

Psychoanalysis and Moral Theory

Joel Smith

University College London

MPhil, 2000

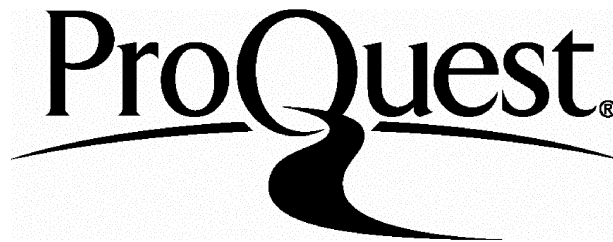
ProQuest Number: 10016054

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10016054

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

The aim of the paper is to assess the impact made by psychoanalysis on moral theory. In particular, I examine the claim that the psychoanalytic theory of the development of the super-ego has important ramifications in ethics. I argue that psychoanalytic theory is most reasonably seen as an extension of everyday 'commonsense' psychology. The psychoanalytic theory, of the development of the super-ego, offers an account of the development of moral sensibilities in the individual. This can be viewed as providing an account of moral motivation. Freudian and Kleinian psychoanalytic theories are examined and three accounts of moral motivation are identified: the 'oedipal model', the 'aggression model' and the 'Kleinian model'. I show that there are two main criteria for the acceptance of psychoanalytic theories: cogency of psychological explanation and independent empirical plausibility. The Kleinian model is found to be the most plausible of the psychoanalytic accounts. I claim that there is good evidence to suggest that the Kleinian model of moral motivation is true. Support for this claim is found, partly, in the fact that the Kleinian model cogently explains an impressive range of moral actions and phenomena. Also, it is shown that the Kleinian model receives support from experimental psychology, in particular the theories of Piaget and Kohlberg. I go on to claim that the Kleinian model of the development of moral sensibilities in the individual provides the basis of a plausibly sophisticated (genetic) naturalism in morals. Psychoanalytic naturalism is shown to be incompatible with rationalism in morals. This is followed by a discussion of the relationship between the Kleinian model and the realist/anti-realist debate in ethical theory. I finish with some thoughts on the place of moral psychology within the field of ethics and conclude that it should be awarded greater importance than has traditionally been the case.

Contents

One	Introduction	4
Two	Commonsense Psychology and Commonsense Moral Distinctions	5
Three	The Psychoanalytic Extension	11
Four	The Oedipal Model	16
Five	The Aggression Model	28
Six	The Kleinian Model	33
Seven	Choosing a Model	42
Eight	Relevance to Ethics	52
Nine	Guilt and Anxiety	53
Ten	Naturalism	58
Eleven	The Humean Account	64
Twelve	Non-Naturalism and Rationalism	69
Thirteen	Realism and Anti-Realism	73
Fourteen	Moral Psychology in Ethics	78
Bibliography		80

Psychoanalysis and Moral Theory

The most personal questions of truth. - ‘What am I really *doing*? And why am I doing it?’

Nietzsche, *Daybreak*, 196

One

Much of psychoanalytic theory might profitably be seen as an attempt to answer questions similar to those above. Freud’s writings overflow with explanations of actions and mental events (such as dreams, slips and symptoms). Among his driving principles was the desire to show us why it is that we do what we do. Freud is, perhaps, the figure who changed, most radically, our view of what sorts of motivations people have. His theories tell us, amongst other things, why we have the dreams we do, why we have the fears and emotions we do and why we perform the actions we do. I will be concerned, chiefly, with what psychoanalytic theory has to say about moral actions. This will involve distinguishing moral actions from non-moral actions; representing the psychoanalytic account(s) of moral motivation; clarifying the grounds upon which the plausibility of the psychoanalytic account(s) turn; and pointing out the extent to which the psychoanalytic account of moral motivation has ramifications in ethics.

The paper will fall into two parts, the first concentrating on the nature and standing of the psychoanalytic account(s) of moral motivation, the second focussing on the contribution that this account makes to moral philosophy. I will have succeeded in my aims if I answer these two questions:

- (1) What reasons do we have for supposing that the psychoanalytic account of moral motivation is true? This will include differentiating between competing accounts (i.e. Freudian and Kleinian¹) and finding reasons for accepting one over the other.
- (2) If correct, what consequences does the psychoanalytic account of moral motivation have for moral philosophy?

¹ I am limiting my discussion of psychoanalysis to the theories of Freud and Klein.

Of course, in answering these questions I shall be ranging over other areas and discussing the issues upon which my answers depend. These will include the nature of psychological explanation in general; the status of moral psychology within the field of ethics; and questions about naturalism and realism/anti-realism in moral theory.

Two

Question (1), above, can be split into two parts, neither of which can be answered without going some way towards answering the other. These two interlocking questions are:

(1a) What *is* the psychoanalytic account; *or* which psychoanalytic account is the most likely to be true?

(1b) What reasons are there for thinking this account correct; *or* what evidence is there to support it?

In answering these two questions it will be necessary to explain exactly what it is that I mean by ‘an account of moral motivation’. This requires a discussion of the nature of psychological explanations in general. Obviously, an account of what motivates us to perform moral actions must rely on a more general account of how we are motivated to perform actions *per se*. So, the aim of this section will be to give an account of psychological explanations of actions, then show how we can narrow our scope to those actions which are regarded as specifically moral.

It is not uncommon, thanks largely to the work of Richard Wollheim and Jim Hopkins, to view psychoanalysis as ‘an extension of commonsense psychology’. This is a view that I share and it has an important part to play in much of what follows. Of course, to say that psychoanalysis is ‘an extension of commonsense psychology’ is to say nothing until we know what is meant by ‘commonsense psychology’ (and what kind of an extension is involved). Commonsense psychology (sometimes called ‘folk-psychology’) is a clarification of the method we use, in everyday life, to understand the behaviour of others. Commonsense psychological explanations tell us why a

person performed a particular action. This is achieved through the attribution to that person of mental states such as beliefs and desires.

The account of commonsense psychology (and its psychoanalytic extension) that I am about to give is not original¹ but the use to which I shall put it (explaining specifically moral actions) is. Commonsense psychological explanations set out to explain behaviour in terms of reasons. By 'reasons' I am referring to mental states such as beliefs, desires, hopes, wishes, fears etc that make it rational for the agent to perform that action. It is through explanations by means of reasons that we understand the behaviour of others. We do this all the time and it provides us with a quick and reliable form of explanation. The structure of the simplest, 'one level', explanations is as follows: an action is explained by specifying the desire and the belief which jointly cause it². In explaining somebody's action I attribute to them a desire to achieve some end and a belief that the action in question is the best way (in the circumstances) of achieving that end. If the agent does, in fact, have this desire and corresponding belief, we say that the action is rationalised. For example, I desire to have a breath of fresh air. I believe that the best way of achieving this is to open the window. Therefore, I open the window. Thus, my action (opening the window) is seen to be rational in the light of my belief/desire structure. Clearly more detail than this can be filled in. For instance, the desire/belief pair above will cause me to desire to open the window. This desire, along with the belief that the best way of doing this is to turn the handle, causes me to turn the handle.

There are two ways in which the belief and the desire in the above example are connected to the action. The first is that they are connected in content. There is a transfer of content from the belief/desire pair to the action. This relation is mirrored in the language we use to describe the belief, desire and action. We describe a desire in terms of the conditions under which it would be satisfied, and we describe a belief in terms of the conditions under which it would be true. These conditions precisely are the successful completion of the specified action. The relation of meaning in the sentences used to describe the reasons and action maps the relation that holds between

¹ See, Davidson (1980), especially essay 1; Wollheim (1991), 'supplementary preface', and (1993d); Hopkins (1982), (1988), (1992a) and (1992b); Gardner (1993), section 8.4, and (1995). For sympathetic, but more sceptical, views see Cavell (1988) and (1993); and Hinshelwood (1995).

² For the classic argument in support of the claim that reasons are causes see Davidson (1980), essay 1.

the reasons (as causes) and the action (as effect). So, we can say that there is a hermeneutic relation between reasons and actions¹.

The second way in which the belief and the desire are connected to the action is causally. Together, the desire and the belief cause the action. This must be so if they are to be given as reasons for the action in psychological explanations. Reasons must cause actions. If they did not an agent might have a reason for performing an action, and might perform that action, but for entirely different reasons (or for no reason at all). But this would not be a good explanation of the action. Psychological explanations of actions *must* give the causes of those actions, otherwise they explain nothing.

Once we have recognised these two ways in which reasons are related to actions, we should be able to see that similarity in content between a desire/belief and an action is *good evidence* for the existence of a causal connection². If we observe an action and know that the agent holds a content related belief and has a content related desire, we have good reason to assume that the belief/desire pair caused the action. Of course the desire and belief must be related to the action in the right way, specified above.

This model can be extended once we notice that desires themselves can be explained in the same way. I gave an example of this above, explaining the desire to open the window by reference to the desire to have a breath of fresh air and the belief that opening the window is the best way of achieving this. In explaining a desire, I attribute to the agent another desire and a belief about how best to satisfy that desire. In the same way as reasons (desires and beliefs) cause actions and determine their content, reasons cause desires (or other mental states) and determine their content. It should be clear that each of my beliefs and desires will be related to countless other propositional attitudes (beliefs, desires, hopes, fears, wishes etc). We might then describe the resulting structure as a 'web' in which (to stick to the simplest

¹ See Hopkins (1992a) p. 92.

² Adolph Grünbaum denies that 'thematic affinity' vouches for a causal connection between reasons and actions, see Grünbaum (1984) p. 55. For a sound argument to the effect that thematic affinity is a mark of causal connection and therefore good evidence for such a connection, see Hopkins (1988).

belief/desire cases) beliefs and desires are connected via their causal roles and their content. Desires are constantly informed by beliefs, causing other desires and actions. Beliefs are constantly informed by reality (the external world or other mental states) and attempt to 'track' the truth. Therefore, in interpreting an action, we can go further and further back into the belief/desire structure of the agent, giving deeper and deeper explanations.

This, then, is the structure of commonsense psychological explanations. Actions, and desires, are explained by specifying the desires and beliefs that jointly caused them and determined their content. I will call the schema to which these explanations fit, the '*desire + belief* → *action* schema'. The question arises, however, as to how one might go about confirming or disconfirming these explanations. A point that Hopkins repeatedly makes is that commonsense explanations are not confirmed or disconfirmed in the same way as scientific explanations¹. Scientific explanations characterise physical events as instances of laws that can be experimentally tested. A scientific hypothesis is confirmed by the success of its predictions. One look at how we actually do tend to confirm our everyday psychological explanations will show how inappropriate the scientific model is in this area. We do not tend to confirm psychological explanations through their predictive power. One way in which this point has been made² is to say that there are no psychological laws. Whether or not this is the case, for our purposes the weaker claim, that we cannot make precise predictions about behaviour (for whatever reason), will suffice. How then do we tend to confirm our psychological explanations? What we actually do is to consider our explanations confirmed when they cohere with other explanations. We consider two explanations mutually confirmed when they postulate the same reasons (desires and beliefs), disconfirmed when they postulate contradictory reasons³. In other words, we crosscheck explanations against each other.

Here is a concrete example: An agent performs two actions, he takes a glass out of the cupboard, and he turns on the tap. In order to explain the first action I attribute to

¹ The assumption that psychological explanations must be confirmed in the same way as scientific explanations has led both Popper (1989) and Grünbaum (1984) to a misunderstanding of psychological and, in particular, psychoanalytic, explanations.

² See Davidson (1980), essays 7 and 11. Also see Cavell (1993), chapter 3.

³ See Hopkins (1982) p. xvii-xviii.

him the desire to open the cupboard, the desire to take out a glass etc. Further back in his chain of reasons, the desire to have a drink of water. In explaining the second action I also attribute to the agent the desire to have a drink of water. Thus, I can view the two explanations as mutually confirming each other. Explanations are disconfirmed in an obviously related way.

It is important to mention that the coherence or dissonance of beliefs and desires does not provide us with an infallible method for confirming or disconfirming commonsense psychological explanations. People can, and often do, have conflicting, inconsistent desires and beliefs. In commonsense explanations we are sometimes willing to attribute conflicting desires to an agent. Dissonance does not always disconfirm, likewise, coherence does not always confirm. This does not mean, however, that commonsense explanations are unsupported, or useless. On the contrary, commonsense explanations provide our most basic and pervasive method of understanding others as rational beings.

I come now to the second aim of this section, to narrow the scope of the discussion to psychological explanations of actions that are considered to be specifically moral. Since we are dealing with commonsense psychology, we require a commonsense distinction between moral and non-moral actions only. Any criterion that we give now as determining the distinction between moral actions and non-moral actions need only have intuitive plausibility. We should expect the distinction to be sharpened, however, once our account of psychology is deepened and extended, along psychoanalytic lines, in the following sections.

Having entered this caveat, I will make the distinction between moral actions and non-moral actions by recourse to the faculty of conscience. The distinction, stated baldly, is this: an action is moral when the reasons for it derive, more or less directly, from 'the conscience'¹. Conscience is necessarily bound up with the concept of duty. The demands of conscience are felt to be duties. Thus, a moral action is one which is done out of a sense of duty. I say that this is a bald statement of the distinction because it leaves much open. In particular I do not wish to use my distinction as a bias

¹ The term 'the conscience' is shorthand for 'the faculty of conscience'.

towards any particular conception of ethics. The two conceptions of morality, which I shall be discussing in the latter half of this paper (the rationalism exemplified by the Kantian tradition and the naturalism exemplified by the Humean tradition) are, I believe, equally served by this criterion of demarcation¹. The question of the nature of conscience, as, for instance, an instrument of reason or natural sentiment, is left open.

It will be instructive to give a concrete example of a moral action, and its reasons, in the terms of commonsense psychological explanation. We will take the example of an agent giving money to charity. The immediate reasons for this action are the desire to help the poor and the belief that giving money to charity is the best way of satisfying this desire. The explanation of this desire might be the desire to do one's duty² and the belief that it is a duty to help the poor. So, we can trace the agent's reasons back to a demand of conscience (what one feels to be one's duty) and the desire to act in accordance with that demand. If we are correct in attributing these reasons to the agent, then the action is moral; its motivation was, in part, a demand of conscience.

To clarify the distinction between moral actions and those that are not moral, we will look at another action that has the same description as the one just given, but which is not motivated by any demand of conscience. An agent gives money to charity. The immediate reasons for this action are the desire to impress his girlfriend and the belief that giving money to charity is the best way to do this. It seems clear that however far back we go into the belief/desire structure of this agent, we will not work our way back to a demand of conscience. This action is, then, not a moral one. I think this is in accordance with most intuitive beliefs.

¹ It might seem, at first glance, that this definition of moral action contains a bias in favour of rationalism. That this is not the case, however, should be seen when it is recognised that even a naturalistic account of moral motivation must do justice to the authoritative phenomenology of moral reasons. Moral actions are *felt* to be duties, and so any psychological explanation of a moral action must include reference to duties as reasons. Of course, the fact that I have a belief that it is my duty to do x, in no way presupposes an answer to the question as to whether that felt duty is objective or subjective, rational or otherwise. We might say of the agent that he has a relevant and distinctive 'ought-belief'. For this point about phenomenology see Deigh (1996) p. ix.

² This explanation looks as though it rules out rationalism (b) (see section twelve) by requiring that there be a desire to do one's duty. It does not, however, as a rationalist could hold that the desire is generated by reason, the belief that the action in question is a duty. It is not necessarily required that there be a *pre-existing* desire to do one's duty.

I have here simplified matters for ease of presentation. In fact, it is likely that the reasons for actions of this type will overdetermine the action. So, typically, the motivation for such actions will include reasons that are moral and reasons that are not. We can, however, keep the distinction between moral actions and non-moral actions sharp by stipulating that moral actions are those which have a demand of conscience as a necessary reason. Actions which would have been performed even in the absence of the demand of conscience, felt as a duty, are not strictly moral actions, i.e. explanations should support counterfactuals¹. This does not include actions, with the same description, that would have been performed for different reasons, reasons that did not figure as overdetermining reasons in the putative moral action, for these would be different actions.

Thus far I have given an account of our everyday method of giving psychological explanations of actions. We attribute reasons, most simply beliefs and desires, to the agent. I have shown how these commonsense psychological explanations are confirmed and disconfirmed, through mutual consonance and dissonance. I have also indicated how we can narrow the scope of the discussion to those actions which are regarded as specifically moral. This is done through recourse to the faculty of conscience and the concept of duty that is tied to it. With this done we can move on to a discussion of how psychoanalysis might be seen as an extension of commonsense psychology and, in particular, what psychoanalytic theory has to say about moral actions and the reasons for them.

Three

There are, broadly speaking, three main philosophical views of the nature of psychoanalysis: psychoanalysis as science, psychoanalysis as hermeneutics, and psychoanalysis as extended commonsense psychology. Which view one adheres to will, to a large extent, determine one's view of the truth-value of the theory. The scientific interpretations of psychoanalysis usually deny its truth; the hermeneutic and

¹ This may seem an overly stringent condition. I agree, and am open to suggestion on this point.

commonsense interpretations tend to have a more favourable view¹. To my mind, the view of psychoanalysis that does most justice to the theory is that which interprets it as an extension of commonsense psychology. This, then, is the only view I shall discuss².

In what way, then, does psychoanalysis extend commonsense psychology, as set out above? As with much in the philosophy of psychoanalysis, it is helpful to begin with the work of Richard Wollheim. Wollheim identifies three ways in which Freud extended commonsense psychology, he ‘deepened’ it, he ‘produced variations’ upon it, and he ‘contextualized’ it³.

What Wollheim means by Freud’s ‘deepening’ of the schema is the addition, to commonsense psychological explanations, of unconscious beliefs and desires. The nature of the explanations remains the same (a desire and belief cause and determine the content of an action), but the desires and beliefs cited as reasons for the action in question are unconscious. The schema is, thus, deepened to *unconscious desire + unconscious belief* → *action*. This, of course, is not uncontroversial and one of the most common philosophical reactions to psychoanalytic theory has been a rejection of the unconscious. When discussing the psychoanalytic concept of (the) unconscious it is often easy to fall into the trap of conflating the *descriptive* sense of ‘unconscious’ with the *dynamic* sense of ‘unconscious’⁴. Philosophical objections to the concept of unconscious mentality are usually directed towards the dynamic sense of ‘unconscious’. This deepening of the commonsense schema, however, requires only the descriptive sense of ‘unconscious’, for it attributes to the agent desires and beliefs that operate in the normal way but lack the property of consciousness. Other

¹ For psychoanalysis as science see Grünbaum (1984). For psychoanalysis as hermeneutics see Habermas (1971). For psychoanalysis as extended commonsense psychology see the list given in note 3, except Davidson (1980).

² For a detailed critique of the hermeneutic view see Grünbaum (1984), part 1. For critiques of the scientific interpretation of psychoanalysis see Wollheim (1993d); Hopkins (1988); and Gardner (1995). Sachs (1992) also criticises Grünbaum’s interpretation. Glymour (1982) attempts to answer the perennial question of the scientific testability of psychoanalysis.

³ Wollheim (1991) p. xxix.

⁴ For this distinction see Freud (1933a) p. 70-2, and (1923) p. 13-18; also see Laplanche and Pontalis (1973), the entry ‘Unconscious’. For a systematic treatment of the philosophical issues involved, and defence of both descriptive unconscious states and the dynamic unconscious, see Gardner (1992a). Also see Wollheim (1991), chapter 6.

psychoanalytic explanations do require the concept of the dynamic unconscious in order to become intelligible. This is the case particularly with instances of repression.

