such alteration is never absolute, the existence of various hidden-variables theories of quantum mechanics (albeit part of a "degenerating" programme perhaps) incorporating some form of "classical" determinism, attesting to this.

A related point concerns the relationships between the various parts or aspects of the metaphysical structure and whether they are corrigible singly, jointly or in certain groupings. Thus, with regard to quantum mechanics, it has been argued that as well as the constitutive Principle of Causality, certain principles of Individuality have to be abandoned also (Post 1963). Whether and in what way these are connected is left unclear. (In fact, the "classical' Principles of Individuality can be preserved in the quantum context but only at the price of introducing a certain non-classical ontological view regarding the set of states available to the quantal particles; see French 1985; French and

Redhead 1988). To give an alternative example it has been claimed that the choice of a particular logic is, in fact, an empirical affair. However, although we agree that classical logic is not simply "given", it is hard to see how one could empirically test for a particular logic without some ontological meat on the symbolic bones (cf. the final comments in French, 1989*c*).

The questions raised by these two points can only be resolved by detailed conceptual and, primarily, historical studies, by examining, in the latter case, actual historical episodes in which one metaphysical framework is exchanged for another, in order to map out the above inter-relationships. In this paper we simply want to emphasize that theories are proposed, developed, etc. within the context of a certain categorial metaphysical framework and that one may rationally appeal to one or other of the parts of this framewark in order to make a choice between theories which are underdetermined in the "Quinean" sense.

That theoretical developments occur within a certain metaphysical context seems so obvious that it hardly needs pointing out. So obvious perhaps that when this fact *is* acknowledged it is usually only the "constitutive" principles which are discussed, causality and determinism being the most cited examples. Thus Worrall, for example, argues that "At any stage—especially in a well-developed science like physics—there will be a whole range of accepted theories, extending from observational and auxiliary theories (the theory of how light affects photographic emulsion, say) through specific theories (Fresnel's wave theory of diffraction, say) to more general theories (the general wave theory of light, say, that light is *some* sort of mechanical medium) and on to *very* general claims of a metaphysical character (light is some sort of mechanical and deterministic phenomenon)". (Worrall 1988, p. 270; cf. Worrall 1986, p. 679; also Shapere 1986, 1987).

Although we would not go so far as Worrall in asserting that "very general" metaphysical claims can be regarded as "theories" in the same sense as "specific theories", we agree with his subsequent point that there exists "a rough pecking order in case of empirical refutation" (Worrall 1988, p. 270) and that "replacements for the general [metaphysical] idea will standardly be sought only once a *series* of specific theories built around them have proved unsuccessful" (ibid., p. 271).

Perhaps it is the philosophical shock induced by quantum mechanics which causes philosophers of science to run always to causality and determinism when discussing science and metaphysics. Or perhaps the importance of the other aspects of our categorial metaphysical frameworks is simply too obvious. Whatever the reason, little attention is usually paid to them. However, it is worth emphasizing that without some prior categorization of the entities of the universe or domain of knowledge concerned,

scientific theorizing could not even get started. Likewise it is only recently that the importance of Principles of Individuality has been pointed out and their status examined (French 1985; French and Redhead 1988) or the question considered of the possible changes induced by embedding a theory within a non-standard logic or set theory (see, for example, da Costa and Doria, forthcoming).

Although *prior* to a particular body of scientific theoretization in the sense of being already presumed before that particular body of scientific development can get under way, the categorial metaphysical frameworks are not, as we have tried to emphasize, strictly *a priori* or *transcendental*, in the sense of being imposed from without or from above, on the "metalevel". Rather they are derived, at least in part, from previous substantive beliefs. This of course, is implicit in their corrigibility. Again, we leave further discussion to another work (for some interesting suggestions, see Shapere 1987).

Worrall goes on to note that these "more general" metaphysical principles play a dual role within science, both substantive and heuristic: "Not only are they accepted parts of scientific theory, making assertions about the world... they also operate as heuristic principles." (Worrall 1988, p. 271). Put in our terms, the categorial metaphysical frameworks not only provide the context for our theoretical "view" of the world, they also constrain future theoretical development. If a certain theory "fails" in some way to account for the relevant domain, then a successor will be sought for within the same metaphysical context. It is only when the search itself fails, when no adequate successor is found (where the sense of 'adequate' needs to be specified of course) that attention will shift to the metaphysical framework in which the theory is embedded (within the framework itself a hierarchy may exist regarding which aspects to be tinkered with first—the constitutive principles perhaps).

