

THEORY AND PRACTICE IN PSYCHOTHERAPY*

Russell Keat[†]

In the preceding chapter (Chapter 4) I examined Habermas's characterization of psychoanalysis as emancipatory. I turn now to the other major element in his view of psychoanalysis as a model for critical social theory, the claim that the criteria of validity for psychoanalytic theory, and of the particular interpretations based upon it, differ significantly from those of both empirical-analytic and historical-hermeneutic sciences. This can be seen as an attempt by Habermas to articulate and justify the general claim made by critical theorists that a distinctive relationship between 'theory' and 'practice' is involved in critical social theories.

Habermas proceeds by examining the relationship between what he calls the 'general interpretations' of psychoanalytic theory, and their employment in the analysis of individual patients in the therapeutic process. These general interpretations are, roughly, the propositions contained in Freud's theories of psychosexual development and the formation of the neuroses. They involve, says Habermas,

assumptions about interaction patterns of the child and his primary reference persons, about corresponding conflicts and forms of conflict mastery, and about the personality structures that result at the end of the process of early childhood socialization, with their potential for subsequent life history.¹

Further:

The learning mechanisms described by Freud (object choice, identification with an ideal, introjection of abandoned love objects) make understandable the dynamics of the genesis of

* [2013] Published as Chapter 5 of *The Politics of Social Theory: Habermas, Freud and the Critique of Positivism*, Basil Blackwell/University of Chicago Press 1981, pp. 133-167; citations should be to this. No changes have been made to the original text, apart from typographical corrections and the addition of bibliographical information. References in the text to other chapters in the book have been retained: chapter 2 ('Value-Freedom and Socialist Theory'), chapter 3 ('Knowledge, Objects and Interests'), chapter 4 ('Psychoanalysis and Human Emancipation'), and the Introduction, are also available at www.russellkeat.net; so, too, is an earlier version of chapter 1 ('The Critique of Positivism').

[†] [2013] School of Social and Political Science, University of Edinburgh: russell.keat@ed.ac.uk. Previously, Department of Philosophy, Lancaster University.

ego structures at the level of symbolic interaction. The defense mechanisms intervene in this process when and where social norms, incorporated in the expectations of primary reference persons, confront the infantile ego with an unbearable force, requiring it to take flight and objectivate itself in the id. The child's development is defined by problems whose solution determines whether and to what extent further socialization is burdened with the weight of unsolved conflicts and restricted ego functions, creating the predisposition to an accumulation of disillusionments, compulsions, and denials (as well as failure) - or whether the socialization process makes possible a relative development of ego identity.²

Apart from the presence here of Habermas's problematic version of the id,³ this is an admirable summary of these elements of psychoanalytic theory. But Habermas goes on to argue that general interpretations only partially resemble the general theories of empirical-analytic science in their criteria of validity, as can be shown by examining their application to the life-histories of individual patients in the therapeutic process. Such applications he terms 'constructions' - for instance, suggestions by the analyst about particular occurrences in the patient's past; and Habermas claims that:

The criterion in virtue of which false constructions fail does not coincide with either controlled observation or communicative experience [i.e. with the criteria involved in either empirical-analytic or interpretive sciences]. The interpretation of a case is corroborated only by the successful *continuation of a self-formative process*, that is by the completion of self-reflection, and not in any unmistakable way by what the patient says or how he *behaves*.⁴

So for Habermas the truth or falsity of these constructions - and thus, derivatively, of the general interpretations upon which they are based - is determined partly by how the patient responds to them: that is, by the success or failure of the therapeutic process, itself conceived as a self-reflective movement towards autonomy.

To assess Habermas's view of the criteria of validity for critical social theories, we need first to consider the place of predictions and practical success or failure in the empirical-analytic sciences. I will argue that Habermas's contrast between such theories and the general interpretations of psychoanalysis involves too simple a view of this issue. It will be convenient to begin by examining Karl Popper's well-known attack on the scientific status of psychoanalytic theory, which reveals a similar form of over-simplification.

1. Psychoanalytic theory as 'science': problems of theory-testing

In a number of autobiographical passages Popper tells us how his early encounters in Vienna with

Freudian psychoanalysis, Marxism and Adlerian psychology, and his comparison of them with Einstein's relativity theory, led him to propose falsifiability as the demarcation criterion between science and non-science.⁵ In the course of these reminiscences he claims that Freudian theory fails to meet this criterion: its propositions are not falsifiable, and thus not scientific.

The significance of Popper's view can be brought out by considering its implications for the vast literature devoted to the empirical testing (both clinical and non-clinical) of various propositions of Freudian theory. In these studies we find considerable disagreement as to whether any of these propositions are actually supported or refuted by the relevant evidence. For instance, Paul Kline argues that for at least some of them - such as the relationship between child rearing procedures and the presence of 'anal' character-traits in adults - there is no disconfirming evidence and some positive support; whilst H. J. Eysenck and Glenn Wilson have claimed an absence of support, or some disconfirming evidence, for all of them so far tested.⁶ But it would seem that the intelligibility of these disagreements presupposes that Popper's criticism of Freud is mistaken, since the disputed studies all involve attempts to do what Popper claims is impossible - to assess the degree of support for Freudian theory by deriving testable predictions from it. According to Popper this cannot be done: the theory is untestable, unfalsifiable and hence unscientific.

To examine this claim, we must first explore certain difficulties in the falsifiability criterion itself, and its associated account of theory-testing. Popper argues that because of the logical problem of induction it is not possible to verify scientific theories. Such theories consist in 'strictly universal statements', whose possible falsity is logically consistent with any finite amount of supporting evidence: thus they cannot be conclusively verified by establishing the truth of particular predictions derivable from them. Inductive argument, he claims, is not rationally justifiable. But fortunately science - which is for him a paradigmatically rational form of enquiry - does not require the use of inductive argument. It can proceed through deductive argument alone, which is involved in the testing of theories at two distinct points: in the derivation from the theory of testable predictions; and, if these fail, in the argument from this to the falsity of the theory. Theories which have not in their tests been refuted are not to be seen as thereby positively confirmed or supported. Instead, they are said by Popper to be 'corroborated', meaning that they *so far* have not been refuted. But it is essential to his overall position that corroboration is not regarded as providing positive support for the belief that a theory will pass future tests: this, for Popper, would involve mistakenly accepting the legitimacy of inductive argument.⁷

I shall ignore here the debates about Popper's conception of science generated by the work of Kuhn, Feyerabend and Lakatos⁸ (though it would be interesting to examine the history of psychoanalysis in terms of their rival accounts), and focus on just two criticisms of Popper's position.⁹ First, it seems

very doubtful whether any adequate account of the assessment of scientific theories can, like Popper's, do without the idea that successful predictions give some form of positive support to the theories from which they have been derived. To this extent, the 'logic' of theory-testing must be partly inductive. Second, there are at best very few scientific theories that could satisfy Popper's demarcation principle of falsifiability, since it is generally the case that in order to derive determinate predictions from a theory, assumptions that are additional to the propositions of the theory itself have to be made. I shall now elaborate briefly on each of these points.

The first criticism can be justified in a number of ways. For instance, it can I think be shown that scientific practice depends upon the assumption that, other things being equal, a theory that has 'passed' a particular kind of test will continue to do so in future tests of the same kind. Further, it is difficult to understand how it can be rational to base one's actions on well-tested theories, unless their past success is seen to provide some positive grounds for believing that predictions derived from them will turn out to be correct in the future: that is, in this context, when they are acted upon.