It is easy to see that this *is* an extension of commonsense psychology, but it is one which is firmly rooted in it. The justification of such an extension is the fact that commonsense psychology contains ‘gaps’. There exists action-like behaviour for which the agent is unaware of the reasons. Thus, if it adds to our understanding of the agent’s behaviour, and the reasons cohere in the right way with the rest of the agent’s mental life, we are justified in attributing, to him, reasons that are unconscious. Freud writes

The data of consciousness have a very large number of gaps in them; both in healthy and in sick people psychical acts often occur which can be explained only by presupposing other acts, of which, nevertheless, consciousness affords no evidence... All these conscious acts remain disconnected and unintelligible if we insist upon claiming that every mental act that occurs in us must also necessarily be experienced by us through consciousness; on the other hand, they fall into a demonstrable connection if we interpolate between them the unconscious acts which we have inferred. A gain in meaning is a perfectly justifiable ground for going beyond the limits of direct experience. (Freud 1915d) p. 166-7.

With this deepening of the commonsense schema, psychoanalysis extends the range of behaviour that we should view as intentional. The most common actions Freud explained in this way are parapraxes. These include forgettings, slips of the tongue, misreading, slips of the pen, bungled actions and mislayings. Psychoanalysis asks us to view these as intentional actions, the reasons for which are unconscious.

The second way in which Wollheim suggests that Freud extended commonsense psychological explanations is by ‘elaborating’ or ‘producing variations’ on the schema. Wollheim gives three examples of variations of the *desire + belief → action* schema. The first might be expressed as follows, *unconscious desire + unconscious belief (or phantasy) + chain of association → displaced action*. Cases such as this occur when the action that would have resulted from the desire and belief is inhibited in some way. A chain of association causes a substitute action, connected to the inhibited action through a chain of association, hence the term ‘displaced action’¹.

This extension of the schema introduces the notion of ‘phantasy’. This will require a brief digression on the nature of phantasy and the closely related concept of

¹ This term is Wollheim’s not Freud’s, see Wollheim (1991) p. xxxi.

‘instinct’¹. An instinct is a somatic force², which aims at the elimination of tension. The biological instincts have ‘psychical representations’ and these are phantasies. Phantasies are unconscious mental states and are the primary content of the id³. Phantasies represent a wish as satisfied. So, in the above extension of the schema, a phantasy (an unconscious representative of an instinct, which represents a wish as satisfied) can replace the belief.

The second variation can be expressed as follows, *unconscious desire + instinct* → *activity*. The term ‘activity’ here replaces ‘action’ to make the point that what is done is not rationalized as there is no instrumental belief, that belief having been replaced by an instinct. The mental processes that Freud explains in terms of this variation of the schema are, for the most part, mechanisms of defence, for instance repression.

The third, and final, variation is, *unconscious desire + precipitating factor* → *expression*. Such cases occur when someone does something without the instrumental belief (so it is not rationalised) and without the presence of a relevant instinct. Such things as dreams and neurotic symptoms form the bulk of what can be described as expressions. Expression is the furthest removed of the three variations from the original schema. It is debatable as to whether we should consider expressions as cases of someone’s *doing something* at all⁴.

These are the three ways in which psychoanalytic theory produces variations of the *desire + belief* → *action* schema. In addition to the extensions of commonsense psychology we have already considered, Freud also ‘contextualized’ it. He placed psychological explanations in the context of a developmental theory of the individual. It is helpful to see Freud as pursuing two interrelated projects. The first, explicated above, is to give cogent psychological explanations of actions, parapraxes, dreams, symptoms etc. The second is to give a developmental account of the mind of the individual. What it is important to see is how these two projects are related. The

¹ The following account of ‘instinct’ and ‘phantasy’ is broadly Kleinian in nature. For the classic statement of the following view of phantasies and their relation to instincts, see Isaacs (1952).

² Freud sometimes said of the instincts that they were themselves psychical, see, for instance, Freud (1905) pp. 82-4. However, the view expressed in what follows appears to be less problematic.

³ As the representatives of the instincts, and the primary content of unconscious states, phantasies underlie every mental process.

⁴ See Wollheim (1991) p. xxxv

developmental account gives us a structure of the mind. In Freud's later work this is the 'structural' account of id, ego and super-ego, alongside the account of the development through the phases (oral, anal, and genital). In the brief points made above, about the justification of the unconscious, psychological explanations require the structural account for their intelligibility. Similarly, with explanations involving phase specific desires, intelligibility is gained only in view of the developmental account. It is only against the background of the developmental account of the individual (in both its structural and phase aspects, which are, in any case, interrelated) that psychoanalytic explanations become plausible. For example, explanations involving sexual regression are plausible only in the presence of a specific account of the individual's development through various phases. Similarly, explanations involving the repression of instinctual urges are plausible only insofar as the distinction between the id and the ego is assumed. Of particular interest to us in this paper is the example of the super-ego. It will be seen that psychoanalytic explanations of moral actions are intelligible only by assuming the existence of the super-ego.

In this way, the developmental account lends support to psychoanalytic explanations; explanations of moral actions rely on the assumption of certain factors postulated in the developmental account. There is, however, a dialectical process at work: *the developmental account is supported by the cogency of the explanations it legitimises*. This mutual support of psychoanalytic explanations and the developmental account is not circular because, firstly, the developmental account must be independently plausible (and empirically supported) and, secondly, the psychoanalytic explanations must be consonant with our everyday, non-psychoanalytic commonsense psychological explanations¹. Therefore support for psychoanalysis comes both from within the theory (the relation of mutual support between the explanations and the developmental account) and from outside it (independent plausibility, empirical support, and psychological cogency/consonance).

¹ Of course, this is not to say that reasons attributed to individuals cannot contradict other reasons (e.g. conscious beliefs, desires etc) that the individual might have. They can. But attributing conflicting beliefs and desires to individuals is already part of ordinary, unextended commonsense psychology. We do not normally think of people as perfectly consistent or rational.

The area we are interested in is no exception to this general point. The explanations that give reasons for specifically moral actions are only intelligible against the background of the developmental account of the super-ego. According to psychoanalytic theory, conscience is a function of the super-ego. Thus, according to the distinction we made above, between moral and non-moral actions, a moral action is one, a necessary reason for which, derives from the super-ego (in its capacity as the agent of conscience). Explanations of moral actions thus depend of the developmental account of the super-ego. They must also be consonant with other psychological explanations attributable to the agent. The developmental account must itself be independently plausible, and empirically supported (or, at least, not at odds with empirical evidence). Also, the developmental account of the super-ego is supported to the extent that it legitimises cogent explanations of moral actions.

This is the form, without the content, of an answer to question (1b), ‘why should we think the psychoanalytic account correct?’. It is here, however, that we begin to see the urgency of answering question (1a), ‘which psychoanalytic account is the most likely to be true?’. For in order to see whether the psychoanalytic explanations are cogent, we must be working with a particular developmental account in mind¹. From what we have said so far, it is clear that the choice of developmental account will be constrained from two directions: independent plausibility/empirical support, and the cogency of the psychological explanations allowed. It is to this important question that I now turn, and this will involve us in our first detailed look at Freudian and Kleinian theories.

Four

Freud’s account of the development of the super-ego, which is at the same time an account of the development of the conscience, is presented in more than one work, and is not always consistent. This means that there is no simple answer to the question as to exactly what the Freudian account is. In fact, there are two significantly different

¹ True, we might evaluate psychoanalytic accounts by concentrating on the features they share (i.e. those features that make the theories distinctively psychoanalytic, as opposed to non-psychoanalytic). However, as we will see in section seven, the explanations of moral actions that the different psychoanalytic accounts legitimise differ enough to make the choice of developmental account pressing.

versions within Freudian theory, the ‘oedipal model’ and the ‘aggression model’¹. I shall treat these in turn before turning to the Kleinian account.

The notion of the super-ego has its roots in one of Freud’s most important papers ‘Mourning and Melancholia’, although the concept was not explicitly developed until later, the first extended discussions of the super-ego appearing in *Group Psychology and the Analysis of the Ego* and *The Ego and the Id*. Freud’s analysis of the condition of melancholia relied heavily on the concept of ‘introjection’. Specifically, the subject ‘introjects’ an object² that he feels he has lost. In introjection one phantasises a transportation of the object, and its inherent qualities, from the ‘outside’ to the ‘inside’ of the self³. Introjection is a method by which ‘identification’, the most important of all mental mechanisms, takes place⁴. Identification is a “Psychological process whereby the subject assimilates an aspect, property or attribute of the other and is transformed, wholly or partially, after the model the other provides. It is by means of a series of identifications that the personality is constituted and specified.” (Laplanche and Pontalis 1973) p. 205.

When a subject loses a loved object (usually a person), he unconsciously phantasises that it is ‘inside’ of him. These phantasies are rooted in the oral phase and find their prototype in the early phantasy “I should like to eat this”, or “I should like to take this into myself”. These early phantasies are of incorporating a good object. ‘Introjection’ is the name for the process by which the subject takes on the characteristics of those objects incorporated in phantasy. Freud originally relied on introjection to explain melancholia but he soon came to realise that it was a much more common phenomenon

¹ For a detailed discussion of the different views to be found within Freud’s work, see Deigh (1996), essay 4, although the distinctions Deigh draws do not coincide *exactly* with mine. As Deigh (1996) points out, on p. 76, Freud does not appear to have realised that he presented two significantly different accounts.

² The term ‘object’ has a rather specific meaning within psychoanalytic theory. For instance, Brenner says, “In psychoanalytic literature the term “object” is used to designate persons or things of the external environment which are psychologically significant to one’s psychic life, whether such “things” be animate or lifeless.” (Brenner 1955) p. 112. The idea that all objects are external becomes modified within Kleinian theory.

³ See Laplanche and Pontalis (1973), entry entitled ‘Introjection’.

⁴ In Freudian theory introjection is the most basic method of identification. However, in the work of Klein, and her followers, projection is increasingly important, the term ‘projective identification’ becoming commonplace. For an extensive discussion of this complicated and controversial concept see Hinshelwood (1991), entry entitled ‘Projective Identification’, also see Sandler (1988).

We succeeded in explaining the painful disorder of melancholia by supposing that an object which was lost has been set up again inside the ego...At that time, however, we did not appreciate the full significance of this process and did not know how common and how typical it is. Since then we have come to understand that this kind of substitution has a great share in determining the form taken by the ego and that it makes an essential contribution towards building up what is called its 'character'. (Freud 1923) p. 28.

So, by the time Freud came to formulate his account of the development of the super-ego, he viewed introjection as a ubiquitous process. Introjection is important for our purposes because it is absolutely fundamental to all accounts of the development of the super-ego.

The three texts upon from which I shall be drawing the oedipal model are Freud (1923) essay 3, (1924a), and (1933a), lecture 31. As my label suggests, the 'Oedipal account' begins from, what is itself, a controversial piece of Freudian theory, the Oedipus complex¹. It is one of the cornerstones of psychoanalytic theory that the child's most important object relations during the phallic phase constitute what is known as the 'Oedipus complex'. The, so called, 'positive' Oedipus complex is well known, the child wishes to have a sexual relationship with the parent of the opposite sex, and wishes to do away with the, rival, parent of the same sex. In harmony, however, with Freud's claim that the infant is originally bisexual, there is also a 'negative' Oedipus complex. In this second complex, the respective positions of the parents are swapped².

Freud claims that each individual's innate bisexuality will be weighted differently (favouring either masculinity or femininity)³. This, and environmental factors, will serve to make either the positive or the negative Oedipus complex the more powerful. Typically this is the positive Oedipus complex. In several places⁴, Freud called the super-ego the 'heir to the Oedipus complex'. The oedipal model is the story of how the Oedipus complex gives way to the super-ego.

In the normal development of the individual the Oedipus complex is overcome. There are two questions to be asked here, Why does this happen? And, How does this

¹ I will not attempt a justification of the universal character of the Oedipus complex here. For relevant material, see Freud (1916-17) lecture 21, and (1905) essay 3. See also, Brenner (1955) chapter 5, and Laplanche and Pontalis (1973) entry entitled 'Oedipus Complex'.

² See Brenner (1955) pp. 120-1.

³ Freud (1923) p. 34.

⁴ Freud (1923) p. 36; (1924a) p. 167; and (1933a) p. 64.

happen? The second question is the most important for our purposes, but I will briefly sketch an answer to the first. In 'The Dissolution of the Oedipus Complex' Freud gives a number of related motivations for the overcoming of the Oedipus complex. One is that the child realises his that his oedipal longings are hopeless, and gives them up. Another is that the complex is genetically programmed to die down at a certain point of development. The most important reason, however, is the fear of castration¹. The child fears castration at the hands of the father as retaliation for the positive complex, and castration is a necessary prerequisite for the satisfaction of the negative complex. This fear is strong enough to motivate a relinquishing of oedipal relationship

If the satisfaction of love in the field of the Oedipus complex is to cost the child his penis, a conflict is bound to arise between his narcissistic interest in that part of his body and the libidinal cathexis of his parental objects. In this conflict the first of these forces normally triumphs: the child's ego turns away from the Oedipus complex. (Freud 1924b) p. 176.

So, for several reasons, the fear of castration being the most important, the Oedipus complex must be surmounted. We now need an answer to the question of *how* this is done. The answer, in short, is that the parents, as lost loved objects, are introjected and set up within the ego as the super-ego. Thus, the super-ego is the heir to the Oedipus complex. It is worth quoting Freud at some length here, as this is a crucial point

We have attributed the function of conscience to the super-ego and we have recognised the consciousness of guilt as an expression of a tension between the ego and the super-ego. The ego reacts with feelings of anxiety (conscience anxiety) to the perception that it has not come up to the demands made by its ideal, the super-ego. What we want to know is how the super-ego has come to play this demanding role and why the ego, in the case of a difference with its ideal, should have to be afraid. We have said that the function of the ego is to unite and to reconcile the claims of the three agencies which it serves; and we may add that in doing so it possesses in the super-ego a model which it can strive to follow. For this super-ego is as much a representative of the id as of the external world. It came into being through the introjection into the ego of the first objects of the id's libidinal impulses – namely, the two parents. In this process the relation to those objects was desexualized; it was diverted from its direct sexual aims. Only in this way was it possible for the Oedipus complex to be surmounted. The super-ego retained several features of the introjected persons – their strength, their severity, their inclination to supervise and to

¹ The part played by the fear of castration given, below applies, only to boys. For the different account of the resolution of the Oedipus complex in girls, see Freud (1925b). Briefly, the fact of the anatomical difference between the sexes results in the relative instability of the female super-ego. "The incentive for abandoning the oedipal position and for establishing the superego is weaker in women because they have already been deprived of the penis. Consequently they remain more on the developmental level of the wish to be loved and the fear of loss of love. Their internalizations are less stable, and they are more tied to the immediate give-and-take of real relations" (Schafer 1960) p. 176.

punish...it is easily conceivable that, thanks to the diffusion of instinct which occurs along with this introduction into the ego, the severity was increased. The super-ego – the conscience at work in the ego – may then become harsh, cruel and inexorable against the ego which is in its charge. Kant's Categorical Imperative is thus the direct heir to the Oedipus complex. (Freud 1924a) p. 166-7.

This passage is incredibly dense, and it covers a lot of material. Essentially it is a summary of the oedipal model and from it we can extract everything we require for an understanding of that account. The first sentence is conceptual. The super-ego is, in part, defined as that part of the mental apparatus which has the function of conscience¹. Similarly a person feels guilty when they fail to meet up to the demands set by their conscience. Freud says of the super-ego that it plays both a demanding role and is an ideal. This point will become more important later during the discussion of the ego ideal. For the moment, it is enough to say that the super-ego is not just a prohibiting agent. Prohibition and the installation of guilt feelings are important, but do not cover the full range of the super-ego's functions².

The main claim in this passage is that the super-ego is the result of an introjection of the parents. This claim is fundamental to all accounts of the development of the super-ego. As we have seen, the super-ego is a structure in the mind defined, partly, by its function as the agent of conscience. For the account to be plausible, the super-ego must have the features that we would normally ascribe to the conscience, and it must be reasonable to suppose that these features would naturally result from an introjection of the parents. The features of conscience that I will consider are these: (i) the conscience is a distinct agency in the mind, acting, seemingly, independently of the agent, (ii) it is intimately bound up with certain 'moral' emotions, such as guilt and pride³, (iii) the conscience is felt to have authority over its possessor, issuing

¹ Defining the mental apparatus functionally is controversial. Wollheim thinks that Freud's move to the structural theory signifies a move away from functional definitions, see Wollheim (1991) chapter 6. On the other hand, Gillman (1982) stresses the usefulness of 'considering the super-ego as a functional concept'. I suggest that it might be instructive to view both function and developmental history as necessary in defining the super-ego. As such, the super-ego is, in part, defined by its role as the agent of conscience. Indeed this appears to coincide with something Wollheim himself says in *The Thread of Life*, "The superego is...singled out...by two criteria: by its history, and by the role that it plays in the life of the person." (Wollheim 1984) p. 200.

² "These functions include (1) the approval or disapproval of actions and wishes on the grounds of rectitude, (2) critical self-observation, (3) self-punishment, (4) the demand for reparation or repentance of wrongdoing, and (5) self-praise or self-love as a reward for virtuous or desirable thoughts and actions." (Brenner 1955) pp. 126-7. Schafer (1960) gives a broad ranging account of the 'loving and beloved' aspects of the super-ego.

³ I will simply assume that guilt, pride are emotions. There are many other moral emotions, e.g. shame, indignation and resentment. Guilt and pride are, however, typical. The question of the nature of emotions is addressed in Gardner (1992b).

demands, (iv) in punishing transgressions, with guilt feelings etc, the conscience can often seem disproportionately harsh. Unless the oedipal model can explain these features of conscience, we shall have to reject it¹.

Why, or how, does the super-ego appear as a distinct agency in the mind? Clearly Freud accepted that the super-ego does appear as a separate agency in the mind, and this can be summed up in the phrase ‘the super-ego *confronts* the other contents of the ego’. We will, however, have to look in several places to answer the question. Freud must explain how it is that the super-ego, qua introjected parents, is felt to be an independent entity within the ego. This matter is particularly pressing as Freud clearly thought that introjection could and, in virtually every other case, did occur without this result. In other words, since most introjected objects simply become part of the ego, and help form its character, why does this not occur with the super-ego?

Freud suggests that the reason that the super-ego is set up as a distinct agency in the mind is that the renunciation of the Oedipus complex is the *first*, and also the *most intense*, instance of the loss of a loved object

The super-ego owes its special position in the ego, or in relation to the ego, to a factor which must be considered from two sides: on the one hand it was the first identification and one which took place while the ego was still feeble, and on the other hand it is the heir to the Oedipus complex and has thus introduced the most momentous objects into the ego. (Freud 1923) p. 48.²

This points us in the right direction, but it is inadequate. The fact that the introjection of, and identification with, the parents is the first and most emotionally charged of such processes does not explain why the super-ego should be felt as acting *independently* of the ego. Why shouldn't the introjected parents rather cause a massive change in the ego as a whole, forming part of the ego's character? There must be something more in the oedipal model if it is to explain the *otherness* of the super-ego. This ‘something more’ is emphasised by Jennifer Church (1992). The super-ego *confronts* the ego because it is comprised primarily of desires, felt as demands *about* other desires, beliefs etc³. This is a result of the fact that the introjected parents act to

¹ This and the following two sections will involve taking each of the oedipal model, ‘aggressive model’ and ‘Kleinian model’ in turn, in order to see whether each can account for these features.