Extending this observation as to the heuristic role of metaphysical principles, we come to the central point of this section: appeal may be made to the categorial metaphysical framework underlying the theories concerned in order to resolve cases of "Quinean" underdetermination. Such cases, by their very nature, will typically involve certain fundamental differences between the metaphysical frameworks in which the theories are embedded. We say "by their very nature" because there are no, nor can there ever be, empirical differences between the theories, which nevertheless are in conflict over certain ontological and hence metaphysical claims. This is clear with regard to the two examples given at the beginning of this section. It is important to emphasize this point since it is only in the context of the attempted elimination of such claims from science that one may assert that underdetermined theories in Quine's sense are merely "notational variants"

of one another; that is, the same theory written in different ways. If the broad ontological implications of our theories, that is, the way they say the world is (albeit perhaps only approximately) are an important and fundamental part of the scientific endeavor, then such arguments are clearly absurd, or at the very least, misguided.

A rational choice may then be made between two "Quinean" underdetermined theories by appeal to their respective underlying metaphysical frameworks, in particular to questions of ontological economy or coherence between the theory and certain aspects of the metaphysical structure, for example. Let us take the underdetermination between field and particle theories as our example.

It can be shown that, contrary to certain claims which are often made, indistinguishable particles in quantum mechanics can be treated as individuals but that problems arise with regard to the Principle of Individuality invoked (French 1985; French and Redhead 1988). Thus it can be demonstrated that considerations of individuality in terms of the Principle of Identity of Indiscernibles or spatio-temporal continuity are ruled out within the quantum context. The only remaining serious contender for such a Principle involves the well-known and much criticized Lockean-type substratum, the perceived philosophical defects in which may, perhaps, be used to argue against any theory based on it (although this is rather a weak basis for theory choice). More serious, perhaps, is the observation that with bosons, fermions and higher-order paraparticles regarded as individuals, in some sense, the results of quantum statistics can only be accommodated by regarding certain sets of states (antisymmetric for bosons, symmetric for fermions) as being "off-limits", unavailable to the particles. That is, the states exist, ontologically speaking, but the particles cannot get into them (French, forthcoming *b*). Furthermore, under this interpretation, it is an ontological commitment of quantum mechanics that every electron, for example, partakes of the state of every other electron in the universe, forming a kind of "global collective" (Post 1963). As has been said elsewhere, "if this sounds too bizarre to be acceptable, it provides another argument for preferring the treatment of indistinguishable particles along the lines provided by quantum field theory" (French and Redhead 1988, p. 245).

The above problems with the metaphysical basis of the particle view all essentially derive from the tension between indistinguishability and individuality inherent in this view. In particular, the particle permutation operators are not regarded as observables in quantum mechanics (putting it simply, particle permutations counted in classical statistical mechanics are not so in quantum form). By treating an assembly of particles as individuals and then permuting the particle labels in constructing the wave-function for the

assembly, the particle interpretation "gets off on the wrong foot", as it were (Post 1963), with the consequences indicated above.

In the field approach, on the other hand, the entities to which the theory is ontologically committed are simply not regarded as individuals; instead of "particles" we have "excitations" of the quantized field (for more on the philosophically interesting aspects of quantum field theory, see Redhead 1983, 1988). That is, quantum field theory takes seriously the fact that the particle permutations are not observable and dispenses with all talk of "individuals" from the very beginning. Instead we have an ontological categorization in terms of "non-individual" field excitations and thus the above problems concerning individuality in the particle view are avoided.

Clearly then it would be rational to choose the field-theoretic approach and thus break the underdetermination, on the basis of the argument that it is free from certain metaphysical difficulties regarding individuality. Of course, this choice should not be regarded as "once and forever" or definite for all time. It may be that the field theoretic view will itself run into difficulties in the future, or the particle interpretation may find a way around the above problems (leading to a new metaphysical framework). Theory evaluation and choice are never "all at once", algorithmic affairs and this is especially true when metaphysical questions are being considered.

The suggestion that the space-time example might be treated in a similar manner is an interesting one. Certain other forms of underdetermination, such as that between linear and cyclic time might also be attacked by way of considerations of individuality and indistinguishability (cf. Reichenbach 1958, pp. 141-143). Certainly there is little worthwhile discussion of such considerations in the literature; however, we shall not further increase the length of our paper by examining them here.

In putting forward this possible use of metaphysical considerations it may be thought, at first blush, that we are closer to Laudan (1984) and Shapere (1986, 1987) in such matters than to Worrall (1988) for example. The former include within the ambit of "methodological principles" exactly these metaphysical considerations, thus lending support to the argument that the history of science has seen changes not only in our substantive beliefs but also in the methodology used to judge them. Worrall, on the other hand, argues that this is too wide and loose a conception of methodology, which should more properly be restricted to an invariant core involving such well-known standards as empirical success, simplicity, etc.