However, in saying that one must therefore regard the logic of theory-testing as at least partly inductive, it is important to emphasize that this does not imply acceptance of certain views often termed 'inductivist'. First, it does not mean accepting that any theory can be conclusively verified by reference to positive evidence. Second, it does not entail the view that the logic of evidential support can be mapped onto the calculus of probability statements and their relations; nor that it can be analysed in terms of a frequency interpretation of that calculus. Third, to claim that successful predictions provide positive support for theories is compatible with accepting an a-symmetry between the significance of confirmatory and dis-confirmatory evidence. That is, it may be recognized that negative evidence is a more powerful reason for rejecting theories than positive evidence is for accepting them, but without thereby rejecting the significance of the latter altogether, as is proposed by 'pure falsificationists' such as Popper.

Finally, it seems that any adequate account of positive evidential support must involve comparisons between alternative possible theories. That is, the extent to which successful tests provide good grounds for accepting (provisionally) a particular theory must be partly determined by the relative plausibility of alternative theoretical explanations of the same evidence. In this respect, I think that most scientific argumentation is concerned with the question of which theory provides the best, or better, explanation; and I doubt whether what philosophers have said about either deductive or inductive argument is of much help here.¹⁰

Moving now to the second criticism, concerning falsifiability, the basic difficulty is that one cannot derive determinate predictions from most scientific theories without additional assumptions, usually

termed 'auxiliary statements'. So whether a theory is falsifiable (and thus 'scientific') may not be knowable at any particular time, since the appropriate auxiliaries may not be available; and it may not be possible to know that they will never become available. One obvious source of this difficulty is the existence of suitable scientific instruments: without these a theory may remain untestable, but clearly this should not by itself lead us to regard it as unscientific. However, once the appropriate auxiliaries are provided, the logical outcome of theory-testing becomes indeterminate with respect to the possible refutation of the theory. For if predictions derived from the theory together with the auxiliary statements fail, all that follows (deductively) is that at least *one* of the premisses in the derivation must be false. But which one is false - in particular, whether it is the theory itself or its auxiliaries - cannot thereby be determined. Thus no conclusive refutation of a theory is possible where statements other than those of the theory being tested are required in the derivation of predictions.

These are the basic difficulties for falsificationism.²² But it will be helpful to add a few further points concerning the character of auxiliary statements. First, it is worth distinguishing two different kinds of auxiliaries: singular statements and theoretical statements. The former are always necessary if predictions about particular phenomena (as opposed to other universal relations) are to be derived from the tested theory. (A simple example of this would be the need to know the actual values of the independent variables in a theory, in order to predict the values of the dependent variables.) The latter, theoretical auxiliaries, may not always be necessary to derive predictions; but very often they are. For instance, any use of instruments will require the theoretical assumptions upon which they are based to be included in the derivation of test-predictions.

The need for auxiliary assumptions (both singular and theoretical) is especially apparent where the tested theory itself involves forms of idealization or abstraction, as most theories in the advanced sciences do.¹² For instance, a theory that uses the concept of a frictionless surface cannot generate determinate 'real world' predictions without additional assumptions about the actual degrees of friction present in a particular surface (singular auxiliaries), and about the deviations from the idealized theoretical relationship that will result from these (theoretical auxiliaries). Finally, a rough distinction can be drawn between those theoretical auxiliary statements that come from the same general theoretical area to which the tested theory belongs, and those that are taken from relatively independent and unrelated areas of theoretical work. An example of the latter would be the auxiliaries involved in using optical microscopes to test a theory about neuronal cell-structures. Here, the auxiliary theory is drawn from optics, and thus from a theoretical area (relatively) independent from neurophysiology.

With these general points about theory-testing and evidential support in mind, we can now return to Popper's criticism of Freudian theory. Its main elements are indicated in the following passages:

I found that those of my friends who were admirers of Marx, Freud and Adler, were impressed by a number of points common to these theories, and especially by their apparent *explanatory power*. These theories appeared to be able to explain practically everything that happened within the fields to which they referred. The study of any of them seemed to have the effect of an intellectual conversion or revelation, opening your eyes to a new truth hidden from those not yet initiated. Once your eyes were thus opened you saw confirming instances everywhere: the world was full of verifications of the theory. Whatever happened always confirmed it.¹³

Popper goes on to describe an episode in which Adler had 'found no difficulty' in analysing a particular case without ever seeing the person concerned, and had justified his confidence in doing so by his 'thousandfold experience'. Popper says he came to think that all Adler had done was to show in each instance,

that a case could be interpreted in the light of the theory. But this meant very little, I reflected, since every conceivable case could be interpreted in the light of Adler's theory, or equally of Freud's. I may illustrate this by two very different examples of human behaviour: that of a man who pushes a child into the water with the intention of drowning it; and that of a man who sacrifices his life in an attempt to save the child. Each of the two cases can be explained with equal ease in Freudian and in Adlerian terms. According to Freud the first man suffered from repression (say, of some component of his Oedipus complex), while the second man had achieved sublimation [I omit Popper's Adlerian explanation, since I will not be discussing his 'equal ease' claim] . . . I could not think of any human behaviour which could not be interpreted in terms of either theory. It was precisely this fact - that they always fitted, that they were always confirmed - which in the eyes of their admirers constituted the strongest argument in favour of these theories. It began to dawn on me that this apparent strength was in fact their weakness.¹⁴

Most discussions of this criticism have assumed that it is fairly clear what exactly Popper is saying. But I do not think this is so, and will therefore proceed by offering a number of different interpretations. The first is that Popper appears to claim that the fact that everything anybody does can be explained in terms of Freudian theory should not be seen as providing massive confirmation for it, but instead as casting doubt upon its scientific status. Taken at face-value, this is a curious argument. For were a theory designed to explain all human behaviour, why should the fact that it can do so

actually count against its scientific status, even if we go along with Popper and accept that passing tests shows only that a theory has not been refuted, and not that it is thereby confirmed?

Popper, after all, adopts the deductive-nomological model of explanation. So if Freudian theory can 'explain everything', this must mean that statements describing every actual (and possible) item of human behaviour can be derived from the theory (together with, of course, appropriate auxiliary statements). Now Popper maintains that scientific status is determined by falsifiability. Falsifiability involves testability; and tests require the derivation of determinate predictions. But if deductive-nomological explanations can be provided, the theory must be testable, and hence scientific: for the logic of explanation and prediction are the same, on this model of explanation.¹⁵ So it cannot be the case both that a theory can explain everything, and that it is unfalsifiable.

If any plausibility is to be given to Popper's argument, then, we must find another interpretation of it. One alternative might be this: Freudian theory is unscientific because 'it can explain everything' in the sense that for any particular item of behaviour, both its occurrence and its *non*-occurrence can be derived from the theory. Now if this were so it would indeed be an overwhelming objection to the theory: for it would follow that it was internally inconsistent. For if it is true that from precisely the same set of theoretical statements both the occurrence and non-occurrence of some phenomenon can be derived, there must be logical contradictions within this set of statements. (And from contradictory statements any statements can be derived, including further contradictory ones.)

Although Popper does not explicitly claim that Freudian theory is internally inconsistent - and one might expect that he would, if this were his view - this might seem to be suggested by the comments he makes about the hypothetical example of the two men, one of whom pushes a child into the river, whilst the other jumps in to save a child's life. For he suggests that both actions could equally well be 'explained' on the basis of Freudian theory; and this must mean that statements describing both events could be derived from it. However, this by itself would not show that contradictory statements could be derived, since there is no contradiction between the statement 'X pushed a child into the river', and 'Y rescued a child from the river', unless it is also true that 'X' and 'Y' refer to the same person, the child and the river are the same in both statements, and so on.