² See also, Freud (1933a) p. 64.

³ The ego ideal is also felt to be opposed to the ego. This is due to its highly idealised nature, it feels just as *other* as the super-ego, in its prohibitive role. The reason for the high degree of idealisation present is the fact that the parents, before they are introjected, are seen as almost omnipotent. I discuss the ego ideal below.

restrict the expression of the child's drives and desires. With the internalization of these restrictions, they are bound to seem opposed to the ego.

The terminology is slippery here. The distinctions between 'introjection', 'identification', 'incorporation' and 'internalization' are subtle, and not altogether sharp¹. Clearly, the parents are introjected, on the basis of a phantasy of incorporation. Since 'internalization' is a process dealing with relations (as opposed to objects), it is the relation between parents and child that is 'internalized'. Should we say that the parents are identified with? Freud uses the term 'identification' when speaking of the formation of the super-ego, for instance in the quotation above. And, following the definition of 'identification' given, by Laplanche and Pontalis, above we should call this process an identification with the parents. But, it looks as though we must distinguish between two kinds of identification. First, an identification where the object is subsumed within the ego² and helps form the subject's character. Second, an identification where the object is set up as a distinct agency within the ego, and is felt as alien to it (as is the case with super-ego formation). We might call them 'central identification' and 'peripheral identification', respectively³. With this distinction made, we can move on to the next feature of conscience that requires explanation.

In what way is the super-ego bound up with moral emotions such as guilt and pride? I shall begin by discussing Freud's account of the arrival of a sense of guilt in the individual. Guilt, and the feeling of guilt, is moral. Someone who is susceptible to feelings of guilt has a moral sense of some kind. Guilt is intimately associated with the conscience; the term 'guilty conscience' testifies to this. We can plausibly say that the emotion guilt is caused by the conscience. Since Freud thinks that conscience is a function of the super-ego, and the super-ego only comes into being with the dissolution of the Oedipus complex, we should expect Freud to claim that before this development, children are immune to guilt. This is exactly what he does

¹ To pick up on some of the difficulties here, see the entries on these concepts in Laplanche and Pontalis (1973).

² Or, possibly, later, the super-ego itself. It is controversial whether later identifications (such as with teachers, heroes etc) are introjected into the super-ego, or into the ego proper.

³ This distinction is similar, though not identical to, Wollheim's distinction between 'identification' and 'mere internalization', see Wollheim (1984). Something similar to Wollheim's distinction will be canvassed when I come to discuss Klein's account of the super-ego. It should be made clear that this distinction between 'peripheral' and 'central' identification is not a sharp one, there will be various kinds of identification ranging from 'completely peripheral' to 'completely central'.

A great change takes place only when the authority is internalized through the establishment of a super-ego. The phenomena of conscience then reach a higher stage. Actually it is not until now that we should speak of conscience or a sense of guilt. (Freud 1930) p. 125.¹

Freud thinks that moral emotions arrive with the super-ego, the agent of conscience. Why, and how, does this occur? To answer this question it will be necessary to look at the situation the child is in prior to the setting up of the super-ego. Conscience is an internal authority, issuing demands, and guilt is felt when these demands are not satisfied. Before the super-ego is formed, there are no such internal demands; prohibitions on behaviour come from the external world, specifically the parents. These prohibitions are usually directed against the desires and instinctual urges of the child, and fear (realistic anxiety²) is the child's response

The part which is later taken on by the super-ego is played to begin with by an external power, by parental authority. Parental influence governs the child by offering proofs of love and by threatening punishments which are signs to the child of loss of love and are bound to be feared on their own account. This realistic anxiety is the precursor of the later moral anxiety. So long as it is dominant there is no need to talk of a super-ego and of a conscience. It is only subsequently that the secondary situation develops (which we are all too ready to regard as the normal one), where the external restraint is internalised and the super-ego takes the place of the parental agency and observes, directs and threatens the ego in exactly the same way as earlier the parents did with the child. (Freud 1933a) p. 62.

So, guilt is produced when the ego fails to live up to the internal demands of the super-ego. The production of guilt is the super-ego's method of punishing the failure of the ego to meet its demands. These demands are a transformation (through internalisation) of the external demands of the parents. The analogue of guilt feeling in the prior situation of external parental authority, is the fear of the loss of love.

Moral pride is a more complex matter, as it introduces the notion of the ego ideal. I pointed out above that there is more to the super-ego than its prohibitive and punitive function. There is also what Schafer calls the 'the loving and beloved super-ego'³. I shall call this other aspect of the super-ego, the ego ideal⁴. The idea is that the

¹ This quotation is actually taken from the main text from which I shall draw the aggression model. However, Freud never deviated from this point of view.

² Throughout this essay I will not refer to Freud's first theory of anxiety, only the second, which is much more satisfactory. For the notion of 'realistic anxiety' and the distinction between the two theories of anxiety, see Freud (1926); Laplanche and Pontalis (1973), entry entitled 'Realistic Anxiety'; and, Rycroft (1995), entry entitled 'Anxiety'.

³ See Schafer (1960).

⁴ The notion of the ego ideal is notoriously controversial and unsettled. Freud's own use of the term is inconsistent. For an excellent summary of the various interpretations of the meaning of the term 'ego ideal', see Laplanche and Pontalis (1973) entry entitled 'ego ideal'. There they point out that "It is

super-ego, as well as reproducing the harsh, prohibiting characteristics of the parents, also takes over their role as objects of admiration, and the ability to dispense gratifying feelings. This is the ego ideal. Some passages of Freud's appear to tell against this interpretation

The super-ego seems to have made a one-sided choice and to have picked out only the parents' strictness and severity, their prohibiting and punitive function, whereas their loving care seems not to have been taken over and maintained. (Freud 1933a) p. 62.

However, Freud was inconsistent on the point of whether the super-ego, or ego ideal, had a loving, non-punitive, aspect. Freud had, six years previous to the above quotation, attributed the function of humour to the super-ego

If it is really the super-ego which, in humour, speaks such kindly words of comfort to the intimidated ego, this will teach us that we have still a great deal to learn about the nature of the super-ego...if the super-ego tries, by means of humour, to console the ego and protect it from suffering, this does not contradict its origin in the parental agency. (Freud 1927b) p. 166.

Schafer points out that Freud had good reasons for emphasising the harsh characteristics of the super-ego¹, but this does not mean that he did not recognise the ego ideal (as I have characterised it). In fact, it is through the loving aspect of the super-ego, the ego ideal, that we can give an account of the moral emotion pride. Freud doesn't give a detailed account of pride, as he does with guilt. It is reasonably simple, however, to extrapolate from what we know of the oedipal model. We can give an account that fits with the model's generation of the super-ego. The super-ego arises, as the introjected parents, on the event of the relinquishing of the child's most intense object relationship, the Oedipus complex. Part of the super-ego corresponds to the harsh, prohibitive aspect of the parental agency, part of it corresponds to their idealised and loved aspect. This ego ideal is seen as an ideal which the ego should strive to become. If the ego succeeds in becoming, in some way, like the ego ideal, the ego will feel pride. Pride is felt when the subject lives up to the demands of the super-ego, whether these be the demands of the harsh super-ego or those of the ego ideal.

difficult to discern any hard and fast meaning of the term 'ego ideal' in Freud's writings" p. 144. My, admittedly vague, interpretation of the ego ideal as an 'aspect of the super-ego' leaves open the question of whether it refers to an independent agency situated within the super-ego. The discussion of Klein, below, should clear up most ambiguities in this area.

¹ See Schafer (1960) pp. 169-171.

With guilt, we found a pre-super-ego analogue in the external demands of the parents. There is a similar analogue in the case of pride. This is the demonstration of love given by the parent to the child when he behaves in accordance with the parents' demands. Before the super-ego is formed, the child longs not only to possess, sexually, the parents, but also to be like them¹. With the ego ideal installed within the mind, this is transformed into a longing to be like the ego ideal. It is from success in this department that the moral feeling of pride springs.

Why is the conscience felt to have an authority over the subject? The explanation of this feature of conscience should be reasonably clear in light of what we have said already about the oedipal model. The authority of the conscience, the authority of the super-ego over the ego, is derived from the authority of the parents over the child. The external authority of the parental agency is transformed into the internal authority of the super-ego

As the child was once under a compulsion to obey its parents, so the ego submits to the categorical imperative of its super-ego. (Freud 1923) p. 48.

The very young child idealises his parents and parental rules are seen as universally binding, as categorical. Also, it is likely that the norms of the parents are the only norms that the very young child comes across, and so there is little chance for him to question their universal authority². It is the internalization of these categorical demands that provides the super-ego, and hence the conscience, with the absolute authority it appears to wield.

Why is it that in punishing transgressions with feelings of guilt, the super-ego can often appear disproportionately harsh? There are many possible answers to this question and I will not be able to exhaust them here. Freud's interest in the super-ego centred on its harsh, sometimes pathological nature, as found in certain mental disorders. This is to be expected from a man whose theories about normal mental functioning often find their root in clinical and pathological material. The disproportionately harsh nature of the super-ego, however, is not confined to

¹ Indeed, the longing to be like the parents, and the resulting pre-super-ego identifications, are a vital aspect of the Oedipus complex and pave the way for the installation of the super-ego.

² Also relevant here is Freud's suggestion (see below) that the child introjects not the parents themselves but their super-egos, see Freud (1933a) p. 67. This would make it easier for a strict, authoritative super-ego to perpetuate itself. It also goes some way towards explaining the similarity between different subjects' super-egos.

pathological subjects; it is a common feature of experience. I will look at the three main reasons the oedipal model suggests might account for this phenomenon: (1) the super-ego can operate unconsciously, (2) the parents' super-egos, not the parent, are introjected, and (3) the child feels there to be a reciprocal aggressive relationship between himself and the parents.

Why does Freud think that the super-ego can operate unconsciously? In what way does this explain the disproportionately harsh nature of its punishments? Part of the answer to the first question is that there is a plausible answer to the second. If we can show that the postulation of unconscious super-ego operation helps to explain an observable phenomena, this is evidence in its favour. As this is the case, I shall take the second question to begin with.

If the super-ego sometimes operates unconsciously, it follows that we will sometimes be unaware of the demands of our own conscience. This means that we will also be unaware that we have transgressed those demands. If we transgress an unconscious demand of conscience, we will be punished with feelings of guilt, even though we are unaware that we have done anything wrong. As J. A. C. Brown points out, this explains the fact that we often feel guilt after having performed an action that reason tells us is not immoral¹. This explanation is itself evidence for the claim that the super-ego sometimes operates unconsciously, but Freud did give independent reasons for this claim, in particular, that the super-ego is responsible for resistance. After asking the question as to which part of the mind unconscious resistance² arises from, Freud says

The resistance can only be a manifestation of the ego, which originally put the repression into force and now wishes to maintain it...Since we have come to assume a special agency in the ego, the super-ego, which represents demands of a restrictive and rejecting character, we may say that repression is the work of this super-ego and that it is carried out either by itself or by the ego in obedience to its orders. If then we are met by the case of the resistance in analysis not being conscious to the patient, this means that in quite important situations the super-ego and the ego can operate unconsciously... (Freud 1933a) p. 68-9.

The second reason for the disproportionately harsh nature of the super-ego is Freud's claim that it is the super-ego of the parent which is introjected. Although this

¹ Brown (1961) p. 29.

² Unconscious resistance in analysis is the patient's resistance to the attempt to make his repressed unconscious wishes conscious to him.

does not explain the harshness of the super-ego (only pushing the problem back to the harshness of the parental super-ego), it does (partly) explain why a child with lenient parents can develop a harsh super-ego. Freud writes

As a rule parents and authorities analogous to them follow the precepts of their own super-egos in educating children. Whatever understanding their ego may have come to with their super-ego, they are severe and exacting in educating children... Thus a child's super-ego is in fact constructed on the model not of its parents but of its parents' super-ego. (Freud 1933a) p. 67.

The third reason for the harshness of the super-ego is the reciprocal aggressive relationship the child feels he has with the parent. The oedipal situation from which the super-ego arises is essentially one of ambivalence. If we take both positive and negative complexes into account, each parent is, on the one hand, loved and, on the other hand, despised as a rival for the love of the other parent. In relinquishing of the Oedipus complex, murderous wishes towards the parents are repressed but find displaced expression in the harsh punishments of the super-ego on the ego

There is no doubt that, when the super-ego was first instituted, in equipping that agency use was made of the piece of the child's aggressiveness towards his parents for which he was unable to effect a discharge outwards on account of his erotic fixation as well as of external difficulties; and for that reason the severity of the super-ego need not simply correspond to the strictness of the upbringing. (Freud 1933a) p. 109.

The aggressive relationship is felt to be reciprocal, to a large extent, because the child projects its aggression onto the parent and, as a result, feels the parent to be aggressive towards him. Thus the parents are experienced as being harsher than they in fact are. This, of course, adds to the disproportionately harsh nature of the introjected parental agency¹. The success of the account will be gauged later. For now, it is sufficient to say that the model cannot instantly be dismissed for being unable to account for the salient features of conscience. Our next step is to set out the aggression model, show how it differs from the oedipal model, and see how well it accounts for the same four features of conscience.

¹ This point is emphasised in the 'Kleinian model'. In fact, it is slightly controversial as to whether Freud held this view concerning the projection of aggression (in either the oedipal or the aggression model). For the view that he did see Wollheim (1991) p. 197, for the opposite view see Deigh (1996) p. 74, n. 18. Even if Freud did accept the view, it plays such a small part in his theories that the question is of no great consequence. The view makes no essential difference to the plausibility of the oedipal or the aggression model.

Five

The account of the origin of the super-ego which I am calling the ‘aggression model’ is contained within the last two chapters of *Civilization and its Discontents*. The first thing that one notices, reading this text, is that reference to the Oedipus complex, so central to the oedipal model, is virtually non-existent. As reflected in my terminology, the central part played by the Oedipus complex is replaced by the child’s aggressive instincts. It might be suggested that this is only a change in emphasis rather than a substantial change in the theory. The oedipal model does contain reference to the aggressive relationship between the parents and the child, and the change in emphasis might be explained by Freud’s different point of departure into the area¹. In the following discussion I hope it will become clear that the aggression model does present us with a significantly different theory and not just a shift of emphasis.

By the time of writing *Civilization and its Discontents* Freud had introduced the concept of the death instinct². Whether or not the concept of the death instinct is acceptable, the attribution of aggressive instincts to the individual seems almost irrefutable and, in practise, the difference (between an aggressive instinct, and a pure death instinct) is minimal. The aggression model relies heavily on this attribution and we shall take it as an assumption. In answer to the question of how the individual’s natural aggression is curbed, Freud says that

His aggressiveness is introjected, internalized; it is, in point of fact, sent back to where it came from – that is, it is directed towards his own ego. There it is taken over by a portion of the ego, which sets itself over against the rest of the ego as super-ego, and which now, in the form of ‘conscience’, is ready to put into action against the ego the same harsh aggressiveness that the ego would have liked to satisfy upon other, extraneous individuals. (Freud 1930) p. 123.

Why is this aggression internalized and redirected towards the individual’s own ego? The answer to this question requires us to look at the child’s early situation. He has certain instinctual urges, both libidinal and aggressive, and it is the role of the parent to curb the more socially unacceptable of these. The parents, then, represent an external restriction on the satisfaction of the child’s urges and desires. This causes the

¹ Freud’s point of departure for the oedipal model is the need to account for paranoid delusions of being observed, Freud (1933a) p. 59-60. The point of departure for the aggression model is the need to explain how civilization inhibits the natural aggressiveness in man, Freud (1930) p. 123-4.

² Freud (1920). Also see Laplanche and Pontalis (1973), entry entitled ‘death instincts’.

child to resent the parents and powerful aggressive instincts are turned toward that agency. However, the child loves the parents, as protectors and providers. This ambivalent attitude towards the parents is unstable, in a way that tends to escalate the aggressive aspect. The frustration of the child's desires angers the child, but this anger cannot be vented on the parents, because they are loved and also because they prohibit such expression. This further denial of instinctual urges heightens the child's aggression, and so on in a way that is viciously circular. Freud explains how this situation is resolved

He finds his way out of this economically difficult situation with the help of familiar mechanisms. By means of identification he takes the unattackable authority into himself. The authority now turns into his super-ego and enters into possession of all the aggressiveness which a child would have liked to exercise against it...If this is correct, we may assert truly that in the beginning conscience arises through the suppression of an aggressive impulse... (Freud 1930) p. 129-30.

The aggression model then, explains the erection of the super-ego by reference to an ambivalent attitude the child finds himself in, regarding his parents. In displacing his aggressive impulses back towards himself, the child gains the instinctual satisfaction that his parents denied him. Two questions must be answered in regard to this aggression model: In what ways is it different from the oedipal model? Can it explain the four features of conscience listed above?

There are, to my mind, two important differences between the oedipal model and the aggression model. The first is that while the oedipal model regards the setting up of the super-ego as the result of the suppression of both sexual and aggressive instincts (the relinquishing of the Oedipus complex), the aggression model relied solely on the suppression of aggressive instincts. One might think that this suggests that the oedipal model is a general account, and the aggression model an elaboration of the role played by aggression in that account. That this is not the case is shown by the second difference between the two models. Namely, that while, in the oedipal model, the aggression towards the parents stems from rivalry, in the aggression model it stems from being the subject of their authority. Although this does not make the two models contradictory, each is alone sufficient to explain the setting up of the super-ego, i.e. if we assume the truth of the oedipal model, the aggression model becomes inessential, and vice versa.

Can the aggression model explain the salient features of conscience? Conscience is a distinct agency in the mind, acting independently of the ego. The oedipal model explained this phenomenon by citing the fact that the Oedipus complex was the first and most intense object relationship of the child, the resulting introjection of the parents thus taking a special place in the ego. Added to this is the fact that the demands of the introjected parents are bound to appear external to the ego, as the ego is their object. The harsh treatment of the ego by the super-ego explains its *otherness*. This last point can also clearly be made concerning the aggression model, the first, however, will either have to be modified or abandoned, as there is no reference to the Oedipus complex in the aggression model. I think that we can safely modify it by changing all reference to 'the Oedipus complex' to 'emotionally ambivalent situation'. The Oedipus complex, to be sure, is also an emotionally ambivalent situation, but the ambivalence arises from erotic love and rivalry, the ambivalence presupposed in the aggression model arises from love (probably with erotic elements) and revenge for the prohibition of instinctual satisfaction. This 'emotionally ambivalent situation' leads directly to the erection of the super-ego; it is the first major introjection and it is clearly the result of an intense object relationship. Thus, with this slight modification, the reasons for the independence of the super-ego are *essentially* the same in both models.

Conscience is intimately bound up with moral emotions such as guilt and pride. The oedipal model claimed that the production of guilt is the super-ego's method of punishing the failure of the ego to meet its demands. This is an internalization of the prior situation in which the child feels fear of the loss of love when it fails to keep to the parents' external moral prohibitions. I see no reason why this explanation cannot apply equally as well to the aggression model. In fact, Freud explicitly gives this explanation in *Civilization and its Discontents*

At the beginning...what is bad is whatever causes one to be threatened with loss of love... This state of mind is called a 'bad conscience'; but actually it does not deserve this name, for at this stage the sense of guilt is clearly only a fear of loss of love... In small children it can never be anything else... A great change takes place only when the authority is internalized through the establishment of a super-ego. The phenomena of conscience then reach a higher stage. Actually it is not until now that we should speak of conscience or a sense of guilt. (Freud 1930) p. 124-5.