However, if we look more closely at the above "metaphysical difficulties" we see that the arguments they generate involve consideration of the complexity or "bizarre" nature of the ontological views concerned. In arguing for

the field interpretation and against the particle view on the basis of considerations of that aspect of our categorial metaphysical frameworks which has to do with Principles of Individuality, we are essentially extolling the greater metaphysical simplicity of the former as compared with the latter. Thus it may be suggested that metaphysical considerations carry any methodological force in such matters only because of the appeal to an element of Worrall's "invariant core" above (given that the metaphysical framework is one further step removed from experience than the theory, it seems natural that the methodological elements concerned should be "non-empirical", with regard to simplicity or economy, for example). Of course, what we have is an appeal to simplicity at the underlying metaphysical level rather than at the level of the theory itself. Again we shall not pursue this point; suffice to note that in this way some kind of approximation might be effected between Worrall and Laudan and Shapere.

Having come this far, however, the reader might suspect that a circularity has entered our account. Earlier we argued for the corrigibility of our metaphysical frameworks, while above we have suggested that they may be appealed to in theory choice. How can this be so without circularity? How can we appeal, on the one hand, to theoretical and empirical considerations in the shift from one metaphysical framework to another and, on the other, to the structure of such frameworks in choosing one theory over another? Such accusations of circularity are, of course, analogous to Worrall's criticisms of Laundan's and Shapere's attempts to "internalize' methodology (Worrall 1988, 1986, respectively). How can we step out of the circle?

The answer lies in our example above; metaphysical considerations only carry methodological force when conjoined with certain principles, such as ontological economy or simplicity, applied at the metaphysical rather than the theoretical level. Such principles are, of course, logically independent of the constitutive or individuating principles associated with the framework concerned and there is therefore no circularity involved (cf. Körner 1969, pp. 206-207).

Our central point, then, is that what are being judged or evaluated in these situations are not theories or metaphysical frameworks on their own but both together, the former being embedded in the latter. In most cases empirical sucess will win out, eventually, both as regards substituting one theory for another within a particular metaphysical framework and as regards shifting from one to another in a different metaphysical structure. In "Quinean" situations, where this is not possible, attention must be focussed on the underlying metaphysics in order to support the choice of one theory over another. In general such choice and evaluation proceeds according to a set of

criteria, assigned different weights depending on the objective context concerned (French, forthcoming *c*; Körner 1969, p. 207).

7. *Conclusion*

Although the empirical underdetermination of theories is a complex and difficult problem in the philosophy of science, within science itself, that is, in practical terms, it is usually resolved, *Some* choice is normally made, on the basis of certain reasons. In this paper we have argued that in the case of "transitory" underdetermination further developments, both theoretical and experimental, will lead to one theory being chosen over another. What is needed, then, to accommodate this "temporary" underdetermination is some notion of "approximate" or "pragmatic" truth and we have put forward just such a notion in this context.

In the case of "Quinean" underdetermination, empirical or predictive success will obviously not work. In such situations, we have argued, appeal must be made to certain aspects of the underlying metaphysical frameworks in which the theories are embedded. This further brings out the importance of metaphysics in science.

Returning to Newton-Smith's four distinctions above, it is clear that we reject 4.1 while adopting various forms of 4.2, 4.3 and 4.4. Taking these in reverse order, it is equally obvious that we accept 4.4, appealing as we do to "extra-theoretical", metaphysical considerations in order to resolve cases of Quinean underdetermination. 4.2 and 4.3 are a little more complicated however. In rejecting the simple correspondence theory of truth in the case of transitory underdetermination we must change the "ontological ingredient" of "naive" realism. However a full explication of this move requires further elaboration of the "pragmatic realist" position, in particular its relationship(s) to Peircean and "convergent" realism and this we leave to another work (cf. French, forthcoming *a*). Finally, regarding 4.2, da Costa and French's "pragmatic" programme also involves a weakening of the claim that we can have warranted beliefs concerning the truth-values of scientific theories, *where truth is understood in the correspondence sense*. What we have, instead, are beliefs concerning the pragmatic truth of our theories, introducing a further fallibilist dimension to discussions of the relationship between belief and truth. Again this is commented on elsewhere (da Costa and French 1989). (It is important to note that in saying that "partial structures do not capture everything in the domain concerned" we are not advocating a form of the "ignorance" response, since the partial structures "lead the way" to a given "total" structure which models the (intended)

domain exactly. Therefore there are no "inaccessible" facts, in Bergstrom's sense, according to this view).