In any case, Popper also says about this hypothetical example that the former action could be explained by reference to repression, and the latter by reference to sublimation. But if this is so, then even were 'X' and 'Y' to be identical, etc., no internal inconsistency in the theory would be demonstrated, since it would no longer be the case that contradictory statements were derivable from the *same* set of theoretical statements. For in the former case, statements about repression would be present in the assumptions, whilst in the latter these would be replaced by statements about

sublimation. Further, if both *explanandum* statements could be derived in this way - and thus the two phenomena explained - how could this show the theory to be unscientific?

So we must try another interpretation of Popper's criticism. Suppose that on the basis of psychoanalytic theory it had been predicted that a person would push a child into a river. And suppose that this does not in fact happen - indeed, that 'the opposite' happens instead. Perhaps what Popper has in mind is that in this and other such cases Freudian theory can remain unrefuted, by substituting 'sublimation' for 'repression' in the set of statements from which the initial prediction was derived. But, if this is Popper's complaint, why should we take it as showing the unfalsifiability of Freudian theory? For all that would be involved would be the substitution of one singular auxiliary statement for another (concerning the occurrence of sublimation instead of repression), together with their related theoretical auxiliaries. In other words, the failure of the prediction is explained by reference to auxiliary statements, not to the main theoretical statements. But as my earlier discussion of theory-testing suggested, there is nothing unscientific about such a procedure.

However - and here we arrive at what seems the only remotely plausible version of his criticism - Popper might go on to claim that such a procedure is indeed unsatisfactory if no justification can be provided for the adjustment other than the fact that by this means the apparent refutation can be avoided: in other words, that the adjustment is entirely 'ad hoc'. But if this is his objection, it would be incumbent upon him to show that Freudian theory is inherently ad hoc in this way. And this seems very difficult to show, though the accusation does point to certain potentially abusable features of Freud's theory.

But before briefly discussing these, it is worth dealing with Popper's criticism in the case of his hypothetical example. To avoid the objection that the procedure is ad hoc, all that is required is the possibility of adducing evidence that sublimation rather than repression had occurred, other than the person's child-rescuing behaviour itself. I can see no difficulty in doing this: psychoanalytic explanations are not normally based upon a single occurrence, and possible independent evidence of sublimation should in principle be available.

However, I think it is true that there is a definite potential for 'ad hocery' in psychoanalytic theory. Consider, for instance, the following (fairly typical) claim by Freud:

We can at any rate lay down a formula for the way in which character in its final shape is formed out of the constituent instincts: the permanent character-traits are either unchanged prolongations of the original instincts, or sublimations of those instincts, or reaction-

formations against them.¹⁶

Now were this to be all that Freud had to say about the relationships between ‘the original instincts’ and adult character formation, one could object that this part of his theory is pretty well untestable, since no indication is given of what it is that leads to one rather than another of these three possible processes taking place. In particular the concept of a reaction-formation (like several other psychoanalytic concepts, such as displacement) is especially problematic since, roughly speaking, its supposed presence will generate precisely the opposite behaviour to what would otherwise have been expected. It is therefore a concept with what might be called a ‘high ad hoc potential’ in explaining away inconvenient phenomena. But this is only a potential: it need not be actualized. In order to avoid the danger, though, Freud’s ‘formula’ here has to be developed in such a way that the conditions in which a reaction-formation occurs be specified independently of the phenomenon which is said to be explained by reference to it. And not only must this be done at a theoretical level, but the attempt must also be made to specify the kind of evidence that would indicate that such conditions were or had been present in concrete situations.

However, as I noted earlier, the absence at a particular time in the development of a theoretical system of the necessary auxiliary statements does not show that the theory is unfalsifiable, and hence unscientific; and the problems just described seem basically to concern the provision of appropriate auxiliaries. So I conclude that there is no good reason for accepting Popper’s judgment that Freudian theory is unscientific. But this discussion has put us in a better position to examine the relationship between theory and therapeutic success in psychoanalysis. Thus I turn now to Habermas’s view of this relationship.

2. The criteria of validity for general interpretations

Habermas claims that the criteria of validity for general interpretations differ from those of both empirical-analytic and hermeneutic sciences (nor are they simply a combination of the two). He concentrates mainly on the differences from the former, and I shall follow him in this. One major element of the contrast is presented in this passage:

general interpretations do not obey the same criteria of refutation as general theories. If a conditional prediction deduced from a lawlike hypothesis and initial conditions is falsified, then the hypothesis may be considered refuted. A general interpretation can be tested analogously if we derive a construction from one of its implications and the communications of the patient. We can give this construction the form of a conditional prediction. If it is correct, the patient will be moved to produce certain memories, reflect on a specific portion of forgotten life history, and overcome disturbances of both communication and behaviour. But

here the method of falsification is not the same as for general theories. For if the patient rejects a construction, the interpretation from which it has been derived cannot yet be considered refuted at all. For psychoanalytic assumptions refer to conditions in which the very experience in which they must corroborate themselves is suspended: the experience of reflection is the only criterion for the corroboration or failure of hypotheses. If it does not come about, there is still an alternative: either the interpretation is false (that is, the theory or its application to a given case) or, to the contrary, the resistances, which have been correctly diagnosed, are too strong.¹⁷

But the contrast stated here depends upon Habermas's failure to recognize the place of auxiliary statements in the derivation of predictions from scientific theories. It is of course true that if a prediction *can* be derived from a single hypothesis (together with initial conditions) and the prediction is false (and the statements specifying the conditions are true), then the hypothesis is refuted. But this is not the usual situation in theory-testing. The indeterminacy of outcome which Habermas presents as peculiar to tests of general interpretations is instead typical in testing scientific theories.

In the kind of case he considers, the psychoanalytic prediction that, say, the patient will accept a particular construction, will require as one of its assumptions the absence of resistance in the patient. Thus, when the patient rejects the construction it may be that this assumption is false; or that the construction itself is false; or that one or more of the general interpretations from which, together with various singular statements about the patient's past, this construction has been derived, is false; and so on. There is nothing methodologically peculiar about this. For instance, from Newton's first law - that a body will continue either at rest or in rectilinear motion in the absence of external forces - we can predict the actual motion of a particular body only with further assumptions about what forces are operating upon it, and about what laws they are subject to. Failure of the prediction does not refute the law, unless we know these auxiliary statements are true.

There is another feature of the passage just quoted requiring comment, which will help to indicate the extent of the 'gap' between general interpretations, predictions about the patient's production of memories, reflections upon portions of forgotten life-history, and the overcoming of communicative and behavioural disturbances. Now the last of these involves, in effect, some elements of 'therapeutic success'. So it seems that Habermas is claiming it is possible to derive predictions about the successful outcome of the therapeutic process from general interpretations, and then saying that failure of such predictions does not refute these interpretations in the way that failed predictions refute scientific theories.

In general terms, my response to this is the same as that just given for the patient's rejection of a particular construction. But it is worth exploring further the kinds of auxiliaries required to derive predictions about therapeutic success from general interpretations. For, roughly speaking, these auxiliary statements would have to include *a theory about the therapeutic process itself*; and this may involve propositions that are to some extent independent of the general interpretations. So this might be a situation where auxiliary statements are drawn from a relatively distinct area of theory, as in the example given in the previous section, of the use of optical theories in testing physiological theories. If this were so, failure of therapeutic practice might indicate deficiencies in the theory of the therapeutic process, and not in the general interpretations of psychoanalytic theory itself.