In the case of pride, the oedipal model claimed that it is felt when the ego lives up to the demands and ideals set by the ego ideal and super-ego, respectively. However,

the setting up of the ego ideal was explained with reference to the relinquishing of the Oedipus complex. Obviously, if we are to save the explanation we will have to modify it, so that reference to the Oedipus complex is unnecessary. So, the question must be, Does the aggression model provide an explanation of the installation of the ego ideal within the mental apparatus? According to the model, the super-ego is formed as a reaction to an unbearable situation in which an escalating aggression towards the parents can't find satisfaction, partly because the parents prohibit aggression, and partly because the child loves the parents. The installation of the super-ego allows for the discharge of aggression, back towards the ego. Since the super-ego is modelled on the parental agency, it is reasonable to suppose that some of the love felt for the real parents will be displaced onto the new mental structure. This love will be directed onto the part of the super-ego that corresponds to the loved aspects of the parents, the ego ideal.

So it is plausible to claim that, on the aggression model, an ego ideal would be erected, enabling us to give an explanation of the moral emotion pride. However, we should note that the setting up of the ego ideal does not follow so easily from the aggression model as it does from the oedipal model. The aggression model describes the setting up of the super-ego purely in terms of the satisfaction of aggressive instincts. The genesis of the ego ideal is only indirectly explained, by the assumption that the whole parent, including loved and loving aspects, is introjected¹. On the other hand, the oedipal model explains the generation of the ego ideal directly, as the parental agency is introjected *qua lost loved object*. This worry aside, we can conclude that the aggression model can explain the close relationship between conscience and the moral emotions of guilt and pride.

The conscience is felt to have authority over its possessor, issuing demands. The aggression model accounts for this feature of conscience in exactly the same way as did the oedipal model. The authority wielded by the parents over the child is taken over by the newly installed super-ego. The rules of the parents are felt to be ultimate and universally binding due to the particular circumstances of childhood. The

¹ Deigh (1996) p. 77-8, has made the complimentary claim that the oedipal model (which he calls the 'standard account') only explains the authority of the super-ego indirectly, 'per accidens', since it is accidental that the objects of the Oedipus complex are the child's authorities. This seems to be a fair point. In section seven I deal fully with the implications of such claims.

internalization of these standards explains the authority of the demands made by the super-ego, and hence the conscience.

In dealing out punishment for transgressions, the conscience often seems to be excessively harsh. In the case of the oedipal model, I outlined three explanations for this phenomenon: that the super-ego can operate unconsciously, that it is the parents' super-egos that are introjected, and, that there is an aggressive relationship between the child and the parents. The first two explanations hold equally for the aggression model. The third, however, requires some discussion. The reason for this is that, in the aggression model, it is the aggressive relationship existing between the parents and the child that is the precipitating factor of the erection of the super-ego, and that the nature of this aggressive relationship is different that in the oedipal model.

According to the aggression model, the *raison d'être* of the super-ego is the satisfaction of aggressive urges. This is not the case for the oedipal model¹. This means that the disproportionately harsh treatment of the ego by the super-ego is, according to the aggression model, the normal state of affairs. The aggression model explains the harsh aspect of the super-ego directly and the loving aspect indirectly, and the oedipal model vice versa. In this case we should not be surprised to find the super-ego formed in line with the aggression model harsher than that formed in line with the oedipal model.

The nature of the aggressive relationship that holds between the parents and the child in the aggression model differs from that in the oedipal model. In the oedipal model the child harbours murderous wishes towards the parents because they are seen as rivals in his erotic life. This means that the self-punishment inflicted by the super-ego will have a necessarily erotic element. This erotic element is not present in, or essential to, the aggression model. The aggressive wishes the child has towards the parents are the result of the suppression of instincts. True, the child has sexual instincts which are suppressed by the parents but, Freud contends, it is only the suppression of aggressive urges that leads to the formation of the super-ego and they supply its power. Freud explicitly rejects the idea that *all* instinctual renunciation gives rise to guilt feelings

¹ The oedipal model holds that the *Raison d'être* of the super-ego is the replacement of the lost loved object with an internal substitute.

In the most recent analytic literature a predilection is shown for the idea that any kind of frustration, any thwarted instinctual satisfaction, results, or may result, in a heightening of the sense of guilt. A great theoretical simplification will, I think, be achieved if we regard this as applying only to the *aggressive* instincts, and little will be found to contradict this assumption. (Freud 1930) p. 138.

This means that, in the aggression model, there is no immediate connection between the punishment dealt out by the super-ego and the sexual instincts. However, this point, important as it will be later, may be put aside for the moment. We can still see that the aggression model can explain the super-ego's tendency to punish the ego in an overly harsh manner; the final feature of conscience is accounted for. At this point we will leave the aggression model, and Freud, altogether, to delve into the theories of Melanie Klein.

Six

Although Kleinian psychoanalytic theory differs from the orthodox Freudian position(s), it is firmly rooted in Freud's theory, and can reasonably be seen as a development of it. One of the most important reasons for Klein's move away from Freud was her observation of guilt feelings in very young children

In the course of my analysis of small children, as I began to get a direct knowledge of the foundations upon which the super-ego was built, I came upon certain facts which seemed to allow of an enlargement in some directions of Freud's theory on this subject. There could be no doubt that a super-ego had been in full operation for some time in my small patients of between two and three quarters and four years of age, whereas according to the accepted view the super-ego would not begin to be activated until the Oedipus complex had died down – *i.e.* until about the fifth year of life. Furthermore, my data showed that this early super-ego was immeasurably harsher and more cruel than that of the older child or adult, and that it literally crushed down the feeble ego of the small child. (Klein 1933) p. 248.¹

Klein attempted to square these findings with Freudian theory by claiming that the mature super-ego develops with the demise of the Oedipus complex, but an earlier super-ego comes into existence at the beginning of the long process of 'working through' the Oedipus complex. However, Klein moved further from Freud's views and eventually came to see the erection of the super-ego as occurring *before* the Oedipus complex. I will run through the Kleinian model of the development of the

¹ Here Klein refers to the oedipal model, but the point applies also to the aggression model as Freud never changed his views about the age at which the super-ego develops.

super-ego. First, however, we must characterise the early condition of the infant, the so-called 'paranoid-schizoid position'¹.

Klein accepted Freud's theory of the death instinct (the self-directed aggressive instinct) and suggested that the first action of the ego was to deflect its death instinct outwards. There are three channels through which this aggression flows: some cannot be deflected and remains directed towards the ego; some is transformed into aggressive wishes towards external objects: and, some is projected onto those external objects and is felt as external aggression directed back towards the infant. The aggression that cannot be deflected, and remains directed towards the ego, is the origin of the aggressive power of the early super-ego

The danger of being destroyed by this instinct of aggression sets up, I think, an excessive tension in the ego, which is felt by it as an anxiety, so that it is faced at the very beginning of its development with the task of mobilizing libido against its death instinct...This apparently earliest measure of defence on the part of the ego constitutes, I think, the foundation stone of the development of the super-ego. (Klein 1933) p.250.

The result of the fact that much of the infant's aggression *is* deflected is that it feels itself to be in a very hostile environment indeed

he perceives his anxiety arising from his aggressive instincts as fear of an external object, both because he has made that object their outward goal, and because he has projected them onto it so that they seem to be initiated against himself from that quarter. He thus displaces the source of his anxiety outwards and turns his objects into dangerous ones; but, ultimately, the danger belongs to his own aggressive instincts. (Klein 1933) pp. 250-1.

This very early situation the child finds himself in (a hostile external environment, whilst also facing aggression from the inside) Klein calls the 'paranoid-schizoid position', in which the pervasive emotion is paranoid anxiety; fear of persecution from external and internal objects². There is, however, another feature of this position that is vital to Klein's understanding of super-ego formation. This is the fact that external objects are experienced as 'part objects', which are 'split' into good and bad objects. For instance, the infant's mother will be formed of many objects; the most important of which is the breast. The infant is, as yet, incapable of comprehending

¹ For discussions of this concept, see Klein (1946); Segal (1978) ch. 3; and, Hinshelwood (1991), entry entitled 'Paranoid-Schizoid Position'.

² See, Hinshelwood (1991) entry entitled 'Internal Objects'. The notion of internal objects is often confusing. Segal makes it clear that "internal objects are not 'objects' situated in the body or the psyche: like Freud, Melanie Klein is describing unconscious phantasies which people have about what they contain." (Segal 1978) p. 12.

that the mother is a single entity, and will relate to parts of her independently according to their differing functions and characteristics. In addition to this, the breast itself will be split into 'good/nourishing breast' and 'bad/persecutory breast'. To the child there are two separate breasts¹. The good breast being that object which satisfies the child, the bad breast being that which does not. Klein writes

The object world of the child in the first two or three months of its life could be described as consisting of hostile and persecuting, or else of gratifying parts and portions of the real world. (Klein 1935) p. 285.

Given the paranoid-schizoid position, how is the super-ego formed? I have said that the portion of aggressive instinct remaining directed towards the ego forms the basis of the super-ego, but this is not yet to explain the genesis of that entity. All we have at the moment is a self-destructive instinct within the child; it has not yet inherited the functions of the super-ego and would not yet be able to account for the features of conscience that we have been considering. This is because we have not yet broached the subject of the introjection of the parental agency. Since the infant's object relations are with part objects, we would expect him to introject these parts, as opposed to the whole objects and this is, indeed, the case. In fact, on the Kleinian model, there are many instances of introjection forming not a single, unitary, super-ego but a group of 'internal objects'. In both Freudian models, there occurs one momentous introjection of the parental agency, installing the super-ego. In the paranoid-schizoid position, however, the parental agency is split into many part objects, which are all introjected individually. On both Freudian models the child phantasises an internal object, the super-ego, which has certain relations to the ego. On the Kleinian model the child phantasises many internal objects, the most important being the good breast and the persecutory breast, each of which bears certain relations to the ego and to other internal objects. Thus Kleinians speak of 'a community of internal objects'.

Freud viewed introjection as a response to the loss of a loved object. Klein accepts this but also extends the range of situations in which an external object is introjected. Both good and bad objects are introjected, but for different reasons. Bad objects are

¹ In fact, since the child has not yet developed the cognitive ability to perceive each object as enduring over time, the breast is encountered as a different object each time it is encountered. Thus, for the infant, there will be countless breasts.

introjected in oral, cannibalistic phantasies. The bad, threatening breast is devoured in phantasy and becomes a member of the world of internal objects. Klein writes

the relation to the first object implies its introjection and projection...These processes participate in the building up of the ego and super-ego...From the beginning the destructive impulse is turned against the object and is first expressed in phantasised oral-sadistic attacks on the mother's breast, which soon develop into onslaughts on her body by all sadistic means. (Klein 1946) p. 2.

The internalized persecuting objects take possession of the remaining internally directed aggressive instinct and become the source of the harsh super-ego. However, the child also introjects good objects. This is done as a defence against the bad objects that the child feels he already contains. These bad objects are not just those recently introjected but also include unpleasant bodily sensations. For instance, a feeling of hunger is experienced as a bad internal object causing pain. Thus the introjection of good objects serves to protect the child from his bad internal objects. These good introjected objects form the basis of the loving aspect of the super-ego. So, bad introjected objects form the basis of the super-ego and good introjected objects form the basis of the ego ideal.

This is, however, a primitive super-ego. On both Freudian models the super-ego is formed at a time when the child can recognise particular parental demands and prohibitions. Thus the demands and prohibitions of the super-ego can be seen as (reasonably directly) descending from these. It is difficult to say, however, in what way the good or bad breast can make *demands* on the child. The bad introjected objects are simply experienced as violently attacking the ego; the good objects experienced as helping and protecting the ego¹. This is what I mean by saying that the super-ego formed in the paranoid-schizoid position is primitive. Klein owes us an account of how this early super-ego develops in to something which we can readily recognise as the agent of conscience. This she does with her characterisation of the 'depressive position'².

The paranoid-schizoid position gives way to the depressive position when the child develops the concept of a whole object. The child now comes to realise that the bad breast that it has attacked is actually the same object as the good breast that it

¹ Due, however, to the ascendance of the child's sadism at this point, the super-ego is experienced as almost exclusively harsh, as opposed to loving.

² See Hinshelwood (1991) entry entitled 'Depressive Position'.

loves. The child now experiences depressive anxiety, which manifests itself in the realisation that he has damaged that which he loves, the fear of a loss of its love, and the desire to make reparation. There is a corresponding change in the super-ego. In the paranoid-schizoid position the child is threatened with violence as punishment for its aggressive excesses

The child then dreads a punishment corresponding to the offence: the super-ego becomes something which bites, devours and cuts. (Klein 1928) p. 187.

The change to the depressive position alters this primitive super-ego, the only real function of which is to punish aggressive acts and phantasies. The onset of depressive anxiety causes the infant to lessen the ferocity of his attacks on the mother

The more the child's sadism is lessened, the more the influence of its unreal and frightening imagos recedes into the background, since they are offshoots of its own aggressive tendencies...there emerge beneficent and helpful imagos, based upon its fixations, in the oral sucking stage, on its generous and kindly mother, which approximate more closely to the real objects; and its super-ego, from being a threatening, despotic force issuing senseless and self-contradictory commands which the ego is totally unable to satisfy, begins to exert a milder and more persuasive rule and to make requirements which are capable of being fulfilled. In fact, it becomes transformed into conscience in the true sense of the word. (Klein 1933) p. 252.

Why, with the beginning of the depressive position, does the child's sadism decrease? One reason is that now the child sees that it has been attacking its loved object, fear of the loss of that object is heightened. This causes the child to lessen its sadism. Wollheim (1984) pp. 211-3, suggests that there is another reason why the child will 'prefer love over hate' and lessen its sadism. He suggests that there is an innate intolerance of aggression, testified to by the fact that aggression is frequently expelled and projected. This is why the child's perception of whole objects tends to lessen his aggression. The ambivalence toward the whole mother is unbearable and either love or hate must take precedence. With this account of how the early super-ego is formed and subsequently develops, we can turn to the question of whether the Kleinian model can account for our four features of conscience. This discussion will illuminate and amplify the basic Kleinian model I have given.

Can the Kleinian model account for the fact that the conscience is felt to be a distinct agency in the mind? In the paranoid-schizoid position, the introjected objects are either phantastically good or phantastically bad. It is this that explains the fact that they are peripherally, as opposed to centrally, identified with. The excessively harsh objects are felt to oppose the ego, and thus feel *other* to the ego. The helpful objects

feel *other* to the ego because of their highly idealised character¹. We meet with a problem, however, with the transition to the depressive position. According to Klein, an important part of this transition is the introjection of good objects and a fuller identification with them

As the ego becomes more fully organised, the internal imagos will approximate more closely to reality and the ego will identify itself more fully with 'good' objects. The dread of persecution, which was at first felt on the ego's account, now relates to the good object as well and from now on preservation of the good object is regarded as synonymous with survival of the ego. (Klein 1935) p. 264.

This transition would seem to rob the super-ego of its otherness and thus deprive the Kleinian model of its ability to account for that particular feature of conscience. I think, however, that this only seems to be the case. The problem is avoided when we note that although, in a sense, there is a transition from the 'paranoid super-ego' to the 'depressive super-ego', it is also the case that the phantastically harsh internal objects that constitute the early 'paranoid super-ego' remain unaltered. They do not disappear, they simply lose their dominance

I am led to believe from the analysis of children that their super-ego is a highly resistant product, at heart unalterable (Klein 1927a) p. 155.²

The peripherally identified-with internal objects that form the early super-ego remain peripheral. The central identification with the good object that is important for the transition to the depressive position must be based on another introjection. It is not the case that the peripherally identified-with object somehow becomes centrally identified-with (and becomes less phantastic). It is more properly seen as a transition of power from peripheral objects to central ones. This view is in line with Klein's assertions concerning the multitude of super-ego identifications

The analysis of little children reveals the structure of the super-ego as built up of identifications dating from very different periods and strata in the mental life. (Klein 1928) p. 187.

So the super-ego retains its otherness because it still very much contains early, peripherally identified-with, internal objects³. In the depressive position, however, the most violent objects lose much of their influence, this being transferred to the more

¹ See the descriptions of the 'fairy-mama and papa' in (Klein 1929) p. 203; and (Klein 1927a) p. 157.

² In a much later paper (Klein 1958) Klein suggest that the most terrifying of the early internal objects are permanently retained in the 'deeper layers of the unconscious'.

³ In fact, these provide fixation points for possible regressions.

realistic, intermediate, objects, and to the centrally identified-with good objects. This means that the super-ego becomes less harsh, less phantastic, less powerful and less alien to the child. However, this does not occur to such a great extent that the phenomena of super-ego as independent agent within the mind is lost. This can be put in another way by saying that the paranoid-schizoid position is never *fully* worked through. If it were, the super-ego would cease to be differentiated from the ego and the demands of conscience would simply be what the agent desired¹.

In what way is the super-ego bound up with moral emotions such as guilt and pride? In both Freudian models, guilt was explained as punishment of the ego by the super-ego for failure to meet with its demands. Its prototype was seen as the fear of the loss of love. In the Kleinian model, however, we have distinguished two phases of super-ego development: paranoid and depressive. As a result, the Kleinian explanation of guilt is more complex. Klein distinguishes between persecutory anxiety, felt in the paranoid-schizoid position, and depressive anxiety, felt in the depressive position. Persecutory anxiety is related to the protection of the ego from aggressive objects, internal and external. Depressive anxiety is related to the protection of good objects, internal and external, against damage caused to them mainly from the child himself. Persecutory anxiety is experienced as fear of both external and internal objects and cannot count as guilt. Guilt relies on an ambivalent attitude towards an object, the child harbours aggressive wishes towards one whom he loves. But this is not possible in the paranoid-schizoid position, objects are cognised as either completely good or completely bad, the necessary ambivalence does not exist. Thus, the moral emotion of guilt can only come into being with the onset of the depressive position². In the depressive position the child realises that the object he

¹ Of course, I would not want to claim on conceptual grounds (although it might be empirically true) that this never occurs in an individual.

² There appears to be a tension in Wollheim's interpretation of Klein's position. He agrees that guilt can only be felt with the onset of depressive anxiety. He also says that with the arrival at the depressive position the super-ego becomes less harsh and more closely identified with the good object (he calls the developed super-ego the 'ego ideal'). However, he also claims that we should think of "shame as standing to the ego ideal as in the same way as guilt does to the super ego." (Wollheim 1984) p. 220. This last claim suggests that, contrary to his previous statements, guilt *is* felt in the paranoid-schizoid position, i.e. when the demands of the (paranoid-schizoid) super-ego are not satisfied, guilt is felt, and when the demands of the (depressive) ego ideal are not satisfied, shame is felt. Thus, guilt is brought into existence with the onset of depressive anxiety, but it also exists prior to this, in the paranoid-schizoid position. I avoid this problem by discarding the claim that guilt is transformed into shame with the development of the super-ego beyond its persecutory state. The claim that guilt, and pride, only develop with the advent of the depressive position is perfectly in line with Klein's statement, quoted above, that conscience only really develops at this point.

hates is the object he loves and this ambivalence makes it possible for the super-ego to punish the ego with feelings recognisable as guilt. Depressive anxiety *is*, partly, guilt¹. So guilt is the super-ego's punishment of the ego for failing to keep to its demands. The precursor of guilt is persecutory anxiety, the fear of persecution from both internal and external objects. This anxiety is transformed into guilt with the infant's ability to perceive whole objects and the ensuing ambivalence².