Of course, further work is needed on all these points, as we have repeatedly said. In particular it is important to look at how scientists themselves resolve apparent cases of underdetermination. Nevertheless we hope that the present work can be considered as a small step towards a more comprehensive understanding of the problem.

State University of São Paulo

State University of Campinas

References

Bergstrom, L. (1984), "Underdetermination and Realism", *Erkenntnis*, **21**: 349-365.

Boyd, R.N. (1973), "Realism, Underdetermination and a Causal Theory of Evidence", *Noûs*, **7**: 1-12.

Buchdahl, G. (1970), "History of Science and Criteria of Choice", in R.H. Stuewer (ed.), *Historical and Philosophical Perspectives of Science,* Minneapolis: University of Minnesota Press.

_____ (1980), "Neo-Transcendental Approaches Towards Scientific Theory Appraisal", in D.A. Mellor (ed.), *Science, Belief and Behaviour,* Cambridge: Cambridge University Press.

Carnap, R. (1956), "O Caracter Metodologico dos Conceitos Teoricos", in *Os Pensadores*, Abril Cultural 1980.

Churchland, P.M. and Hooker, C.A. (1985), *Images of Science: Essays on Realism and Empiricism*, Chicago: University of Chicago Press.

da Costa, N.C.A. (1986), "Pragmatic Probability", *Erkenntnis*, **25**: 141-161.

_____ (1987), "An Outline of a System of Inductive Logic", *Theoria,* **7**: 3-13.

_____ and Doria, F.A. (forthcoming), "Structures, Suppes Predicates and Bolean-Valued Models in Physics", to appear.

_____ and French, S. (1988), Review of *In Contradiction*, by G. Priest, *History and Philosophy of Logic*, **17**: 240-246. (A different version has been published in *The Philosophical Quarterly*, **39**: 498-501 (1989).)

_____ (1989), "Pragmatic Truth and the Logic of Induction", *British Journal for the Philosophy of Science*, **40**: 333-356.

_____ and French, S. (forthcoming *a*). "The Model-Theoretic Approach in Philosophy of Science", to appear in *Philosophy of Science*.

_____ and French, S. (forthcoming *b*), "Ontology and Paraconsistency", to appear in H. Burkhardt and B. Smith (eds.), *The Handbook of Metaphysics*, Philosophia Verlag.

Dorling, J. (1972), "Bayesianism and the Rationality of Scientific Inference", *British Journal for the Philosophy of Science*, **23**: 181-190.

Dummett, M. (1973), *Frege: Philosophy of Language*, London: Duckworth.

Ellis, B. (1985), "What Science Aims to Do", in Churchland and Hooker 1985 pp. 48-74.

French, S. (1985), "Identity and Individuality in Classical and Quantum Physics", Ph.D Thesis Univ. Of London, unpublished.

_____ (1987), "Metafísica e Ciência", forthcoming in *Anais do Simpósio: Atualidade dos Principia de Newton, Rio de Janeiro 1987.*

_____ (1988), "Models, Pragmatic Virtues and Limited Scepticism: The Three Pillars of Constructive Empricism", *Manuscrito*, **11**: 27-46.

_____ (1989*a*), "A Peircean Approach to the Realist-Empiricist Debate", *Transactions of the C. S. Peirce Society*, **25**: 293-307.

_____ (1989*b*), "Why the Principle of the Identity of Indiscernibles in Not Contingently True Either", *Synthese*, **78**: 141-166.

_____ (1989*c*), "Individuality, Supervenience and Bell's Theorem", *Philosophical Studies*, **55**: 1-22.

_____ (forthcoming *a*), "Towards a Pragmatic Realism", to appear.

_____ (forthcoming *b*), "Identity and Individuality in Classical and Quantum Physics", to appear in *The Australasian Journal of Philosophy*.

_____ (forthcoming *c*), "Razão e Racionalidade na Ciência", to appear.

_____ and Redhead, M. (1988), "Quantum Physics and the Identity of Indiscernibles", *British Journal for the Philosophy of Science*, **39**: 233-246.

Friedman, M. (1983), *Foundations of Space-Time Theories*, Princeton Univ. Press.

Giere, R. (1985), "Constructive Realism", in Churchland and Hooker 1985, pp. 75-98.

Glymour, C. (1977), "Indistinguishable Space-Times and the Fundamental Group", in J. Stachel, C. Glymour and J. Earman (eds.), *Foundations of Space-Time Theories*, Minneapolis: University of Minnesota Press, pp. 50-60.