The fact that predictions about therapeutic results may require auxiliary statements concerning the nature and outcome of therapeutic processes, in addition to general interpretations, also has important consequences for the question of how much confirmation can be provided for these interpretations by successful therapeutic results. I shall comment on this issue later on. But first, the comments just made about the relations between general interpretations and a theory of therapeutic processes must be somewhat qualified, at least in the case of psychoanalysis. For a comparison with the relations between physiological and optical theories may suggest a greater degree of independence between the two theoretical areas than probably exists here.

For instance, psychoanalytic therapy makes considerable use of the phenomenon of 'transference', where the patient's feelings towards others, such as his or her parents, are projected onto the analyst. This provides both an aid for the analyst in interpreting the patient's activities, and a means by which the patient can become aware of the nature of these feelings and 'work through' them. So not only is there a partial mapping of the two sets of relationships, with the same theoretical claims potentially explaining them, but the concept of projection employed to explain the process of transference is itself a central one in psychoanalytic theory. Somewhat similarly, the concept of resistance functions both in the theoretical explanation of the therapeutic process and also, partly through its logical relationship to the concept of repression, in these general interpretations. So there is by no means a complete divorce between psychoanalytic theory and the auxiliary statements concerning the therapeutic process.¹⁸

None the less the derivation of predictions about therapeutic results requires a good deal more in the way of theoretical assumptions than the theory of psychosexual development that Habermas regards as being 'tested' (in a methodologically peculiar manner) in the therapeutic process. And when we take account of these additional assumptions, I believe this supposed peculiarity disappears, depending as it does upon a contrast with an oversimplified view of the logic of scientific theory-testing. However, this is not the only way in which Habermas tries to show what is distinctive about

the criteria of validity for general interpretations. He also offers a more positive account (involving a further contrast) as an alternative to their assimilation to the testing of theories in empirical-analytic science, and I will now examine this.

Habermas says that:

Whereas in other areas theories contain statements about an object domain to which they remain external as statements, the validity of general interpretations depends directly on statements about the object domain being applied by the 'objects', that is the persons concerned, to themselves [Thus] analytic insights possess validity for the analyst only after they have been accepted as knowledge by the analysand himself.¹⁹

This contrast between scientific theories and general interpretations is further specified as follows:

When valid, theories hold for all who can adopt the position of the inquiring subject. When valid, general interpretations hold for the inquiring subject and all who can adopt its position only to the degree that those who are made the object of individual interpretations *know and recognize themselves* in these interpretations. The subject cannot obtain knowledge of the object unless it becomes knowledge for the object - and unless the latter thereby emancipates itself by becoming a subject.²⁰

In these passages, Habermas appears to make at least two claims: that it is a necessary condition for the truth of any particular psychoanalytic interpretation (and thus, derivatively, of general interpretations) that the patient accepts it; and that, since this self-recognition itself either constitutes or generates 'emancipation', therapeutic success (defined as emancipation) is also a necessary condition for the truth of such interpretations. I begin by commenting on the former claim.

First, Habermas's insistence upon the patient's acceptance as a necessary condition for true interpretations cannot be due to his regarding psychoanalytic theory as a hermeneutic science, and therefore (on some views of the criteria for correct hermeneutic interpretations) requiring agreement by the 'object' of interpretation. For he wishes to show that the criteria of validity for critical social theories, such as psychoanalysis, are distinct from those of both empirical and hermeneutic sciences. Thus he says that:

As with other forms of knowledge, the testing of hypotheses in the case of general interpretations can follow only those rules that are appropriate to the test situation. Only they guarantee the rigorous objectivity of validity. Whoever demands, to the contrary, that general

interpretations be treated like the philological interpretation of texts or like general theories and subjected to externally imposed standards, whether of a functioning language game or of controlled observation, places himself from the very beginning outside the dimensions of self-reflection, which is the only context in which psychoanalytic statements can have meaning.²¹

Second, despite his claim that the patient's acceptance is required for the truth of psychoanalytic interpretations, Habermas also endorses Freud's view that a patient may reject an interpretation which is in fact correct, or indeed accept one that is not.²² The former case is explained by Freud in terms of resistance: he argues that it is only in the absence of resistance that a patient will agree to the correct interpretation. The overcoming of resistance is, of course, related to the dissolution of repression. For Freud, correct interpretations will only be accepted by the patient when the therapeutic process has succeeded in achieving this. But if this is so, it must be possible to determine the truth or falsity of an interpretation independently of the patient's agreement to it. When resistance and repression are overcome the patient, as it were, no longer has any interest in rejecting correct interpretations. Thus only when the therapeutic process has achieved its aim will acceptance follow.

However, this is not to deny that there are serious problems about the validation of psychoanalytic interpretations, which basically hinge upon the extent to which patients may be said to have some ultimate epistemic 'authority' about the nature of their own unconscious states; and the following brief comments may serve to indicate some of those problems. In the course of the therapeutic process patients may do all sorts of things which, to the analyst, indicate the presence of unacknowledged, unconscious items. For instance, a patient might respond angrily to a particular suggestion as to his or her unconscious feelings towards one of his or her parents. How is this anger to be understood? It might be a defence against a correct interpretation; it might display frustration at the apparent stupidity of the analyst; it might be due to a recent encounter between the patient and someone else, in which a row took place about the personal character of that parent; and so on. Which of these explanations is correct is enormously difficult to judge.

Similar difficulties are common in 'everyday life', as well as therapeutic situations. For instance, it seems we sometimes find someone else's interpretation of our motives or attitudes, which we had previously denied or ignored, 'feels' or 'seems' right. And it may be that this 'first person authority', involving some intuitive sense, is the ultimate criterion for assessing such interpretations. But it seems that we often come to revise the judgments previously made with such authority; and also that, in certain situations, and in relation to certain people, we are highly suggestible. Furthermore, our intuitive sense of what is plausible as an interpretation of our unconscious feelings is itself partly influenced by exposure to various theories about the character of unconscious processes, such as psychoanalysis itself. Thus recourse to the patient's acceptance may give a spurious degree of support

to psychoanalytic interpretations just because of the cultural influence of psychoanalytic concepts.

The one point that seems clear is that without some resolution of these kinds of difficulties, Habermas's view that the patient's acceptance and self-recognition is a necessary condition for the truth of psychoanalytic interpretations cannot be accepted. So I turn now to his other claim, about the successful outcome of the therapeutic process as a criterion of validity. To assess this in detail one needs to give a specific meaning to the idea of therapeutic success. I shall discuss this issue in the last section of the chapter. For the moment it will be sufficient to assume a fairly loose definition which follows Habermas's emphasis upon achieving autonomy, itself understood as the absence of compulsiveness (discussed in section 3 of the previous chapter).

If Freudian theory is correct, the sources of compulsive activity lie in certain repressed, unconscious states or processes generated in the course of the patient's life-history; and the theory attempts to specify how in general these arise, through what mechanisms they produce neurotic systems, and so on. It follows that if by suitable therapeutic techniques these determinants of the patient's present behaviour can be removed, the goal of psychoanalytic therapy will be at least partly achieved. But the hypothetical is crucial here; for it cannot be said that psychoanalytic theory by itself guarantees either the effectiveness of any particular set of therapeutic techniques, or indeed that any such effective techniques can be devised.