What of pride? Klein says very little about the emotion of pride, so we will have to try to work out what she *would* have said. Both Freudian models explained pride as a transformation of the feeling of being loved. When the ego lives up to the demands of the super-ego, and comes to closely resemble the ego-ideal, the ego feels pride. The explanation is essentially similar on the Kleinian model. As is the case with guilt, pride will only be felt by the child once it has reached the depressive position. The reason for this is the phantastic nature of the primitive super-ego. The bad objects of the early super-ego are excessively aggressive and violent, the good objects utterly perfect. This means that the child will have no chance of either living up to the demands of the super-ego or coming to resemble the ego ideal. It is only with the transition to the more realistic super-ego of the depressive position that the infant will gain the capacity to feel pride.

How does the Kleinian model explain the fact that the conscience is felt to have authority over the subject, issuing demands? This feature of conscience is a direct result of the fact that the infant is utterly dependent on his parents. The introjected objects are felt to have an authority over the ego because their external counterparts have such authority. The super-ego takes its authority from the authority of the parents

As we know, the parents are the source of the super-ego, in that their commands, prohibitions and so on become absorbed by the child itself. (Klein 1927b) p. 179.

The super-ego issues demands but these only really become intelligible and realistic with the onset of the depressive position. The demands of the primitive super-

¹ It also, importantly, includes the desire to make reparation.

² Because the transition from paranoid-schizoid position to depressive position is gradual and has many intermediate stages, it follows that there will be intermediate feelings between persecutory anxiety and guilt. Thus there will be a sliding scale with persecutory anxiety at one end and depressive anxiety (which includes guilt) at the other.

ego are phantastic in character and generally consist of incredibly aggressive and forceful restrictions on instinctual impulses. Again, this leads us to the conclusion that although it is correct to regard this structure as an early super-ego, it is not plausible to attribute to it the function of conscience. Conscience arrives with the depressive position.

How does the model explain the, often excessive, harshness of conscience? Two of the answers given by the Freudian models to this question also hold for the Kleinian model, but they are of diminished importance. Parts of the super-ego are unconscious¹ making its demands appear disproportionately severe. Also, in bringing up children, parents are likely to act in accord with their own super-egos, making the introjected parent appear harsher than they might otherwise. It is clear, however, that, on the Kleinian model, by far the most important reason for the excessive cruelty of the super-ego is the child's own sadism. In the first few months of life, during which the paranoid-schizoid position predominates, the child's aggressive instincts run rife. I have already described how the child projects aggression on to the mother and attacks her in a way that only serves to escalate his hostility. As a result of this², most of the objects that the child introjects are utterly ferocious. The early super-ego's harsh character is thereby explained.

But we have said that the onset of the depressive position weakens the intense severity of the super-ego. Does this not tell against our claim that the super-ego is often disproportionately harsh? I don't think so. It was noted above that the early, phantastically bad, super-ego constituents are not erased from the mind. Every individual will, due to the pressures of reality, particularly during moral and emotional crises, be prone to regress to an earlier stage; the paranoid-schizoid position. Thus, episodes of over-severe punishment by the super-ego can be seen as the work of the degraded, but still operative, primitive super-ego. The conclusion being that, on the Kleinian model, the excessive harshness of the super-ego is explained by the child's own aggression, the viciously circular way in which this aggression increases and his propensity to project that aggression onto external objects.

¹ In fact it is the particularly harsh internal objects that remain 'in the deepest layers of the unconscious'. See Klein (1958) p. 241.

² And also of the fact that the child splits his objects into 'good' and 'bad'.

Seven

The oedipal model claims that the super-ego originates in the child's relinquishing of the Oedipus complex. The aggression model claims that the super-ego originates in the child's inability to express his aggressive instinctual impulses. The Kleinian model claims that an early super-ego originates in the infant's cannibalistic phantasies, and that it becomes the super-ego proper when the child enters the depressive position. We can now ask the question: which, if any, of these models should we accept?

As I made clear in section three, there are two directions from which our choice of developmental model is constrained. Firstly, there is the cogency of the psychological explanations the model allows. Explanations of moral actions, desires and emotions will differ according to which developmental model we are working with. As a result, we will tend to favour that model which allows the most cogent and coherent explanations. Secondly, the developmental model must be independently plausible and be empirically supported. The closer a model approximates to the empirical findings of developmental psychology, the more we will tend to favour it. We can, then, choose between our three competing models using these two criteria.

One of the main differences between the psychological explanations given on the oedipal model and those given on the aggression model, concerns reference to the sexual instincts. In the Oedipus complex, the child has erotic designs on both parents, and each parent is also seen as a rival for the love of the other. Since, according to the oedipal model, the super-ego is 'the heir to the Oedipus complex', it replaces both the aggressive and the erotic elements of that complex. The aggression model, on the other hand, sees the installation of the super-ego as the result of the suppression of the aggressive instincts alone. This difference has the result that the oedipal model can better account for 'the sexualization of morality'.

In his discussion of 'moral masochism'¹ Freud gives an explanation of those people whose ego's seek punishment from their super-ego. This arises, he claims, from the infantile wish to be punished (beaten) by the father (for whom the super-ego is the internal substitute). He writes

¹ Freud (1924a) pp. 165-170.

We know that the wish, which so frequently appears in phantasies, to be beaten by the father stands very close to the other wish, to have a passive (feminine) sexual relation with him and is only a regressive distortion of it. If we insert this explanation into the content of moral masochism, its hidden meaning becomes clear to us. Conscience and morality have arisen through the overcoming, the desexualization, of the Oedipus complex; but through moral masochism morality becomes sexualized once more, the Oedipus complex is revived and the way is opened for a regression from morality to the Oedipus complex. (Freud 1924a) p. 169.

This gives, I think, a reasonably cogent explanation of the phenomenon of finding sensual pleasure in being punished by one's conscience. Since this explanation relies on a regression from the super-ego back to the Oedipus complex, the aggression model cannot allow it. The explanation depends on factors that are essential to the oedipal model. True, the aggression model does allow that the child has unrequited sexual wishes for the parents, but these play no *direct* role in the genesis of the super-ego and so the explanation of moral masochism would not 'fall out' of the aggression model so neatly.

This last point raises the issue of 'direct' and 'indirect' explanations. A model explains a phenomenon directly when that phenomenon is entailed by an essential part (or parts) of that model. A model indirectly explains a phenomenon when that phenomenon is entailed by a contingent, or periphery, aspect of that model. In section three I pointed out that the aggression model can only indirectly explain the erection of the ego ideal¹, and the oedipal model can only indirectly explain the authority of the super-ego. Can this fact help us to decide between the two models? We will assume that a model which directly explains a phenomenon is preferable to one which explains that phenomena only indirectly. Moreover, it is more important to explain (directly) some phenomena than it is to explain (directly) others. It is my contention that the authority of conscience (explained directly by the aggression model but only indirectly by the oedipal model) is 'more important' than both the ego ideal and the sexualization of morality (both explained directly by the oedipal model but only indirectly by the aggression model). This is so, I claim, due to the pervasiveness and remarkable nature of the authority of conscience.

How is the fact that a model explains a phenomenon (such as the authority of conscience) directly, related to the cogency of the psychological explanations it legitimises? For a psychological explanation to be cogent, we must be able to

¹ And, as we have just seen, 'the sexualization of morality'.

crosscheck it against other explanations. It is also better for a large number of actions to be explained by a small number of base mental states, than vice versa. Thus, if a model directly explains a phenomenon (it results from the *essential* elements of that model), this second criterion is more likely to be fulfilled. This, in turn, means that the first criterion is more likely to be fulfilled, since if a large number of explanations refer to the same base states, it is conducive to their cohering with each other. This leads me to suggest that the aggression model is more favourable than the oedipal model on the grounds of cogency of psychological explanation. It is the fact that it directly explains the authority of conscience, while the oedipal model does not, that gives it this edge.

How can we compare the aggression model with the Kleinian model when it comes to the cogency of psychological explanation? It seems that such a comparison is rather straightforward and tells in favour of the Kleinian model. The aggression model directly explains the authority of conscience as, according to it, the parental agency is introjected *qua* authority figure. Similarly, the Kleinian model directly explains the authority of conscience as, in the paranoid-schizoid position, part objects are introjected *qua* authoritative, persecuting figures. We saw above, however, that the aggression model is only able to indirectly explain the erection of the ego ideal. The Kleinian model, on the other hand, directly explains the erection of the ego ideal as, during the depressive position, objects are introjected *qua* loved objects. Thus, the Kleinian model will be able to give more cogent explanations, involving the ego ideal, than can the aggression model¹. We should now turn to a concrete example of moral action, giving (very rough) explanations of it in the light of each of the three psychoanalytic models².

In section two I gave the example of an agent giving to charity, and explained this action by attributing to him a particular demand of conscience and the desire to act in accordance with that demand. What more does the oedipal model say in this case? The oedipal model, as do all the psychoanalytic models, interprets the demand of conscience as a command issuing from the super-ego. That the super-ego issues this

¹ Also, given the erotic element in early oral, cannibalistic phantasies, there is a direct relation between the individual's libido and his (early) super-ego. Thus, the Kleinian explanations of the sexualization of morality are likely to be more cogent than those of the aggression model.

² The following examples are still, to a certain extent, abstract. This must be so if we are to appreciate the scope of the explanations involved.

demand is explained by the original introjection of the parents (along with the influence of later introjections). This original introjection of the parents is explained by the situation in which the individual felt (or feels) his loved objects to be lost in a certain sense. The oedipal model also says, of the desire to act in accordance with the demand of the super-ego, that it is explained, to a large extent, by the fear of punishment by the super-ego.

The aggression model explains the action in a similar, but importantly different, way. Again, the demand of conscience is interpreted as a command issuing from the super-ego. Again, the super-ego's issuing of this demand is explained by the original introjection of the parents (along with the influence of later introjections). However, this original introjection is explained in terms of the natural aggressive impulses of the individual towards his parents. But the natural aggressive instincts of the individual play another explanatory role. They go to explain why the individual fears the punishment of the super-ego (that fear being postulated as an explanation of the desire to act in accordance with the demand of the super-ego). We can see then, that the aggression model is more explanatory (giving a direct explanation of the fear of the super-ego) and is so in a way that postulates fewer base motives (just the natural aggressive instincts). This gives us reason to accept the aggression model over the oedipal model.

We can now compare this 'aggression explanation' with a Kleinian explanation of the same action. The Kleinian model explains the action in essentially the same way as the aggression model. Natural aggression in the individual is the base explanation for both the fear of punishment by the super-ego and the particular demands of conscience (via the introjected objects). So, the present case cannot tell between these two models. As I pointed out above, it is in explanations involving the ego ideal that we should expect the Kleinian model to be more persuasive. Thus, given a different structure of explanation, we might want to explain the agent's desire to do his duty by attributing to him the desire to emulate his ego ideal. In this case the Kleinian model will fare better than the aggression model due to its more direct explanation of the existence of the ego ideal itself. The Kleinian model is economical in its attribution of base motivations relying on aggressive and libidinal instincts.

So, the Kleinian model allows more cogent psychological explanations than either the oedipal model or the aggression model. We have yet to look, however, at whether each model is independently plausible in the light of empirical developmental psychology.

Jean Piaget is the most influential figure within empirical developmental psychology. In the field of moral development the work of Lawrence Kohlberg leads the way. I will look, briefly, at an aspect of Piaget's theory and then, more closely, at Kohlberg's theory. It will become clear that the work of these two experimental psychologists parallels, in certain important respects, the Kleinian model of the development of the mental life and, in particular, the moral development, of the child. This fact, I claim, lends empirical support to the Kleinian model, and gives us reason to accept that model over either the oedipal or the aggression models.

In his paper, 'Synthesis in the Imagination: Psychoanalysis, Infantile Experience and the Concept of an Object', Hopkins draws attention to an idea which is central to both Kleinian theory and Piaget's theory of child development. This is the idea that, "the infant's use of the concept of identity – the application of the idea of a single, enduring object of perception, emotion and action – plays a pivotal role in both cognitive and emotional development." (Hopkins 1987) p. 140. We saw, in section six, that, in the paranoid-schizoid position, the infant does not have the concept of a unitary, persisting object. The infant experiences 'part objects' which are 'split' in a variety of ways. It is only with the onset of the depressive position that the child gains the requisite ability to cognise objects as unitary, persisting wholes. This cognitive achievement plays an important part in Kleinian theory, causing the move from the paranoid-schizoid position to the depressive position, which, according to Klein, is of fundamental importance to the development of the individual. We need not go into the detail of Piaget's theory, suffice it to say that a similar process is said to occur and that it is of vital importance to the development of the individual. The transition from viewing the world as a "fluctuating tableaux without spatio-temporal consistency" to viewing it as "a universe of permanent objects" (Piaget and Inhelder 1969) p. 26, means that the infant can see people as people for the first time. This is clearly an important development.

How does this claim of Piaget's fare in the light of more recent work in developmental psychology?¹ There are a large number of developmental psychologists who would deny Piaget's claims outright. They are those who take recent experimental results to show that the conceptual apparatus and knowledge that Piaget thinks develops, actually we are born with (or, at least, have from an age much earlier than Piaget suggests)². This view holds that infants have an innate ability to represent objects and also an innate ability to represent objects as permanent (as continuously existing when hidden). Thus, there is no significant conceptual development in infancy. However, this is not an uncontroversial claim. Most recently, it has been challenged by Meltzoff and Moore (1999), who argue, rather plausibly, that this conclusion is not warranted by the data. They deny that experiments with infants show that they have the concept of permanence from such an early age, but they agree that it has been shown that infants have both the ability to represent objects and the concept of identity, and that they have it natively³.

Where does this leave Klein's claims? The first point is that the Kleinian account *requires* that infants are able to represent objects. The representation of part-objects is fundamental to projection and introjection etc. Secondly, it is important to be clear about just what it is that developmental psychologists are claiming. The nativists claim that infants have the concept of *the same object* and that they have the ability to reason about objects which are out of view. Meltzoff and Moore accept the first claim, and deny the second. It appears that the Kleinian account requires the truth of this first claim. If the infant did not have the concept of *the same object* how would an introjected object bear the relation it is supposed to bear to its external counterpart? What of permanence, does the Kleinian account require that infants have the ability to reason about objects which are out of view? I do not see that it is required, but it is probably consistent with a modest Kleinian account. The Kleinian would have to drop his claim concerning the lack of object permanence during the first few months, but the claim that infants experience part objects only would be left untouched. In fact, the reason for Klein's claim that the infant lacks object permanence, is her

¹ It will not be possible to go into great detail on this point. For a more in-depth account, readers should consult the references in the following footnotes.

² See, Spelke and Van de Valle (1993), and Carey and Spelke (1994).

³ Hopkins (1987) claims that the concept of identity requires the concept of permanence. On this point I find him unconvincing, see Meltzoff and Moore (1999) pp. 62-68, for the contrary view.

observation that infants introject part objects many times; they think there are many mothers/breasts (or at least two, good and bad). This claim, which is of course related to considerations of object permanency, is supported, “the young infant (less than five months old) thinks that it has a multiplicity of mothers, whereas the older infant knows that it has only one”¹. I will not pursue the question as to whether or not this phenomenon is necessarily due to lack of object permanence in the young infant. Suffice it to say that there are elements of the developmental tradition that tend to confirm Klein’s theory.

Finally, what of the Kleinian’s claim that, during the first few months, the infant experiences part objects, only gaining the ability to perceive and reason about whole objects around the fifth month of life? Unfortunately, there is precious little experimental evidence concerning infants’ reasoning about persons, most research is directed towards reasoning about simple, inanimate objects. However, what evidence there is tends to support the Kleinian claim. Hopkins (1987) quotes Campos *et al* (1983) as producing persuasive results, showing that between the ages of four and seven months the infant moves from directing his emotions towards parts of the mother’s body to directing those emotions towards the mother as a whole². Although this area is not well researched, we can tentatively say that Klein’s claim is, again, supported.

So much for cognitive development, What of specifically *moral* development? Kohlberg claims that, according to the experimental data, individuals begin life as egocentric agents and only become recognisably *moral* agents through the internalisation of external values and discipline. He suggests that there are six stages through which the individual passes, each of these stages leads on to the next and there is no regression. Each stage represents both a form of cognitive and moral development. These stages are: (1) ‘Punishment’, (2) ‘Reward’, (3) ‘Interpersonal relations’, (4) ‘Social order’, (5) ‘Social contract’, and (6) ‘Universal rights’.

An individual at stage (1) obeys in order to avoid punishment, his reasons for acting morally are purely egotistical, to avoid external punishment. An individual at

¹ Bower (1982) p. 217.

² See Campos *et al* (1983) p. 824.

stage (2) is essentially similar to a stage (1) individual, except that it is now recognised that others have their own interests to follow and it is right for them to do so. At stage (2) it is not so much the fear of punishment for wrongdoing as the desire for reward for rightdoing that motivates the individual. An individual at stage (3) is motivated to act morally because he wants his social group to like him. At this stage it is possible for the interests of one's group to take precedence over one's own interests. At stage (4) an individual is motivated to remain loyal to one's social institutions, doing what is right means maintaining these institutions. An individual at stage (5) is motivated to act morally in accordance with rules of behaviour, that transcend particular social institutions, because they respect the rights and interests of others/society as a whole. At stage (6) an individual has a conception of, and accepts, universal ethical principles which are binding on all rational creatures.

Kohlberg's theory has many critics and many supporters¹. In general, however, there seems to be a wealth of empirical support for Kohlberg's theory², even those who are dissatisfied with much of what Kohlberg has to say admit that the main aspects of the theory are well confirmed³. There are, however, some well-documented problems with the theory. The first is that stage (6) seems to be very rare indeed, it is reached by few, if any people. This means that the empirical evidence supporting stage (6) is weak. Secondly, many studies have shown that young children have a richer sense of positive morality than suggested by stage (1) reasoning, which is claimed to be predominant in virtually all pre-school children⁴. With these problems in mind, we can turn back to the Kleinian model in order to see how consistent it is with the empirical material presented by Kohlberg.

At a first glance the transition from (1) to (2), and to an extent (3), looks similar to the Kleinian model's transition from the paranoid-schizoid position to the depressive position. Fear of punishment is replaced with the desire for reward and then with acceptance within a group. Thus it would seem that the empirical evidence supports Kleinian theory in this respect. However, things are not so simple. The Kleinian

¹ See Rich and DeVitis (1985) pp. 87-98.

² "Validating studies include a twenty-five year study of Chicago area boys, middle and working class; a six year study of Turkish village and city boys of the same age; and various cross-sectional studies in Britain, Canada, India, Israel, Honduras, Taiwan, and Yucatan" (Rich and DeVitis) p. 90.

³ "Kohlberg's theory is one of the best articulated and most thoroughly supported theories we have in all developmental psychology." (Matthews 1987) p. 187.