_____ (1980), *Theory and Evidence*, Princeton: Princeton University Press.

_____ (1985), "Explanation and Realism" in Churchland and Hooker 1985, pp. 99-117.

Harman, P.M. (1982), *Metaphysics and Natural Philosophy*, Harvester Press.

Hesse, M. (1978), "Action at a Distance", in E. McMullin (ed.), *The Concept of Matter in Modern Philosophy*, Notre Dame: Notre Dame University Press.

Hiskes, A. (1986), "Friedman on the Foundations of Space-Time Theories", *Erkenntnis*, **25**: 111-126.

Horwich, P. (1982), "How to Choose Between Empirically Indistinguishable Theories", *Journal of Philosophy*, **69**: 61-77.

Laudan, L. (1984), *Science and Value*, Univ. of California Press.

Körner, S. (1969), *Fundamental Questions in Philosophy*, Penguin 1969; rep. 1973.

Malament, D. (1977), "Observationally Indistinguishable Space-Times" in J. Stachel, C. Glymour and J. Earman (eds.), *Foundations of Space-Time Theories*, Minneapolis: University of Minnesota Press, pp. 61-80.

Mikenberg, I., Chuaqui, R. and da Costa, N.C.A. (1986), "Pragmatic Truth and Approximation to Truth", *Journal of Symbolic Logic*, **51**: 201-221.

McMullin, E. (1978), *Newton on Matter and Activity*, Notre Dame Univ. Press.

Mundy, B. (1986*a*), "Embedding and Uniqueness in Relational Theories of Space", *Synthese*, **67**: 383-390.

_____ (1986*b*), "On the General Theory of Meaningful Representation", *Synthese*, **67**: 391-438.

Newton, I. (1982), "De Gravitatione et Aequipondio Fluidoram", in I. Hall and M. Hall (eds.), *Unpublished Papers of Isaac Newton*, Cambridge Univ. Press.

Newton-Smith, W. (1978), "The Underdetermination of Theory by Data", *Proceedings of the Aristotelian Society*, Supplementary Vol. 52, pp. 71-91.

_____ (1980), *The Structure of Time*, Routledge and Kegan Paul.

Post, H.R. (1963), "Individuality and Physics", *The Listener* Oct. 10 1963, available as departmental reprint, History and Philosophy of Science, King's College, London.

Priest, G. (1987), *In Contradiction*, The Hague: Martinus Nijhoff.

Quine, W.V.O. (1975), "On Empirically Equivalent Systems of the World", *Erkenntnis*, **9**: 313-328.

_____ (1979), "Comments on Newton-Smith", *Analysis*, **39**: 66-67.

Quinton, A. (1973), *The Nature of Things*, Routledge and Kegan Paul.

Redhead, M. (1983), "Quantum Field Theory for Philosophers", in P.D. Asquith and T. Nickles (eds.), *PSA 1982*, East Lansing: Philosophy of Science Association, vol. 2, pp. 57-99.

_____ (1988), "A Philosopher Looks at Quantum Field Theory", in H.R. Brown and R. Harré (eds.), *Philosophical Foundations of Quantum Field Theory*, Oxford: Clarendon Press, pp. 9-23.

Reichenbach, H. (1958), *The Philosophy of Space and Time*, New York: Dover.

Shapere, D. (1986), "Objectivity, Rationality and Scientific Change", *PSA 1984*, East Lansing: Philosophy of Science Association, vol. 2, pp. 637-663.

_____ (1987), "Method in the Philosophy of Science and Epistemology", in N.J. Nersessian (ed.), *The Process of Science*, The Hague: Martinus Nijhoff, pp. 1-39.

Torretti, R. (1987), "Physical Theories, II", *Diálogos*, **49**, pp. 147-188.

van Fraassen, B. (1980), *The Scientific Image*, Oxford Univ. Press.

_____ (1983), "Glymour on Evidence and Explanation", in J. Earman (ed.), *Testing Scientific Theories*, Minneapolis: University of Minnesota Press, pp. 165-176.

_____ (1985), "Empiricism in the Philosophy of Science", in Churchland and Hooker 1985, pp. 245-368.

Watkins, J. (1975), "Metaphysics and the Advancement of Science", *British Journal for the Philosophy of Science*, **26**: 91-121.

Worrall, J. (1982), "Scientific Realism and Scientific Change", *The Philosophical Quarterly*, **32**: 201-231.

_____ (1986), "The Background to the Forefront: A Response to Levi and Shapere", *PSA 1984*. East Lansing: Philosophy of Science Association, vol. 2, pp. 672-682.

_____ (1988), "The Value of a Fixed Methodology", *British Journal for the Philosophy of Science*, **39**: 263-275. (Review of L. Laudan, *Science and Value*).