That is, the truth of psychoanalytic explanations of the formation of the neuroses is consistent with the failure of therapeutic practices guided by psychoanalytic theory. For although that theory both indicates what needs to be achieved in the therapeutic process, and also provides some theoretical claims about what kinds of techniques may prove effective (such as the use of transference, free association in the interpretation of dreams, and so on), it is quite possible that additional theoretical knowledge would be required, which is not derivable from psychoanalytic theory, to enable effective therapeutic techniques to be developed. Thus successful results of psychoanalytically guided therapeutic processes cannot be regarded as a necessary condition for the truth of psychoanalytic theory: indeed its truth is compatible with a considerable degree of pessimism about the possibilities for therapeutic success.²³

The question of pessimism will be further explored in the next section, where I note how some ways in which the relationship of a critical social theory to successful practice has been defined seem logically to exclude the possibility of a true theory with pessimistic predictions. But before pursuing this, two more comments need to be made about Habermas's account of psychoanalytic theory's criteria of validity. First, not only is therapeutic failure compatible with the truth of general interpretations, but therapeutic success through the use of typical psychoanalytic techniques may not

give very strong support to them. For it is possible that the effectiveness of such techniques may be explained by a non-psychoanalytic theory. That is, there may be features of the psychoanalytic process which are responsible for whatever success is achieved through it, and the theory that explains why this is so may neither support, nor be supported by, psychoanalytic theory itself. There may, for instance, be ways of removing the compulsive character of neurotic activity which do occur in the therapeutic process, but whose possible theoretical rationale is either unknown, or clearly divergent from the psychoanalytic rationale. This of course is simply a special case of the general point about theory-confirmation noted in the previous section: positive evidence supports a particular theory only to the extent that other theories are unable to explain it or, if they can, to the extent that they are in other respects less plausible.

Second, Habermas appears to believe there is a close link between the patient's acceptance of a psychoanalytic interpretation, and a successful therapeutic outcome: 'the subject cannot obtain knowledge of the object unless it becomes knowledge for the object - *and unless the latter thereby emancipates himself by becoming a subject*' [my italics]. But it is unclear whether he is claiming here that 'becoming a subject' directly results from self-knowledge, or whether the former is an additional process, for which the latter is only a necessary condition, so that both are necessary but partly independent conditions for the truth of psychoanalytic interpretations. I have argued that neither is a necessary condition. But it is worth adding that there is nothing in Freudian theory to suggest that acceptance by itself guarantees a successful therapeutic outcome, since this would ignore (as Habermas tends to throughout) the affective or emotional processes that must be involved in successful therapy, in addition to the cognitive ones. The sense in which acceptance could be a sufficient condition of therapeutic success, in psychoanalysis, seems to be that only when resistance has been overcome, will the patient in fact accept the interpretation. But this would make acceptance the result rather than a determinant of therapeutic success, as I suggested earlier.

3. Theory and practice: the place of self-fulfilling predictions

So far I have argued that Habermas fails to show through his account of psychoanalysis how the relationship of theory to practice differs from that of empirical-analytic or hermeneutic sciences. This does not of course mean there is nothing valuable in his conception of critical social theory, in particular of how such a theory should be guided by an interest in (practical) emancipation; but this may be accepted without also believing that such a theory has distinctive criteria of validity. So I will now examine Brian Fay's account, in *Social Theory and Political Practice*, of how a critical social theory's criteria of validity are tied in a distinctive way to successful practice. Fay's position is strongly influenced by Habermas's, but it has the virtue of including a more explicit account of the supposedly distinctive criteria.²⁴ However, this virtue has its costs, since I think it reveals more clearly

the difficulties of this position.

Fay argues that critical social theories should be addressed to those already experiencing various frustrations and dissatisfactions in their lives, and be designed to aid them in finding a way out of this situation. Partly, this will involve helping them achieve a better degree of self-understanding; and because of this a critical social theory must include hermeneutic elements. Further, such a theory must help its 'clients' come to recognize how various structural features of their society contribute to their plight, and thus help dispel the ideological concealment of such processes. But also, Fay says, a critical social theory must indicate to its clients 'a means by which they can solve the problems which are facing them'; and the truth of such a theory '*is thereby judged partially on the basis of whether the satisfactions which it promises are forthcoming*'.²⁵

As Fay himself partly acknowledges, it is not clear from these requirements alone how this conception of critical social theory differs from positivist or hermeneutic accounts.²⁶ For it could be argued that, at least with appropriate auxiliary statements (and I shall exclude this complication here, since it has been sufficiently explored in the previous section), it will often be possible to link scientific theories with 'guides to practice' by deriving predictions about the likely outcomes of possible courses of action. It is of course true that many theories are developed without attempting to specify such practical links; but the demand that these be provided makes no difference to the epistemological status of the theories, even if it may often be a rather difficult demand to satisfy. What is it then that does make critical social theories' criteria of validity distinctive? Fay's answer is contained in the following two passages, the first of which carries on directly from the quotation above:

It follows from this that the objects of study of this science - the social actors about which it seeks to provide an understanding - actually help to determine the truth of this science's theories by their reaction to them. According to the critical model, one criterion of the truth of social scientific claims is the response which these people make to these claims, and it is for this reason that the application of its purported truths requires a central and determining role for the actors who are to be affected by it.²⁷

The kind of 'response' Fay has in mind here is presumably indicated by this:

the theories of such a science will necessarily be composed of, among other things, an account of how such theories are translatable into action, and this means that the truth or falsity of these theories will be partially determined by whether they are in fact translated into action.²⁸

This seems a strange requirement. Not only must a critical social theory inform its clients what they can do to reduce their dissatisfactions, but its truth is to be determined, not just by whether, if the recommended action is taken, these are reduced, but also by whether in fact this course of action is performed. Thus a critical social theory will be refuted by its practical recommendations being ignored rather than (or in addition to) these activities failing when they are performed. I suggest a number of objections to this claim.

Fay says that this distinctive criterion of truth actually follows from the fact that critical social theories are designed to be ‘translatable into action’. I cannot see why this is so. It may be possible to derive predictions about the outcomes of courses of action from a theory; but there is no reason for expecting that from the very same set of statements it would also be possible to derive the prediction that when this recommendation is offered to the clients the advice will be taken. To make this kind of prediction we would normally have to add to the initial set of premisses a further set of statements describing some theory about the effects of the dissemination of (various kinds of) theories upon the activities of (various kinds of) human agents. And it is quite possible that this theory would be largely independent of the theory from which the initial ‘advice’ was derived. If Fay wishes to propose that such theories about the effects of theory-dissemination be developed, well and good; but this is a different matter from his claim about actual translation into practice as a condition for the truth of action-guiding theories. One might perhaps revise Fay’s position by saying that it is *pointless* to produce theories which are or will be ignored by their clients; but this would be a significantly different claim that has no implications for the criteria of validity of critical social theories.