⁴ See Damon (1999) pp. 58-9.

model locates the transition to the depressive position round about the fifth month of life whereas, on the Kohlbergian picture, the transition from stage (1) to stage (2) might not be effected until the age of 10, or even as late as 13¹. Can this discrepancy be accounted for? There are two routes we might take to try and square the results of experimental psychology with the Kleinian model. The first, mentioned above, is to stress those studies which suggest that very young children have a more developed sense of positive morality than the Kohlbergian theory suggests. The second, mentioned in section six, is to point out that the transition from the paranoid-schizoid to the depressive position occurs gradually, in stages *beginning* in the fifth month of life².

This 'age gap' might suggest that Kohlberg's theory supports the Freudian models to a higher degree than the Kleinian model since, on both Freudian models, the super-ego develops around the age of five. But this is not the case. On both Freudian models, stage (1) thinking and stage (2) thinking come about at the same time as the super-ego and the ego-ideal come into existence simultaneously. On the Kleinian model, the early, paranoid-schizoid super-ego is a prior structure to the depressive super-ego and corresponding ego-ideal. Thus, stage (1) thinking is prior to stage (2) thinking.

To a certain extent stage (3), and certainly stages (4), (5) and (6) are of less interest to us here. These later stages represent developments that occur after the basic super-ego has been erected. These developments are explained by psychoanalytic theory as being effected through the introjection³ of the ideals of authority figures, which occurs throughout life. Because these later stages are not directly relevant to the discussion (and stage (6) presents its own problem of the lack of empirical support) I shall restrict my discussion to stages (1) and (2)⁴.

Another similarity between the Kleinian model and Kohlberg's theory is the idea that cognitive and moral development are linked. Each of the stages (1) to (6)

¹ See Damon (1999) p. 58.

² Also, Kohlbergian methodology primarily involves assessing the verbal behaviour of its subjects. It is a moot point, however, how long the attitudes of the primarily emotional depressive position take to 'filter through' to the domain of conscious, verbally oriented behaviour.

³ Into the ego or the super-ego, see page 22, note 2.

⁴ In section eleven I give a brief account of how a Humean theory might account for stages (1) to (5), the psychoanalytic account would be essentially similar.

represents both a cognitive and a moral level. Kohlberg claims that reaching a given cognitive stage is a necessary condition of reaching that moral stage, although it is not sufficient. Klein also claims that cognitive development is a necessary precondition for moral development. The cognitive ability to relate to whole objects, as opposed to part objects, is a pre-requisite of achieving the mature super-ego characteristic of the depressive position. Thus, both theories suggest that cognitive and moral development go, to some extent, hand in hand.

It is worth mentioning one further piece of empirical material, as it tells against the oedipal model. On the oedipal model the super-ego is the 'heir to the Oedipus complex'. However, according to Freud, the relinquishing of the Oedipus complex is achieved through different means in girls than it is in boys due to differences in castration anxiety¹. This results in girls having different super-egos than boys and a corresponding difference in male and female moral sensibilities. But this claim has not held up to the scrutiny of empirical psychologists, there do not seem to be any substantial differences between male and female moral orientations². This can be seen, I think, as evidence against the oedipal model.

So, in the light of experimental, developmental psychology what can be said about our three psychoanalytic models? It seems clear that the Kleinian model is better supported than either the oedipal or aggression models. This is all very well, but can we go further and say that the Kleinian model is actually well supported by the empirical evidence? Firstly, we saw that the Kleinian claim that there is a transition from 'part object experiences' to 'whole object experiences' around the fifth month of life is supported by experimental psychology. Secondly, Kleinian theory agrees with Kohlberg's findings that stage (1) thinking ('punishment') precedes stage (2) thinking ('reward'). Lastly, the idea that cognitive development is a condition of moral development is common to both the Kleinian model and the findings of experimental psychology. This, I claim, amounts to good empirical support for the Kleinian model

¹ See page 19, note 1.

² "Well designed studies of American children...rarely detect differences between boys' and girls' ideals. Even for adults, when educational or occupational levels are controlled, the differences disappear. Female lawyers have almost the same orientations as their male counterparts; the same can be said for male and female nurses, homemakers, scientists, high school dropouts and so on." (Damon 1999) p. 60. Of course this point of view has been famously contested by Carol Gilligan (Gilligan 1982), who argues that there is a distinctively masculine 'ethic of justice' and a distinctively feminine 'ethic of care'. However, the sway of opinion within experimental psychology is against her. See, Walker (1984) and (1991); and, Walker, Pitts, Henning and Matsuba (1995).

(or, at least, central aspects of it). This means that we can view the Kleinian model as being plausible, independently of the explanations it allows. When this is added to the conclusion above, that the Kleinian model allows more cogent psychological explanations of moral actions than either of the other two models, it amounts to a sound reason for accepting the Kleinian model as true.

Eight

The rest of the paper will work with the assumption that the (Kleinian) psychoanalytic account¹ of the development of the faculty of conscience in the individual is, at least in broad outline, correct. As I made clear in section two, the concept of conscience is of fundamental importance in distinguishing moral from non-moral actions. Since the topic of moral action clearly falls within the domain of ethics, the question arises as to whether the psychoanalytic account has an important contribution to make to moral philosophy. Those who have addressed themselves to this question have come up with a broad spectrum of different answers. David Jones has claimed that psychoanalysis can make no contribution to moral philosophy because, stated baldly, Freud tries to reduce the moral emotion of guilt to the non-moral feeling of anxiety, and this can't be done (Jones 1966). At the other end of the spectrum, Lawrence Friedman has claimed that philosophy, "must relinquish to psychology the study of the meaning of ethical statements" (Friedman 1956) p. 19. These positions are extreme and the remainder of this paper will, partly, be concerned with steering a course between them.

Much recent discussion of the relevance of psychoanalytic theory to ethics concerns the opposition between rationalistic and naturalistic accounts of moral reasoning, motivation and action². This is not the kind of naturalism that defines moral terms in non-moral terms, the kind that G. E. Moore famously argued against in *Principia Ethica*. Rather this is the kind of naturalism which claims that, "morality originates in certain natural movements of the psyche, which do not themselves

¹ From now on, the term 'psychoanalytic account' will refer to that which I have been describing as 'the Kleinian model'. Although, due to the object of Jones' attack, section nine will continue to refer to Freud.

² See Scheffler (1992a) and (1992b), chs. 4 and 5; Deigh (1996), ch. 6; and Wollheim (1993a) and (1993b).

require reference to morality either to describe or explain them” (Wollheim 1993a) p. 9. At first glance, the psychoanalytic account outlined in the previous sections appears to present a naturalistic theory of this kind. One of the main projects in the following sections will be to approach this claim with two specific questions in mind: Does psychoanalytic theory provide a good basis for a reasonably sophisticated and plausible naturalism? And, if psychoanalytic theory does provide such a basis, in what respect might it provide arguments against non-naturalistic positions, i.e. is psychoanalytic theory incompatible with non-naturalism?

The first of the above questions invites a comparison between psychoanalytic theory and other competing naturalistic accounts. The second requires a characterisation of what is required of a theory for it to be *non*-naturalistic. It is often claimed that the defining characteristic of non-naturalistic theories is that they place fundamental importance on the authority of reason. Since this is the case, it will be important to trace the role that reason plays within the (supposedly naturalistic) psychoanalytic account itself.

A related question is that of realism/anti-realism. In particular I will consider the anti-realist theory known as ‘projectivism’. Section thirteen is concerned with the relationship between psychoanalytic theory and the status of value ascriptions.

These considerations will lead us towards a general conception of the status of moral psychology within the field of ethics. The conclusions reached will answer the question, posed in section one, as to the consequences, for moral philosophy, of the psychoanalytic account of moral motivation. My first task, however, is to address an objection that threatens to undercut any claim to the effect that the psychoanalytic account can have any relevance at all to moral philosophy. This objection has been forcefully put by David Jones, and discussion of it will occupy the next section.

Nine

In his paper, “Freud’s Theory of Moral Conscience”, David Jones comes to the conclusion that Freud does not satisfactorily explain the phenomena of bad conscience. Jones’ argument can be condensed as follows: (1) Freud claims that the

moral emotion of guilt is a species of anxiety, (2) Anxiety is an occurrent-feeling¹, (3) The moral emotion of guilt is not an occurrent-feeling but a 'generic disposition', explained by consciously held beliefs, (4) Contra Freud, the moral emotion of guilt cannot be a species of anxiety².

It should be clear why this argument threatens to undermine the claim that psychoanalytic theory has an important contribution to make in moral philosophy. If psychoanalytic theory's talk of super-egos etc is really only talk of anxiety, and anxiety is different from guilt, the relevance of psychoanalytic theory to moral philosophy becomes obscure. So, should we accept Jones' argument?

That Freud thought guilt to be a species of anxiety is patently true. When talking of the 'unconscious sense of guilt', he is explicit

Even in obsessional neurosis there are types of patients who are not aware of their sense of guilt, or who only feel it as a tormenting uneasiness, a kind of anxiety...Here perhaps we may be glad to have pointed out that the sense of guilt is at bottom nothing else but a topographical variety of anxiety; in its later phases it coincides completely with *fear of the super-ego*. (Freud 1930) p. 135.

In fact, each of the psychoanalytic models relies on the idea that the moral emotion of guilt arises as a transformation from, externally directed, fear of the loss of love, which is itself a species of anxiety. Guilt, on the psychoanalytic model, is a particular kind of anxiety, moral anxiety.

What of Jones' claim that anxiety is an occurrent-feeling? The reason for this claim is to distinguish anxiety from guilt (which is seen as a disposition). Thus, it amounts to the claim that anxiety is *not* a generic disposition. If one is not currently and consciously experiencing anxiety, then one cannot be said to feel anxious. This is contrasted with guilt; even though someone may not currently and consciously be experiencing guilt feelings, it is still possible to truly describe that person as feeling guilty³. Two questions arise. Firstly, is it plausible to view anxiety as a generic

¹ By the term 'occurrent-feeling', I take Jones to mean a currently occurring sensation with its own phenomenological qualities, an example of which would be a pain.

² According to Jones, this argument can be generalised to moral emotions other than guilt (shame, remorse, consciousness etc).

³ This claim is more fully explained and justified below.

disposition, as opposed to an occurrent-feeling? Secondly, would the view that anxiety is a generic disposition in some way contradict Freud's theory?

What would it be for anxiety to be a generic disposition? Following Ryle's characterisation of 'generic disposition'¹, we can say that it would involve a tendency to perform various kinds of behaviour, and tendencies to experience certain occurrent-feelings. This is the view Jones takes of guilt, and I see no obvious reason why it is implausible to view anxiety in the same way. There seem to be, and this is also true of guilt, two ways of understanding the term 'feeling of anxiety'. One is as an occurrent-feeling, and one is as a generic disposition. Saying this allows us to make sense of the fact that it often seems natural to say of someone that they feel anxious even though they are not currently and consciously experiencing anxiety feelings.

Of course, this does not settle the matter. It might be the case that psychoanalytic theory *requires* us to view anxiety simply as an occurrent-feeling. But this does not seem to be the case. Psychoanalytic theory often remains silent on the question of whether anxiety is an occurrent-feeling or a generic disposition. However, I think it is reasonable to read Freud as holding something corresponding to the account I have sketched above

Anxiety is always present somewhere or other behind every symptom; but at one time it takes noisy possession of the whole of consciousness, while at another it conceals itself so completely that we are obliged to speak of unconscious anxiety or, if we want to have a clearer psychological conscience, since anxiety is in the first instance only a feeling, of possibilities of anxiety. (Freud 1930) p. 135.

Here Freud contrasts the occurrent-feeling sense of 'anxiety' (which he simply calls 'anxiety') with the generic disposition sense of 'anxiety' (which he calls 'possibilities of anxiety'). That to which Freud refers when he speaks of 'possibilities of anxiety' looks remarkably like anxiety being viewed as a disposition. In the special case of guilt, this possibility of anxiety is called the 'need for punishment' and it involves a tendency to perform various actions and a tendency to experience various occurrent-feelings. Pinning an occurrent-feeling theory of anxiety on Freud is, it would

¹ Ryle (1949) pp. 117-8.

appear, to misrepresent his views¹. Far from conflicting with psychoanalytic theory, a dispositional view of anxiety is part of the theory itself.

What should we make of Jones' claim that guilt is a generic disposition, explained by the existence of consciously held beliefs? I have already indicated that viewing anxiety as a generic disposition is reasonable, and I see no problems with viewing guilt in the same way. Of course we must be sure to distinguish between guilt as a disposition and guilt as an occurrent-feeling (because surely there is such a sense of 'guilt'), but this is as it should be. There is, however, a problem for the psychoanalytic theory; this is the idea that guilt is explained by the existence of consciously held beliefs. Jones thinks that we can only truly say of someone that they feel guilty if they consciously believe that they have done something that is morally wrong. If this is correct, then guilt is not anxiety, because (moral) anxiety is not explained by conscious beliefs, but by the action of the super-ego (which often operates unconsciously)².

It must be noted that Jones does not argue for this view. On his account it is simply through an analysis of the concept of moral guilt, that we can see that to feel guilty a person *must* consciously believe that they have done something reprehensible. Is this correct? Presumably the psychoanalyst would deny that it is, but we must be given some reasons for rejecting this *prima facie* plausible claim. We might think that Jones has mistaken feeling guilty with *justifiably* feeling guilty. Jones' conceptual analysis correctly picks out the conditions in which someone can be said to be justified in feeling guilty. It is reasonable to hold that only in the presence of the relevant beliefs are guilt feelings rationalised³. This does not, however, show that guilt feelings are only correctly attributable on the assumption of such beliefs. Psychoanalytic theory gives a suggestion as to how it is that we come to have either the occurrent-feeling of guilt, or the generic disposition of guilt. If the psychoanalytic

¹ The misrepresentation is, however, understandable. Freud restricts his use of the term 'anxiety' to the occurrent-feeling sense, for instance, "Anxiety, then, is in the first place something that is felt. We call it an affective state" (Freud 1926) p. 132. This does not mean, however, that there is no dispositional notion in operation.

² This point, and the preceding paragraphs, suggest that the real work, in Jones' argument, is being done by his claim that guilt is explained by consciously held beliefs, and that his distinction between occurrent-feelings and generic dispositions is something of a red herring.

³ Although whether those beliefs must be conscious, is another question.

theory is correct then it is simply not the case that guilt feelings only come into existence through consciously entertaining moral thoughts (although, on many occasions, they may).

The psychoanalyst will point to cases in which an individual has a disposition which operates in the same way as the generic disposition of guilt, i.e. the person acts as though he feels guilty and experiences 'pangs' of guilt, but where the rationalising beliefs are absent, or unconscious. Psychoanalytic theory can give cogent explanations of these cases, in terms of the action of the super-ego. Jones' claim that, in the absence of rationalising beliefs, such dispositions *cannot* be guilt, is unconvincing. He is, in effect, stipulating against naturalistic forms of explanation, not arguing against them.

So Jones' argument, to the effect that psychoanalytic theory does not concern ethics, fails. Jones does not show that guilt cannot be a species of anxiety. Neither does he show that the conceptual analysis of guilt tells against the psychoanalytic account. There are no conceptual barriers to viewing guilt as a species of anxiety. That the psychoanalytic account views guilt in this way is not an argument against it. So what can we learn from this encounter with Jones? I think that we should conclude that it is unwise to take the frontiers of philosophical discourse as too rigidly delineated¹. To put matters crudely, Jones' approach consists in looking up a term (guilt) in a dictionary and drawing conclusions from the conceptual analysis of the meaning given there. This is precisely what psychoanalytic approaches, amongst others, deny is the correct way of proceeding. Again crudely, the psychoanalytic approach claims (contra Jones) that the concept of guilt is not fixed in this way. Understanding the concept of guilt requires an account of why that emotion exists in the first place, and this must be given within a moral psychology. If correct, this approach would give moral psychology an important role to play in moral philosophy. Of course, it may turn out that this approach, exemplified by psychoanalytic theory, is mistaken, but it should be clear that we cannot argue against it simply by assuming the correctness of the kind of approach apparent in Jones' paper. These comments will be amplified below.

¹ For this point, see Wollheim (1993b) section 1.

Ten

I turn now to naturalism and non-naturalism in moral philosophy. The obvious starting point is the question of what is meant by these two terms. In section eight I indicated that by ‘naturalism’ I do not mean definitional naturalism. I take definitional naturalism to be the claim that goodness can be defined in terms of non-moral properties, for instance utility¹. Rather, I suggested that Wollheim’s definition² of naturalism better captured my meaning. However, it looks as though Wollheim’s definition is rather uninformative. We are asking what is meant by the term ‘naturalism’ and the definition we are given relies on the phrase ‘natural movements of the psyche’. But what makes a ‘movement of the psyche’ natural? I think that the short answer here is that a natural movement of the psyche is one which is both intrinsic to human psychological development³ and that we would not tend to call moral. Therefore, ‘naturalism’ is the thesis that morality originates in certain non-moral psychological developments or emotions etc⁴. The claim is that ‘the moral’ has its origin in ‘the non-moral’⁵. We can then, define non-naturalism negatively as the claim that morality does *not* originate in non-moral psychological developments or emotions etc⁶.

We can call this type of naturalism, ‘genetic naturalism’. It will, I think, be fruitful to place genetic naturalism within the context of other types of naturalism. There is a common motivation towards pursuing naturalism in morals and once we have seen some of the standard shortcomings of other types of naturalism we will be in a position to see what might motivate us towards pursuing a genetically naturalistic

¹ This could involve either a reduction or a supervenience claim.

² See page 52-3.

³ This clause is intended to rule out, for example, divine command theories, of the origin of morals and/or theories which locate the origin of morals in *sui generis* social entities.

⁴ Here, as in the whole of the paper, I am talking of the origin of morality in the sense of the origin of morality *within the individual*. No claim about the historical development of morality is intended. The term ‘originates’ might be thought to be rather vague. However I am going to leave it as such. See page 72, note 1, for some relevant remarks.

⁵ This does not entail that talk of ‘the moral’ is really nothing over and above talk of ‘the non-moral’ or that moral properties can exist only in the presence of non-moral ones, i.e. that some kind of definitional naturalism is true.

⁶ Later I raise the question of giving non-naturalism a positive characterisation. In particular I will examine the idea that rationalism and non-naturalism coincide.

theory. So, what are the common types of naturalism in morals and in what way do they count as naturalistic? I have already mentioned definitional naturalism (the view that moral terms can be defined in non-moral terms). A criticism often levelled at definitional naturalists is that they attempt to derive an 'ought' from an 'is'¹. Whether definitional naturalists commit a logical fallacy of this kind is certainly questionable². However definitional naturalists do have to face a problem that is strikingly similar to Moore's 'open question argument'³. It is always possible to wonder whether the moral term being defined *really is* x (the proposed definition). This question seems intelligible, but how can we use the redefined term to state it? Another kind of naturalism is Aristotelian, or neo-Aristotelian, naturalism, which concerns itself with what it is to be a good/virtuous human being, or to perform good/virtuous actions. What it is to lead a virtuous life is specified in non-moral terms, this gives us an account of morality which can certainly be labelled 'naturalistic'. Such accounts, however, fall prey to the view, common to our century, that there is no specifiable 'human teleology', no specifiable 'good human life'⁴. The kind of naturalism that I propose the psychoanalytic account of moral sentiments is a variant of, is genetic naturalism. This naturalism claims that morality *originates* in that which is itself non-moral.

What is it that links these theories as naturalistic? And what is the impetus behind them? These are difficult questions to answer. 'Naturalism' names a family of theories which resemble each other in various ways. Certainly, naturalists share a suspicion of irreducible moral properties such as 'goodness' and 'badness' actually 'out there' in the world. Naturalists claim that these properties are reducible to, or supervene on, or originate in etc non-moral properties. This suspicion is often fuelled by a strict scientific outlook, the view that everything must be describable by science⁵, but this is not always so⁶. In any case, genetic naturalism is one of the family of

¹ See Moore (1959).

² See Pigden (1991).

³ For the 'open question argument', see Moore (1959); Darwall, Gibbard and Railton (1997); and Pigden (1991).

⁴ "it is hard to believe that an account of human nature – if it is not already an ethical theory itself – will adequately determine one kind of ethical life as against others." (Williams 1993) p. 52. See also Darwall, Gibbard and Railton (1997) pp. 24-26. Another problem with neo-Aristotelian naturalism is that it relies heavily on the attribution of character traits to people. This kind of attribution has recently been challenged by philosophers and social psychologists. See, for instance, Harman (1999).

⁵ Freud probably held something like this view.

⁶ See Williams (1993) pp. 121-2; and Pigden (1991) p. 422.

naturalistic theories and it is the one that I shall limit my discussion to. From now on, the term 'naturalism' will refer to genetic naturalism only.

We are now in a position to begin to answer the question posed in section eight, Does psychoanalytic theory provide a good basis for a reasonably sophisticated and plausible naturalism? I think that the answer to this question must be in the affirmative.

All three psychoanalytic models are naturalistic by the above definition; I will restrict myself, however, to the Kleinian model. For a theory to be naturalistic, it must find the origin of the individual's morality in specifiable psychological developments which do not themselves require reference to moral concepts or phenomena. Throughout the course of this paper I have been associating an individual's morality with his conscience and the cluster of moral emotions that are bound up with the conscience. The Kleinian model views conscience as a function of the super-ego. So, if the formation of the super-ego is explained in terms of psychological developments that are non-moral, the Kleinian model can be classed as naturalistic. I hope that it is clear that this is the case. The main concepts used to explain the emergence of the super-ego in the individual are: the aggressive (and sexual) instincts; introjection; identification; anxiety (fear of the loss of love); and, the external parental agency. None of these concepts require any reference to morality. So, the psychoanalytic model is, clearly, a naturalistic account of the origin of morality.

What criteria must a naturalistic account satisfy in order to qualify as plausibly sophisticated? And does the psychoanalytic account satisfy them? I will outline what I see to be the four most important of such criteria¹, these are:

- (1) The resonance of morality.
- (2) The fragility of morality.

¹ The following list is by no means exhaustive, but it will suffice as a guide. It is also important to make clear that these criteria are supposed to help us decide between competing naturalistic theories. They are *not* supposed to be decisive between naturalistic and non-naturalistic theories, since (1), (2) and (4) are clearly weighted in favour of naturalism.

- (3) The phenomenology of conscience (in particular, the apparent authority of morality).
- (4) The empirical evidence provided by developmental psychology.

To what do I refer when I speak of ‘the resonance of morality’¹? Simply the fact that moral concerns are to be found throughout our psychological and social lives. Morality does not appear to us to be a domain separate from the rest of our lives. Morality is implicated in a vast range of attitudes and (moral) emotions, and a naturalistic account *must* explain how moral concerns become so enmeshed within our psychological makeup. It should not be difficult to see that the psychoanalytic account easily accounts for this resonance. Moral emotions such as resentment and guilt are explained by the action of the super-ego on the ego. Sentiments such as moral approval are explicable in terms of a similarity of the object of the sentiment with the ideals of the super-ego. Moral concerns are the result of relationships between central structural elements of the mind (the ancestors of which being relationships between the individual and others), so it isn’t surprising that they resonate throughout our psychological and social lives.

What is ‘the fragility of morality’? The fragility of morality is the well-known phenomena of the disfigurement of morality and moral motivation. For instance, self-righteousness seems, in some way, to be motivated by moral concerns, but in such a way we think has somehow ‘gone wrong’. There are many more examples of such disfigured moral concerns, excessive moralising and ‘pale criminality’² being two. It seems that morality and moral motivation are ‘fragile’, they are easily disfigured. The psychoanalytic account seems to be in a good position for accounting for this phenomenon. Because the individual’s conscience is so dependent on contingent developmental factors, the super-ego is not necessarily a stable structure and may develop in imperfect ways. A sentiment such as self-righteousness might be explained by saying that the love felt for the ideals of the super-ego/ego-ideal (which is a continuation of the admiration and love felt for the parents) is, due to unfortunate

¹ The terms ‘resonance’ and ‘fragility’ belong to Scheffler, and the following owes much to his work. See Scheffler (1992a) and (1992b) Chs. 4 and 5.

² A pale criminal is someone who commits a crime *because* he feels guilty, in order to assuage that guilt. See Freud (1916) part 3; Klein (1927b); and Wollheim (1993e).

developmental factors, excessively strong and a kind of narcissistic attitude develops.

Similar explanations of specific disfigurements of moral motivations can be given.

For example, of the phenomenon of pale criminality, Klein writes

It is just anxiety and the feeling of guilt which drive the criminal to his delinquencies... everything seems to point to the conclusion that it is not the lack of a super-ego but a different development of the super-ego – probably the fixation of the super-ego at a very early stage – which will prove to be the main factor. (Klein 1927b) p. 184.

A naturalistic theory must be able to account for the phenomenology of conscience and, in particular, the apparent authority of morality. To a certain extent this issue has been dealt with in the sections detailing the success of the three psychoanalytic models in explaining what I said were four important features of conscience (the ‘otherness’ of conscience, its intimate relation to moral emotions, its apparent authority and, its severity). I indicated that the psychoanalytic models were convincing on this point. In this section I will expand a little on the ability of the Kleinian model to account for the apparent authority of conscience and morality.

One philosopher who has written convincingly about the authority of morality is Immanuel Kant. In section 1 of *Groundwork of the Metaphysics of Morals*, Kant gives an account of what he considers to be our pre-philosophical intuitions about moral motivation. He thinks that we only attribute to actions a specifically *moral* worth if they are done out of a sense of duty. This sense of duty is independent of any of our inclinations, sentiments or self-interest. This is obvious when we see that we are inclined to attribute moral worth to an action when it is done contrary to that person’s inclination. In fact, these are the most clear-cut cases of specifically moral motivation. Furthermore, the sense of duty which motivates moral actions has an authority over our inclinations etc. Kant expresses this authority in his claim that to act morally is to do one’s duty. In this way, moral concerns are overriding.

Suppose, then, that the mind of...[a] philanthropist were overclouded by his own grief, which extinguished all sympathy with the fate of others, and that while he still had the means to benefit others in distress their troubles did not move him because he had enough to do with his own; and suppose that now, when no longer incited to it by any inclination, he nevertheless tears himself out of this deadly insensibility and does the action without any inclination, simply from duty; then the action first has its genuine moral worth. (Kant 1997) p. 11-12.

This seems to me to be a reasonable characterisation of our pre-philosophical intuitions about moral motivation. It is part of what I have called ‘the phenomenology

of conscience' that morality is felt to have authority, to be overriding¹. The question remains as to whether the psychoanalytic account can explain this felt authority.

The psychoanalytic account does not have to show that moral concerns are, in actual fact, overriding. It can remain agnostic on this point². What it must show is how we come to regard morality as overriding for, rightly or wrongly, we do so regard it. In section six I said that moral considerations are imbued with authority as a result of their origin in the commands of the authoritative parents. Indeed, this is the crux of the psychoanalytic position, but can we say more about authority and overridingness?

On the psychoanalytic picture, before the super-ego is set up, the parents are seen as virtually omnipotent beings, their commands are categorical. However, the reasons for acting on these commands at this stage are hypothetical, the infant fears the loss of love. So, even though a demand is categorical (it is of the form "you should do x"), the reasons motivating the infant to act on it are hypothetical (they are of the form "I should do x, otherwise y will occur"). This means that when the parental agency is internalised, the super-ego issues demands that resemble categorical imperatives but the individual's actions will be based on hypothetical imperatives, the condition being avoidance of the super-ego's punishment. The psychoanalytic account does not rule out the *bare possibility* that an agent may act on a categorical imperative. The claim is simply that, in many or most cases, what looks like an action based on a categorical imperative is, in fact, an action based on a hypothetical one. In this way, I think the psychoanalytic account can explain the apparent authority of morality. The demands of conscience present themselves as overriding, but reasons motivating the individual to act morally usually include fear of punishment by the super-ego (or desire for reward by the ego ideal).

Above, I said that psychoanalytic theory must account for the empirical evidence presented by developmental psychologists. It should be clear from section seven that

¹ Whether we normally feel morality to be *absolutely* overriding or just overriding in most cases, I am unsure. I am inclined to deny the stronger claim, but I do not rely on this in what follows.

² In fact, as Wollheim points out, the psychoanalytic account gives us reason to suspect that morality is not actually overriding. Wollheim (1984) p. 225.

it succeeds in doing this. I have shown that the Kleinian account is consistent with, and supported by, Piaget's theories and also more recent work in cognitive development. I also showed, in section seven, how the psychoanalytic account is consistent with, and supported by, Kohlberg's empirical studies on moral development¹. Thus, I conclude that psychoanalytic theory can, and does, provide us with a naturalistic theory that is sophisticated enough to account for some of the most salient features of morality, features it had to account for if it was to be in any way plausible.

Eleven

How does psychoanalytic theory fare in a comparison with other naturalistic theories? Naturalistic theories are mostly located within, what is often called, the Humean tradition². David Hume's *Treatise of Human Nature* (particularly Book III) and his *Enquiry Concerning the Principles of Morals* have proved a valuable source of theories in moral psychology and philosophical psychology. This Humean tradition continues to the present day and is evident in various diverse positions. As I don't want to become involved in a discussion and evaluation of the various strands of the Humean tradition I shall go directly to Hume himself. I do not, however, want to present Hume's theory in all its rigorous detail, rather I shall concentrate on certain doctrines that are essential to Hume's position, and that have been influential. Thus the position I will outline is certainly Humean, but not necessarily identical with Hume's³. In order to make a comparison with psychoanalytic theory I will see whether this Humean theory can account satisfactorily for the four salient features of morality outlined above.

For a theory to count as naturalistic it must show how morality originates in certain psychological developments and emotions, a full characterisation of which

¹ See, Kohlberg (1984); Rich and DeVitis (1985) pp. 87-98; Matthews (1987); Thomas (1991); and, Damon (1999).

² Indeed, it is certainly plausible to view the psychoanalytic theory of the origin of moral sentiments as falling within the Humean tradition. In this case, the Kleinian model itself can be seen as a development of Hume's theory. This, of course, does not undermine my enterprise in this section since it is perfectly reasonable to contrast the 'base model' with a particular development of it (particularly when, as in the present case, the differences between the two models are so noticeable).

³ For an excellent discussion of the relationships between Hume's position and other 'Humean' positions see, Snare (1991).

does not require reference to morality. The following quotation from the *Treatise* suggests that a Humean account will be naturalistic in this way

The most probable hypothesis, which has been advanc'd to explain the distinction betwixt vice and virtue, and the origin of moral rights and obligations, is, that from a primary constitution of nature certain characters and passions, by the very view and contemplation, produce a pain, and others in like manner excite a pleasure. The uneasiness and satisfaction are not only inseparable from vice and virtue, but constitute their very nature and essence. (Hume 1978) p. 296.¹

So Hume suggests that morality originates in certain passions which are natural to humans. If a description of these natural passions does not require reference to morality, the theory will count as naturalistic. What, then, are the passions that the Humean thinks give rise to moral sentiments? The natural passion in question is sympathy, humanity, or 'fellow-feeling with others'²

The same endowments of the mind, in every circumstance, are agreeable to the sentiment of morals and to that of humanity...By all the rules of philosophy, therefore, we must conclude, that these sentiments are originally the same; since, in each particular, even the most minute, they are governed by the same laws and are moved by the same objects. (Hume 1975) pp. 235-6).

So, the Humean theory contends that the sentiment of humanity, or sympathy, is natural and basic in humans and that in this sentiment we can find the origin of morality. It is reasonable to admit, with Hume, that humans have a natural sentiment of fellow-feeling with others and that this is describable in terms that are not specifically moral³. Similarly, we can allow that a plausible story can be told (Hume is certainly a great storyteller) leading from the natural passion of humanity to morality and moral sensibilities. The more specific question I want to answer is whether this Humean position is capable, maybe with modification, of accounting for the four salient features of morality given above.

¹ This last sentence suggests that Hume is proposing a reductive account, saying that morality is nothing over and above 'certain characters and passions'. But it is more in line with the rest of Hume's account to say that 'certain characters and passions' are a necessary condition of morality. For this point see Norton (1993) p. 179, n. 24. In any case, since I am presenting a Humean, not Hume's, account, we can afford to be generous and stick with the weaker, and more plausible, claim.

² In the *Treatise Concerning Human Nature* Hume speaks of the sentiment of sympathy, whereas in the *Enquiry Concerning the Principles of Morals* Hume speaks of the sentiment of humanity. This may, or may not, signify a change in Hume's views as to which passions are natural to man. For our purposes, however, we can treat them as equivalent.

³ Psychoanalytic theory posits a similar sentiment in its instinct theory. In Kleinian theory, love (more accurately, a particular kind of love) is treated as basic. See, Klein (1952) p. 115; Hinshelwood (1991), entry entitled 'Love'. See also, Wollheim (1984) pp. 211-213.

In what way might the Humean theory account for the resonance of morality?¹ The Humean account must be able to explain how moral concerns and emotions become so subtly woven throughout our entire psychological makeup. In particular, it must show how the distinctively moral aspect of the moral emotions, for instance guilt, are based on sympathy. If the Humean theory is to be an acceptable account of the origin of morality in the individual, it must show how an individual's feelings of guilt, pride, indignation etc can be explained in terms of his sympathy. I would not wish to claim that this cannot be done, but it should be clear that it is no easy task. However it is construed, as humanity or fellow-feeling, it is, *prima facie*, only plausible to explain 'positive' moral emotions and attitudes (e.g. benevolence, generosity and fairness) in terms of sympathy.

Can the Humean theory account for the fragility of morality? The Humean must explain how it is that moral attitudes are so easily disfigured, resulting in attitudes such as self-righteousness and pale-criminality. Again, the Humean theory appears to fall into difficulties. Hume suggests that the sentiment of morals originates in the sentiment of sympathy, but as soon as we admit that self-righteousness has a moral element, this claim becomes doubtful. It is very difficult to see how a natural sympathy for others can lead to self-righteousness. Again, I am not claiming that the Humean account *cannot* explain the fragility of morality, but there seems to be no easily identifiable way in which it could.

Might the Humean theory have more success in accounting for the phenomenology of conscience, in particular the apparent authority of morality? I think that this question can be answered in the affirmative. The problem for naturalistic theories is that they seem unable to account for the fact that moral demands and considerations present themselves to us as authoritative, or as overriding our particular desires etc. According to naturalism, our moral sensibility originates from these extra-moral desires etc so it seems that they could not override them. The psychoanalytic theory answers this objection by pointing to the authority of the introjected parental authority. How does the Humean theory proceed? The answer lies

¹ This paragraph and the next follow the arguments presented in Scheffler (1992a) pp. 93-5.

in Hume's distinction between natural and artificial virtues¹. In Hume's view, virtues such as generosity and temperance are natural human characteristics, they will be found, to a greater or lesser extent, in every normal person. On the other hand, virtues such as justice and fidelity are artificial. These virtues, of which justice is the most important and most widely discussed, are not to be found naturally occurring in humans, but arise only under a particular circumstance. This circumstance is society. Briefly, that humans are self-interested leads them to create rules of justice for preservation within society. These rules of justice become moral through the workings of the natural sentiment of sympathy. The particularities of Hume's account need not bother us here. What is important is that if Hume is correct in basing moral codes of justice on the sentiment of sympathy, he can answer the problem of the authority of morality. Moral demands are felt to override the desires of particular individuals because they are embedded in the rules of justice which an individual accepts. Justice overrides the partiality of our sentiments, it gives us an impartial perspective which, once accepted, leads to moral considerations which are overriding

'tis easily conceiv'd how a man may impoverish himself by a single instance of integrity, and have reason to wish, that with regard to that single act, the laws of justice were for a moment suspended in the universe. But however single acts of justice may be contrary, either to public or private interest, 'tis certain, that the whole plan or scheme is highly conducive, or indeed absolutely requisite, both to the support of society, and the well-being of every individual. (Hume 1978) p. 497.

How does the Humean theory fare when measured up to the empirical evidence presented by developmental psychology? As we saw in section seven, there are clear parallels between psychoanalytic theory and developmental psychology. However, since the Humean account does not claim to give an account of the development of morality in the child, the relationship between it and developmental psychology is less easy to trace. We must, then, limit ourselves to judging whether anything in, or entailed by, the Humean account explicitly contradicts the findings of developmental psychology.

Following Kohlberg, developmental psychologists claim that moral development progresses through stages². It seems to me that each of these stages ('punishment',

¹ For a clear discussion of this distinction see Mounce (1999) Ch. 7. A more detailed discussion can be found in Mackie (1980) Chs. 6 and 7.

² Because stage six raises its own problems, it would be unfair to ask the Humean theory to give an account of it. Therefore I will limit the following discussion to the first five stages.

‘reward’, ‘interpersonal relations’, ‘social order’, and ‘social contract’) is perfectly compatible with the Humean theory. Developmental psychology does not identify any moral attitudes that cannot easily be accommodated within the Humean framework. The interesting challenge comes when we ask whether the Humean theory can account for the transitions between the various stages.

The transition from stage 1 to stage 2 involves a move from the perceived irrelevance of the concerns of others to the recognition that the instrumental exchange of goods or services can help achieve one’s aims. This is a move from pure self-interest to enlightened self-interest, and from being motivated by the fear of punishment to being motivated by the desire for reward. Although Hume gives no explicit account of how an individual makes this transition, it is clear to see that he could have done without contradicting himself. Both pure and enlightened self-interest play a role in Hume’s philosophical psychology.

The transition to stage 3 involves moving to a position in which one does what is right in order to be liked by others. In this way stage 3 is still based on self-interest but it also allows that the concerns of one’s group can take primacy over one’s own interests. This would, according to Hume, still be classed as a natural virtue. It is a natural characteristic of humans that they are sympathetic to those immediately surrounding them, friends and family. This stage still remains partial, however, in that one’s group is considered above any others. The transition from stage 2 to stage 3 would, presumably, involve the development and ‘coming to the fore’ of one’s natural sentiment of sympathy.

The transition to stage 4 involves submitting oneself to the rule of law; one behaves morally because of the rules of one’s society or situation. This is the first stage at which the Humean must invoke artificial virtues. The move to stage 4, he would say, comes about with entry into human society or, on the level of the individual’s development, something like, the realisation that one is part of such a society. The rules of the society in which the individual is stationed are accepted as mutually useful conventions.

Finally, the transition to stage 5 introduces the notion of obligation and a sense of justice that transcends the particular rules of one's own society. Hume's account of the nature and origin of obligation is by no means detailed or unproblematic, but he does give one¹. The idea of a sense of justice that transcends that of one's particular society develops naturally from stage 4 when the individual's sympathy is offended by some law or rule within their society. Hume says that, "men receive a pleasure from the view of such actions as tend to the peace of society, and uneasiness from such as are contrary to it." (Hume 1978) p. 533. So, when the rules of an individual's society do not tend to promote peace then he will feel uneasy and a sense of justice will arise that is not limited to the laws of that society².