It might further be objected that Fay’s position is not only epistemologically mistaken, but could also be politically dangerous. For there are hints in what he says that theories can be validated by the self-fulfilling character of their predictions; and if this were to be accepted it might seem to legitimate certain abuses of social theory which, I am sure, Fay would not wish to encourage. In a footnote to the passage I have been commenting on, he says that one of the reasons he has for saying that translation into action only partially determines the truth or falsity of a critical social theory is that:

a theory might be constructed but then lost before it was disseminated, and one would hardly want to claim that, because it never affected social life, it was therefore false. A critical social theory is one which offers an account of future social developments and how they will occur partially because of the existence of this account itself, and if, as a result of the account’s not becoming known, the social order develops along lines other than the theory predicted, then the truth of the theory is indeterminate [i.e. rather than its being refuted].²⁹

I think it follows from this that where a social theory is disseminated, its predictions about a future social development can be confirmed by the fact that those who take notice of it then act in such a way that this development occurs; and presumably if they act so as to prevent its occurrence this would refute the theory. Now imagine the following situation. Some social theorists adopt as their clients a group which feels frustrated and dissatisfied by the presence of non-white immigrants in Britain. They develop a theory to explain how immigration has resulted from the economic relationships between Britain and various other countries, and predict that racial tensions will result in armed conflict between immigrants and non-immigrants. This prediction causes the existing tensions to increase, their clients join the National Front, and eventually the armed conflict results. The predicted rivers of blood begin to flow; and the theory from which the prediction was derived is thereby confirmed in an epistemologically impeccable manner - if this distinctive criterion of validity for action-guiding critical social theory is accepted.

It is important here to be clear about precisely what sorts of claims are contained in the theory from which this self-fulfilling prediction is derived. For although it is true that any prediction that something will happen is confirmed by the occurrence of what is predicted - that is, the future-tense statement is true if what it says will happen, does happen, whatever the reason for this occurring - it does not follow from this that the theory from which the prediction was derived is thereby confirmed. For it may be that this theory does not in fact explain the occurrence of the predicted phenomenon, since what actually caused this to occur (in this case, the dissemination of the theory and/or its prediction) is not included in the premisses from which this prediction was derived. To illustrate this general point a different kind of example may be helpful. From a theory about the causes of cancer it may be predicted that someone will die within a certain period of time. The person does die, but from a road accident and not from cancer. Clearly, the truth of this prediction does not confirm the theory about cancer; nor would it be confirmed if the person died from an anxiety-related disease brought on by coming to hear of the prediction derived from the cancer-theory.

If these comments about the confirmatory significance of self-fulfilling predictions are correct, then a critical social theory whose predictions succeed as a result of their dissemination is not thereby confirmed. What might be instead confirmed is a different theory, namely one about the effects of the dissemination of social predictions. Now from the standpoint of this latter theory, all that matters is the consequences for people's actions of their coming to hold certain beliefs - such as the belief that what has been predicted from the original theory will indeed happen. It is irrelevant in this context whether this belief is true or false. That is, the theory about the consequences of prediction-dissemination, which may enable us to predict that if people believe a certain prediction is true they will act in ways that make it come true, is indifferent to the *truth* of that belief, being concerned only with the consequences of its being held.

Thus the possibility of explaining why (and predicting that) a disseminated prediction may come true does nothing to justify the theory from which that prediction was initially derived. Indeed, it seems that the reverse is generally the case. For if it turns out that a necessary condition for the predicted event's occurrence is that the prediction be disseminated, and affect people's actions, then the success of this prediction must suggest that its initial derivation was not justified, or that its premisses were not true.

It might be argued against this that it may be possible to develop a critical social theory which itself includes propositions specifying the effects of the dissemination of predictions, and that this would avoid the above objection, which depends upon a strict separation of the two theories. This, I think, has to be accepted. But such a theory would not then be able to be used in the way that Fay advocates for a critical social theory. It could only function successfully in what he would surely (and rightly) regard as a manipulative fashion. For this theory would depend for the success of its predictions on being able to predict the consequences of their dissemination. But it would not at the same time allow for the consequences of people not only being informed of these predictions, but also being informed of the basis upon which these have been made, since this would introduce another possible influence on their behaviour, a further factor which would not have been included in the theoretical premises from which these predictions had been derived. That is, dissemination of this theory (which includes statements about the consequences of dissemination of predictions) will alter the situation upon the consequences of which the original predictions were partly based. And it can be seen, I think, that any attempt to take account of this additional influence by yet further premisses must fail, if the additionally 'augmented' theory is itself to be disseminated.

Thus any theory which tries to take account of the social consequences of the dissemination of predictions can do so only by a strategy of at least partial concealment, of non-dissemination of (some of) its own assumptions. Yet it is a central feature of Fay's account of a critical social theory that those who develop it operate in a relationship of openness and honesty with their clients. Indeed, he argues that this is one of the crucial differences between critical and positivist social theory: the latter, he believes, implies a form of political practice in which social science is used by an élite of experts to control people's behaviour, in a way that excludes the latter's active participation in the process of change, through an open dialogue with theorists. But if my argument about the use of self-fulfilling predictions to confirm social theories is correct, this confirmatory role can only exist in precisely the kind of theorist-client relationship that Fay wishes to reject. Thus it seems either that self-fulfilling predictions are logically irrelevant to the truth or falsity of a social theory, or that they can only be relevant in theories which require for the success of their predictions a 'manipulative' attitude of

concealment on the part of theorists towards their clients.

There is one last point to make about Fay's account of critical social theory, which concerns the possibility of pessimism. He argues, as do many other critical theorists, that such a theory must show its clients a 'way out' of their present condition. It must discover the possibilities of change in the existing social structure, and must guide the activities of those who are frustrated, oppressed, and so on.³⁰ And it appears that this requirement is to be regarded as at least one element in the criteria of validity for such a theory. This seems to me epistemologically bizarre. For it would entail that a critical social theory cannot both be true and pessimistic in its diagnoses. That is, it would mean that any theory which failed to discover possibilities for change in the desired direction was thereby falsified. Optimism becomes an epistemological precondition of a correct social theory.

But this would rule out the possibility of social theories whose consequences suggest that certain otherwise desirable developments have little or no possibility of realization, at least within the reasonably foreseeable future. One might of course argue that theories which are designed to point the way to a better future, but which fail to indicate that and how this might be achieved, are of little value. But to suggest that just because of this, their analyses of the past or present are to be rejected as incorrect or inadequate, seems absurd and unnecessary. We can adopt views as to the most useful or beneficial aims and directions for theoretical work, without at the same time allowing the criteria of validity for the results of such work to be (partly) determined by the justifiability of these views. (This is related to my discussion of the place of values in social theory, in section 3 of chapter 2, where I argued that the adoption of normative standards of significance to determine the problems to which a social theory is directed should not affect the criteria of validity for assessing the content of that theory.)

Thus the epistemological character of critical social theory must allow the logical possibility of pessimism, the discovery that there are limits in the present or future to social transformations. As I noted in section 2 of the previous chapter, it is a feature of Freud's theory of the id that there are biologically determined, innate limits to human activity. I am not concerned here with whether this is in fact so, or whether his particular claims as to the nature of these limits are correct. Rather, I am saying that we should not adopt an epistemological account of critical social theory which logically excludes the truth of any such claims.

4. Theory, technique and theory of technique

At various points in this chapter I have emphasized the logical gaps between psychoanalytic theory and the success or failure of psychoanalytic therapy, so that the failure of therapeutic techniques is compatible with the truth of this theory, whilst the success of those techniques may provide little

support for it. I have argued that this is primarily due to the fact that in deriving predictions about therapeutic outcomes from psychoanalytic theory, a number of auxiliary statements must typically be assumed, whose own truth or falsity may display various degrees of independence from the explanatory claims made within this theory. Such auxiliaries may usefully be said to comprise a 'theory of technique': that is, an attempt to specify and explain the effects upon the patient of various elements of the therapeutic process. Thus even in those cases where predicted therapeutic success is achieved, it is possible that neither psychoanalytic theory nor its associated theory of technique are significantly supported, since it may be that this success is better explained by an alternative theory of technique.

I want now to explore some implications of these points for understanding the relations between 'theory' and 'practice' in psychotherapy, an exploration which will be continued in the following chapter. I will no longer confine myself to the discussion of psychoanalysis alone, but will consider other types of therapeutic approach such as behaviour modification, Reichian body-therapies, various kinds of group-therapy, and so on. My basic suggestion will be that there is a legitimate place for a considerable degree of eclectic pragmatism about therapeutic techniques, which partly divorces the assessment of their effectiveness from questions both about their own theoretical rationales, and about the adequacy of possible explanatory theories of the formation of the 'neuroses'.

And I now put this last term in scare-quotes, since I do not wish to imply that the concept of therapy should be limited to what is regarded within psychiatry as 'pathological' human activity. For one of the major developments in the field of psychotherapy in the past couple of decades has been the extension of therapeutic techniques to people whose 'problems' do not fall within the standard psychiatric classifications of either neurotic or psychotic disorders. In, for example, the consciousness-raising activities of women's and men's groups, and in the general phenomenon of the 'growth movement' and 'humanistic psychology', precisely this extension has occurred. Overall I think these have been desirable and fruitful developments, with important implications for traditional conceptions both of political activity and of the distinction between 'normal' and 'pathological' human action.³¹

Consider the following example. A man comes to notice that in his relationships with certain other men he experiences an unusual degree of anxiety which, in various ways, makes him unhappy or dissatisfied in his encounters with them. For instance, he may find that in situations of mutual disagreement he is unable to conduct himself as he would otherwise wish: he becomes angry too quickly, his mind goes blank at the crucial moment when he needs to respond to what has just been said to him, he becomes confused, feels panic-stricken, withdraws resentfully, or suchlike. He is puzzled as to why this happens, does not know how to deal with it, and gains neither insight nor ways

of improving such encounters from reflection, conversations with friends, and so on. So he becomes a member of some kind of therapeutic group to explore these difficulties. A lot of different things may then happen, some of which might prove helpful.

He might try 'acting out' a typical past encounter, with another member of the group whom he chooses, and being told to state at various points precisely what he is feeling. It might emerge that he suddenly fears that the other person is going to hit him. He is then asked if he can remember being hit by anyone in the past, and who that was. He remembers a childhood episode in which that occurred, involving his father. This episode is then itself acted out, with perhaps other members of the group taking the parts of his mother and sister. He re-experiences something of the fear he had then felt, and perhaps the resentment at not having been rescued or protected by the others. This resentment may 'feed back' into his feelings about other members of the group, in terms of his present relationships with them. He might then also become aware of further dissatisfactions he feels in his relationships with other adults: perhaps that they do not look after him sufficiently, that he is unsupported, vulnerable and so on.

I present this as a reasonably typical example of the beginnings of a therapeutic process. A certain kind of self-exploratory insight is perhaps achieved, in that present difficulties are related to past experiences, and connections between previously unrelated features of his life are suggested. There is nothing sacrosanct about the particular techniques that enabled this to happen: others might have done equally well, and indeed might have 'revealed' something quite different. For instance, he might instead have focused upon another feeling generated in the initial acting-out, say of humiliation or defiance, and have been instructed to follow through that feeling, aided by a particular breathing-technique, to see what emerged; or to engage in an 'internal dialogue' between the humiliated and defiant 'parts' of himself. In all these kinds of ways, some form of therapeutic insight may result.

However, the question I am most interested in about this example is: what is the relationship between the use of these therapeutic techniques, the explanation of this person's difficulties in relating to other men, and the successful resolution of those difficulties? My answer, roughly, is quite likely very little. More specifically I suggest that both of the following are possible: that as a result of this kind of therapeutic experience these difficulties either fade away or are significantly diminished despite the fact that the apparent source of those problems which emerged in that process does not correctly explain their origins; or, that although their origins were correctly identified, he finds that he is no better off when he takes part again in the previously problematic encounters. In other words, therapeutic experience may prove to be successful despite the explanations it 'reveals' being mistaken; or it may prove unsuccessful despite those explanations being correct. And in the former case, why it is that those techniques were effective may remain quite unknown; or their supposed

theoretical rationale may be incorrect. Thus therapeutic success, explanatory theories, and theories of technique, may display a high degree of logical independence from one another.

It will be helpful in pursuing this claim further, to consider the contrary view: that therapeutic success requires that the techniques which produce it are based upon a correct explanation of the origins of the person's problematic experiences. As we shall see, whether this view is accepted depends partly on how the success of therapeutic outcomes is defined; and this, I will argue, is itself partly a normative issue.

One way of defending this contrary view would be to claim that success in changing human activity ultimately depends upon the correct identification of its causal determinants. Although we may occasionally come across an effective technique without an adequate explanatory theory, this will always be a more or less random matter; and to do better than this, to have the best chance of arriving at a reliable technique, we must understand the actual sources of the problematic experiences. Thus, in the example considered, the effectiveness of therapeutic techniques must remain a hit or miss affair unless it is discovered whether, for instance, the pattern of anxiety is rooted in the past relationships of this man to his father, and whether this itself must be understood in terms, say, of Freud's theory of the Oedipal conflict.

Against this I suggest that it is often possible to develop quite reliable 'rule of thumb' techniques for change, which do not depend upon acquiring explanatory knowledge. For instance, suppose that in this example, the person reports a few months later that he is still experiencing precisely the same difficulties as before. It might then be suggested to him that he deals with these by training himself in certain anxiety-reducing techniques, such as meditational or breathing exercises, and that he uses these before or during his problematic encounters; or, that he rehearses different ways of dealing with these episodes by acting out preferable scenarios with other members of the group. He follows one of these suggestions, and finds that his anxiety disappears. The successful use of such techniques requires no explanatory theory, not even a theory as to why the technique works. (This would be similar to the present use of acupuncture in treating physical illnesses.) Furthermore, even if a correct explanation of why such situations provoked anxiety had been provided, such non-explanatorily based techniques might still prove necessary.

Two objections might be offered to this. First, it might be argued that the person's initial difficulties may be shown, for instance through a psychoanalytic explanation, to be only one specific manifestation of an 'underlying problem', so that techniques addressed purely to reducing the anxiety will leave unresolved all the other ways in which, say, the unsatisfactory outcome of the Oedipal conflict are or will be expressed. Otherwise, there may result simply a process of symptom-

substitution. Second, it might be said that the ‘successful’ use of the anxiety-reduction techniques cannot be regarded as a genuinely successful therapeutic outcome: to regard it as such misrepresents the proper goal of therapeutic processes. I will consider these replies in turn.

The first one, I believe, depends not upon any *general* claim that can be made about the relationships between explanatory theories, the success of techniques, and theories of technique, but upon the actual truth or falsity of various substantive explanatory theories, such as psychoanalysis itself. For instance, if it is in fact true that repressed unconscious feelings ‘live on’ within us, maintain their causal power, and can be re-directed in an indefinitely large number of ways (emerging as different ‘symptoms’), then certain kinds of therapeutic technique are very unlikely to succeed. But, of course, if this general prediction can be derived from psychoanalytic theory, we must accept the converse point: namely, that if it in fact turns out that these kinds of techniques are successful, this must cast doubt upon the truth of psychoanalytic theory.

So we cannot dismiss the possible success of techniques that are not based upon correct explanatory theories *a priori*. Rather, whether or not they are successful will be relevant (though only indirectly, for the reasons outlined earlier in the chapter) to the substantive truth or falsity of particular such theories. For instance, it might turn out that after a period in which the person in my example used his anxiety-reduction technique as a ‘prop’, he found that he could increasingly often do without it. The sources of the anxiety had, as it were, atrophied. If this were so, it would count against an explanation in terms of unconscious forces which, according to Freud, do not wither away.

I turn now to the second reply, that such a technique would not produce ‘genuine’ therapeutic success. This depends upon how such success is to be defined, as can be seen by considering the following possible definition of therapeutic success: that it should reduce the extent to which our activities are determined by influences which we do not recognize, and which reduce our capacity to exercise conscious control over our lives. Although this is only a rough outline of a possible goal of therapeutic processes, it is sufficiently specific to generate an interesting consequence: that, *by definition*, no technique can be successful unless it involves some form of self-understanding about the sources of one’s present discontents.

In other words, the goal of therapy is defined here in terms of an ideal of self-understanding, so that the sort of technique for dealing with anxiety mentioned above would be logically excluded from counting as ‘successful’ (though such a technique might still be useful to provide the ‘emotional breathing-space’ necessary for the process of self-understanding to develop satisfactorily). But it is important to note that this way of showing that technique must be tied to explanatory theories is quite different from another way in which this might be argued: namely, by claiming that *as a matter of fact*

it is only possible to achieve therapeutic success (defined this time, say, as ‘the removal of the distressful feelings’) if the persons concerned come to understand the sources of their difficulties. It is possible that this is so: but only if a particular kind of explanatory theory, and associated theory of technique, are actually true.

It must not be thought that there is anything suspect about the former definition of the goal of therapy as self-understanding. Rather, I want to emphasize the diversity of possible definitions, and that what counts as therapeutic success (and thus also what is acceptable as a therapeutic technique) is partly, and necessarily, determined by goal definitions. If, for instance, we adopt as the therapeutic goal the removal of some form of emotional distress, then what might be an ‘effective’ technique may differ from, say, techniques that are ‘effective’ in increasing people’s range of choices, or that make them ‘better adjusted to the demands of society’. That is, the effectiveness of a technique is determinable only in relation to a definition of therapeutic success. And it seems to me that this latter question is ultimately normative. Distress-reduction, more autonomy, better social adjustment, and so on: these are different and often competing therapeutic goals, and - if my arguments about values and social theory in chapter 2 are correct - they are not themselves scientifically resolvable. Further, it is I think clear that many of the debates about the relative efficacy of different therapeutic techniques - such as psychoanalysis versus behaviour modification - are due partly to concealed normative disagreements about therapeutic goals.

The issues raised by questions about psychotherapeutic goals, and their relations to other political and moral values, will be a central concern of the next chapter. But there is one final point to be made here, to avoid a possible misunderstanding of what I have suggested so far. In talking of the effectiveness of techniques being necessarily judged by reference to a normatively specified therapeutic goal, I am not endorsing the view that the means-ends relationship can be straightforwardly mapped onto the science-values dichotomy.³² For not only will the norms that specify a particular goal be relevant also in evaluating the legitimacy of the means for achieving it; but other norms may also be relevant to this evaluation. In the case of therapeutic technique the same is true.

For instance, even if our primary therapeutic goal is the relief of distressful feelings (such as anxiety or depression), it does not follow that the only relevant standard in evaluating techniques is their ‘effectiveness’ in achieving this. Suppose, say, that the use of electro-convulsive or chemo-therapy could be shown to relieve depression: we might still legitimately reject such techniques on other grounds, such as the absence of self-reflective engagement by the person treated, the power-relations involved in the therapist-patient relationship, and so on. It is entirely mistaken to claim, as some advocates of these techniques seem to, that objections on these grounds merely counterpose ‘value-

judgements' to effective practice'. For effectiveness cannot be specified independently of values, and the character of a therapeutic technique, as well as the goal to which it is said to contribute, must be subject to normative evaluation.

Notes and references

1 Jürgen Habermas, *Knowledge and Human Interests*, trans. J. J. Shapiro, Heinemann 1972, p.258.

2 Ibid., p.258.

3 See ch. 4, sec. 1, above.

4 Habermas, *Human Interests*, p.266.

5 E.g. Karl Popper, 'Science: Conjectures and Refutations', pp.33-8, in *Conjectures and Refutations*, 3rd edn, Routledge and Kegan Paul 1969, pp 33-65.

6 See Paul Kline, *Fact and Fantasy in Freudian Theory*, Methuen 1972, and H. Eysenck and G. D. Wilson, *The Experimental Study of Freudian Theories*, Methuen 1973.

7 See Popper, 'Science: Conjectures and Refutations'; for a succinct statement of his view of induction see also ch.1 of Karl Popper, *Objective Knowledge*, Oxford University Press 1972.

8 For an excellent account of these, see A. Chalmers, *What is this thing called Science?*, Open University Press 1976.

9 Both have been presented forcefully in H. Putnam, 'The "Corroboration" of Theories', in P. A. Schilpp, ed., *The Philosophy of Karl Popper*, Book I, Open Court 1974, pp. 221-240. See also Popper's response to this in 'Replies to my Critics', pp. 993-9 (in P. A. Schilpp, ed., *The Philosophy of Karl Popper*, Book II, Open Court 1974, pp. 961-1197), which involves I believe some misunderstanding of Putnam's claims.

10 For discussion of some of the issues mentioned here, see C. G. Hempel, *Philosophy of Natural Science*, Prentice-Hall 1966, chs. 2-4; and W. V. Quine and J. S. Ullian, *The Web of Belief*, Random House 1970, chs. 5-7.

11 See also Chalmers, *What is this Thing called Science?*, for a much fuller discussion of falsification.

12 This point is excellently argued in David Bohm, *Chance and Causality in Modern Physics*, Routledge and Kegan Paul 1957, chs. 1 and 5, in the context of a 'process' philosophy of nature, according to which the concept of a 'thing' involves abstraction.

13 Popper, 'Science: Conjectures and Refutations', pp.34-45. But Popper has a further criticism, on p.38, footnote 3, of the 'theory-loadedness' of clinical evidence, which I do not discuss.

14 Ibid., p.35.

15 See ch. 3, sec. 1, above.

16 S. Freud, 'Character and Anal Erotism' (1908), p.215, in *On Sexuality*, Penguin Books 1977, pp. 207-25.

17 Habermas, *Human Interests*, p.266.

18 On transference in psychoanalysis see, e.g., D. Brown and J. Pedder, *Introduction to Psychotherapy*, Tavistock Publications 1979, pp. 58-66. Freud discusses its central importance in, e.g. *Introductory Lectures* (1917), Penguin Books 1973, lecture 27.

19 Habermas, *Human Interests*, p.261.

20 Ibid., pp.261-2.

21 Ibid., p.269.

22 Ibid., pp.267-9.

23 Cf. Freud's discussion of the prospects for therapeutic success in his *New Introductory Lectures on Psychoanalysis* (1933), Penguin Books 1973, lecture 34.

24 Brian Fay, *Social Theory and Political Practice*, Allen and Unwin 1975, ch.5.

25 Ibid., p.108.

26 Ibid., p.102.

27 Ibid., pp.108-9.

28 Ibid., p.95.

29 Ibid., p.95, footnote 6; see also p.100, footnote 8.

30 Ibid., p.100.

31 Useful accounts of the many post-Freudian forms of psychotherapy are J. Kovel, *A Complete Guide to Therapy*, Penguin Books 1978, Part Two; and Brown and Pedder, *Introduction to Psychotherapy*, Part II.

32 See the discussion of this point in ch. 1, sec. 3, above.