It seems then that the Humean theory does not conflict with the empirical evidence. Furthermore, it seems able to account for the apparent authority of morality³. On the other hand, it is difficult to see how the Humean theory could account for either the resonance or fragility of morality. Thus, we can conclude that it fares slightly worse than the psychoanalytic theory. Of course, what I have been calling 'the Humean theory' is a bare sketch of a position, lacking any real detail. Indeed, I have not even approached those sophisticated, modern theories which would count as Humean. This can be taken to show that I have not *proved* that the psychoanalytic theory is more plausible than the Humean one. However, I think I have shown that the psychoanalytic theory is a force to be reckoned with within the naturalistic tradition. The question I began this section with was, 'how does psychoanalytic theory fare in a comparison with other naturalistic theories?' The answer, I think, is 'very well indeed'.

Twelve

Thus far I have been concerned solely with naturalistic theories. I shall now turn my attention to non-naturalism. The following discussion will pursue four obviously

¹ Hume (1978), Book III, Part II, Sections ii and v. See also Mackie (1980) pp. 96-104; and Norton (1993) pp. 168-171.

² This has the consequence that an individual living in a perfectly just society would probably not reach level 5. This does not seem to me to be problematic.

³ I have not examined to what extent the Humean theory is able to account for the other features of the phenomenology of conscience: its otherness, severity etc. At a glance, it looks as though Hume would have difficulties in this area.

related questions: Is psychoanalytic theory (of the development of morality in the individual) incompatible with non-naturalism in morals? Do moral non-naturalism and moral rationalism coincide¹? Is psychoanalytic theory incompatible with rationalism? And, what is the role played by reason within the psychoanalytic theory itself?

I think that the psychoanalytic theory of the development of moral sentiments in the individual² is clearly incompatible with non-naturalism. Non-naturalism is the claim that the individual's moral sensibilities do *not* originate in non-moral psychological developments or emotions etc. According to psychoanalytic theory, however, the individual's moral sensibilities originate from a conjunction of natural, non-moral, human instincts and the emotional situation of the child. Thus, asserting the truth of the psychoanalytic account is one way of denying non-naturalism.

Before we can give an answer to the question as to whether non-naturalism and rationalism coincide, we will have to say exactly what is meant by the term 'rationalism'. I think there are three interesting theses which could, in this context, lay equal claim to the label 'rationalism':

- (a) The origin of morality is the faculty of reason.
- (b) Every moral action is motivated solely by reason.
- (c) Every moral action is rational.

Rationalism (a) holds that an analysis of the faculty of reason will reveal the origin, or foundation, of morality and moral sensibilities. Rationalism (b) holds that either belief alone can motivate an agent to act morally or, moral action can be motivated rationally independent of desire. Rationalism (c) maintains that, in every given situation, the most rational course of action will also be that which is morally required.

¹ This question may seem strange since rationalism is often defined as the contrary of naturalism. However, it should be clear from what follows that genetic-non-naturalism and rationalism are *not* coextensive.

² i.e. the Kleinian model.

For each of these three versions of rationalism we can ask two questions: firstly, does it contradict naturalism¹? And, secondly, does it contradict the psychoanalytic account? In the case of rationalism (a), an answer to the first question requires a characterisation of ‘the faculty of reason’. If the faculty of reason can be construed as natural (i.e. describable in terms which are non-moral) then rationalism (a) will be naturalistic. Of course, many would want to deny that the faculty of reason is natural in this way. There is a question as to where the burden of proof lies; many naturalist will want to claim that the faculty of reason is natural², the rationalist may well deem it obvious that this is not so. For instance, a Kantian (a typical proponent of rationalism (a)) will claim that ethical considerations are introduced with the very idea of a rational agent. This debate goes very deep and is too complex to pursue here. I will leave it as an open question whether naturalism is compatible with rationalism (a)³.

At first glance it appears that rationalism (b) *is* compatible with genetic naturalism as defined above (morality originates in the non-moral), but I think that this is misleading. The positions are compatible because, as defined, genetic naturalism is silent on the question of the source of motivations to act morally. This is misleading, however, since I believe that any plausible genetically naturalistic theory will, in fact, contradict rationalism (b). In section ten I claimed that any plausible naturalistic theory must be able to account for the resonance of morality. This means accounting for the fact that morality is implicated in a vast range of attitudes and (moral) emotions. I suggest, although I will not argue for it here, that to do so a naturalistic theory *is going to have to* include desire, or some other affective attitude, as a source of moral motivation. If this is the case, then rationalism (b) will be contradicted.

Rationalism (c) is compatible with genetic naturalism. Naturalism, as defined above, makes a claim about the origin of morality in the individual. It is, however,

¹ Recall, we are talking about *genetic* naturalism only.

² For instance, see Cavell (1992), who claims that reason (‘thinking’) arises out of desire and the experience of lack.

³ Rationalism (a) bears a certain similarity with Kohlberg’s explanation of the progression from one moral stage to another. This feature is shared, to a certain extent, by the Kleinian model; see section seven. The similarity between rationalism (a) and either Kohlberg’s or Klein’s theory should not be over-emphasised. Rationalism (a) claims that morality *originates* in reason. Kohlberg and Klein claim only that moral development *is conditional* on the development of the powers of reason.

silent on the question as to whether moral actions are rational or otherwise. Since rationalism (c) is compatible with genetic naturalism, non-naturalism and rationalism (of kind (c)) do not coincide.

We can, however, ask the further question, Are the different versions of rationalism compatible with the psychoanalytic account? Rationalism (a) is not compatible with psychoanalytic theory. According to the psychoanalytic picture, an essential part is played by the (non-rational) instincts and by the emotional situation of the child, in the origination of morality. Thus, if the psychoanalytic account is correct, morality cannot have originated in the faculty of reason, i.e. rationalism (a) is false.

Rationalism (b) contradicts the psychoanalytic account. On the Kleinian model, in specifying the motivations of a moral action, there is an essential reference to instinct and desire¹. Psychoanalysis denies that reason alone motivates us to act morally.

Rationalism (c) probably contradicts psychoanalytic naturalism. On the psychoanalytic account there is no reason to suggest that every moral action, that is every action whose motivations originate or bear a necessary reference to the super-ego, will be rational. Furthermore there will be ‘activities’² that count as moral but do not have rationalising beliefs. So we can certainly say that not every moral action (or activity) is rationalised. In order for the psychoanalytic account to contradict rationalism (c), however, we would also need to claim that a (non-rationalised) activity is not rational. Although I do not intend to give a definition of ‘rational’ here, I think that, given any reasonable such definition, a (non-rationalised) activity will

¹ Gardner (1992b) p. 45, points out that is a philosophically open question as to whether an unconscious phantasy must be present on each and every case of a particular emotion, or whether it is enough that the phantasy is present for the evolution of a natural emotion kind, after which it is no longer needed. This point can equally be made about moral motivation. One possibility is that every moral action requires the causal contribution of an instinctual impulse. Another is that unconscious instinctual impulses ‘set up’ kinds of moral sensibilities, emotions and motivational states, but are not necessarily present on each and every occasion. Both of these options contradict rationalism (b) because there will be *at least some* moral actions which are not motivated solely by reason. Importantly, these moral actions *require* the motivating desire or instinctual impulse. As Gardner also points out, the Kleinian account (at least in the case of emotions) takes the second of these views, see Klein (1959).

² See section three.

probably fail to qualify as rational. In fact, activities are more likely to fall into the class of irrational phenomena¹.

From the above discussion we can draw the following conclusion: if the psychoanalytic account of the development of moral sensibilities in the individual is true, rationalism (a), (b) and (c) are all false².

Thirteen

This section might be considered as something of a departure. For the most part, it leaves behind talk of the various models of development and concentrates on a particular metaethical question. This question is the status of value ascriptions. When we ascribe value to something in the world, what is it that we are doing? Are we describing primary qualities of things? Describing secondary qualities of things? Or are we *endowing* things with particular qualities? According to Richard Wollheim³ psychoanalytic theory has a positive contribution to make to this debate. This section is devoted to evaluating that claim and assessing exactly what the contribution might be.

Wollheim's account of value relies on the mechanism of projection. We have come across this notion at various points so far but I have not given a detailed account of it. One of Wollheim's most important contributions to this issue is the distinction he makes between two types of projection, simple and complex. Simple projection is an unconscious phantasy of expulsion whereby a mental disposition, or state, is imputed onto another person. Thus, as Wollheim says, 'I hate him' becomes 'he hates me'. This phantasy serves the purpose of an alleviation of anxiety. There are two ways in which this can occur, either the mental disposition, or state, is a source of anxiety and its projection reduces this source, or the disposition, or state, is felt to be

¹ Gardner gives a definition of irrational phenomena as follows, "a person exhibits irrationality when he does not, or could not...think about himself in a way that would both make sense of his own thought and/or action, and at the same time avoid exhibiting incompleteness, incoherence, inconsistency, lapse into unintelligibility, or some other defect of a kind to signify, in a suitably broad sense, *self-contradiction*." Gardner (1993) pp. 3-4.

² Bearing in mind the proviso above, concerning the relation between 'rational' and 'rationalised'.

³ The texts necessary for an understanding of Wollheim's position are, Wollheim (1980), (1984) pp. 213-8, (1987) pp. 82-5, and (1993f). As will be noticed in these texts, there is, for Wollheim, an essential parallel between ascriptions of moral value and ascriptions of aesthetic value.

under attack from bad internal objects and its projection is a protective measure. In either case anxiety is diminished. The result of projection is a change in the beliefs of the individual, a new belief arises, the content of which is that some other person in the environment possesses the very same mental disposition, or state, that has been projected.

Complex projection is best described by way of those aspects of it that differ from simple projection. Firstly, in complex projection, the projection is not onto a person but onto a natural part of the environment, something which does not itself possess a psychology. Secondly, in complex projection it is not just the individual's beliefs that are altered but the actual experience is altered. The individual gains a new way of experiencing the world. Thirdly, in complex projection the property ascribed to the part of the external world is not the very same property that was projected (as is the case with simple projection). This new property ascribed to a part of the external world can be called a 'projective property', and it is described either by a novel term or by the same term as the projected property, but in this latter case the old term is used metaphorically. So, to take one of Wollheim's aesthetic examples, when I project my own state of melancholy onto a landscape, I do not experience the landscape as literally melancholic, but '*of a piece with*' my melancholy¹. However, complex projection is not haphazard². For complex projection is triggered by the awareness of a match between the mental disposition, or state, to be projected and the part of the outer world onto which it is to be projected³. Like simple projection, complex projection describes an unconscious phantasy of expulsion, but this is only the first stage. This initiating expulsive phantasy sets up within the individual a disposition to phantasise in a particular way which leads him to experience the world in a particular way⁴. So, to return to the aesthetic example, the match between the

¹ The view that we ascribe the very same psychological property to the landscape is called by Wollheim the 'Predication view'. His argument against it can be found in Wollheim (1993f) pp. 145-8. At its most persuasive the argument claims that although a landscape can correspond in the correct way to depression, we do not and would not call a landscape depressed.

² As is, it would appear, simple projection. Cf. Wollheim's remark that almost anyone can turn out to be a paranoid's enemy, Wollheim (1984) p. 214.

³ This match may seem to be mysterious, but as Wollheim points out, (1993f) p. 154, it is difficult to see what else could be said to describe the type of affinity that holds, without invoking the projective property involved.

⁴ This dispositional aspect of the theory allows it to account for cases in which I experience projective properties but lack the corresponding mental disposition, or state. It is, however, true on this theory that

landscape and my melancholy triggers off an expulsive phantasy, aimed at the alleviation of anxiety, the result of which is that I experience the landscape as being *of a piece with* my melancholy. I express this by saying that the landscape is melancholic. This phantasy sets up, within me, the disposition to experience the world in this way so that the same landscape, or similar landscapes, will tend to evoke in me the same, or similar, responses.

This, then, is Wollheim's account of simple and complex projection. It remains to be seen exactly how this connects with the status of ascriptions of value. According to Wollheim, the central claim of chapter XII, section 9, of *The Thread of Life* is that

The assignment of value is...projection of a complex form, which on the level of judgement is represented by the application of a new predicate introduced for this very purpose...it originates in the projection of archaic bliss, of love satisfied. (Wollheim 1984) p. 215.

It is clear that, if correct, this claim will have important consequences for the metaethical question I posed at the beginning of the section. The idea that ascription of value is an instance of the complex projection of love satisfied, I will call 'psychoanalytic projectivism'. Before I examine the ramifications of psychoanalytic projectivism, I should like to briefly address a challenge that has been put to it by Marcia Cavell, in her paper 'Knowing and Valuing: Some Questions of Genealogy'¹.

Cavell asks, of complex projection, whether there is any important difference between saying that *x* is suitable to support a perception of the world as *y* (which is required by psychoanalytic projectivism), and saying that *x* is suitably perceived as *y*. She thinks that there is no such important difference and concludes that the mechanism of projection is doing little explanatory work and that the perception of *x* as *y* can be regarded as due to the nature of *x*. To admit this, however, is to accept realism about value. An initial response might be as follows: although Cavell's

I must *have had* the corresponding mental disposition, or state, at some point, otherwise the disposition could not have been set up.

¹ The point that I am concentrating on is only one of many that Cavell makes in her paper. Her main thesis is that there exists a line of thought, within the writings of both Freud and Klein, supporting realism about values, as opposed to psychoanalytic projectivism. So far as I understand her position, it appears to rely on a cognitivist account of emotions, whereby emotions are explained by reference to the possession of certain concepts. As I pointed out in section nine, I think this approach is mistaken. Incidentally, this cognitive account is the main reason for her scepticism concerning the young age at which Klein attributes phantasies and emotions to infants.

suggestion appears to be economical, it is metaphysically extravagant, in so far as it postulates a realm of value properties. Added to this is the problem the realist faces in providing an account of the perceptual mechanisms involved in ‘acknowledging’¹ the property of value. This problem does not arise for the psychoanalytic projectivist. Secondly, and I think this point is persuasive, what Cavell regards as a weakness in psychoanalytic projectivism can be regarded as a strength. Psychoanalytic projectivism is an anti-realist theory, but it is also a quasi-realist theory: it attempts to explain, and to justify, the ‘realist’ appearance of ordinary moral thought². Psychoanalytic projectivism explains why it is that we usually think of value as a real, objective property of things, but it does so *as an anti-realist* theory. So, the fact that leads Cavell to suggest failure on the part of psychoanalytic projectivism is actually that aspect of the theory which explains ‘the realist mistake’ that we all tend to make.

What, then, is the relation between psychoanalytic theory and the realism/anti-realism debate in moral theory? What contribution does psychoanalysis make? Firstly, psychoanalytic projectivism (a theory which can be gleaned from Kleinian psychoanalytic theory) is clearly an anti-realist theory. If value is the projection of an internal mental disposition, or state, onto an external object, then value is neither a primary nor a secondary property of objects³. Thus, if psychoanalytic projectivism is true, both ‘realism’ and ‘objectivism’⁴ are false. Psychoanalytic projectivism is a quasi-realist theory. There are certain positive contributions that psychoanalytic projectivism makes to quasi-realism. The first of these is that the causal history of the experience of value refers to both the external object (a property of which triggers of the projection) and to specifiable states of the subject. Since these states of the subject⁵ are claimed by psychoanalytic theory to be universal, these two causal factors help explain the high degree of agreement in ascriptions of value that realism would predict. The second contribution that psychoanalytic projectivism makes to quasi-realism is that it specifies the actual projective mechanisms at work, and it does so in a way that embeds those mechanisms in a broader theory. It is not the case that the

¹ See Cavell (1992) p.78.

² The project of quasi-realism is explained in Blackburn (1984), ch. 6, and (1985).

³ That is, of course, unless one takes the, admittedly minority, view that secondary properties are themselves projected in some sense.

⁴ This is Wollheim’s terminology, see Wollheim (1980). ‘Realism’ is the view that value is a primary property, ‘objectivism’ is the view that value is a secondary property.

⁵ Unconscious phantasies of love satisfied.

mechanism of complex projection has been postulated in order to support the anti-realist project in moral theory (or aesthetics). The idea that internal mental dispositions, or states, are projected onto aspects of external reality is an integral part of both Freudian and Kleinian theory. Thus, the projective mechanisms that do the work in psychoanalytic projectivism have a certain level of plausibility independent of any success they might have in forming a theory of the status of value ascriptions.

John McDowell is unimpressed with the quasi-realist's claim that he can justify our apparently realist commitments on a projectivist picture¹. He claims that projective properties are not real in the same sense that either primary or secondary properties are, yet our practice is to treat them as if they were. This means that projectivism cannot do justice to the phenomenology of value. Thus, if we accept the 'realist' phenomenology, we must be either realists or error theorists². Can psychoanalytic projectivism help the quasi-realist overcome this objection? According to psychoanalytic projectivism, a particular property of an external object displays a certain 'match' with a psychological condition and this triggers the projection of that psychological condition onto that external object. What the subject then experiences is a projective property. Now, the projective property (in our case value) *is not* the very same psychological property projected (love satisfied) and neither is it the 'matching' property of the external object, for this is simply the surface onto which the projective property is painted. But, we must admit that even though the projective property is not *the same* as either the projected property or the 'matching' property of the external object, it is intimately related to them. Each particular experience of a projective property is caused, via the mechanism of projection, by these two other properties. It is the fact that the projective property is, in part, caused by the external object that justifies our saying that it is a property *of* that object, that object plays an essential causal role in our experience. This, I think, answers McDowell's objection. I return to the point made by Cavell, there is no important difference between saying that *x* is suitable to support a perception of the world as *y*, and saying that *x* is suitably perceived as *y*. But saying that *x* is suitably perceived as *y* is to do justice to the 'realist' phenomenology.

¹ McDowell (1985) p. 124, n. 4.

² Mackie is the most well known advocate of the 'error theory', see Mackie (1977) Ch. 1. McDowell (1985) defends, contra Mackie, a view of values as secondary properties.

To sum up, there is a theory concerning the status of value ascriptions, psychoanalytic projectivism, that can be gleaned from psychoanalytic theory (particularly that of Klein). This theory is an anti-realist one, but it is also quasi-realist. In fact, psychoanalytic projectivism has several contributions to make to the project of quasi realism. The most important, and unique, of these is the actual specification of the sophisticated projective mechanisms involved. This must be viewed as a positive contribution that psychoanalytic theory makes to moral theory.

Fourteen

Speaking of the frontiers of philosophical discourse, Richard Wollheim has said that, “instead of disputing where the frontiers should run, one is more profitably engaged in showing how the existing lines may be safely transgressed” (Wollheim 1993b) p. 39. One of the things that I have been trying to achieve throughout the course of this paper is just this. Instead of taking the boundary between moral philosophy and moral psychology as rigidly defined, I have attempted to show how the two are interrelated and how, to a certain extent, the boundary is blurred. I have been concentrating on the results that considerations of moral psychology can have for moral philosophy. But it should be equally clear that philosophical considerations can, and should, have implications for moral psychology. I think that moral philosophy has much to learn from moral psychology and vice versa.

In section nine I suggested that the approach to moral philosophy which takes conceptual analysis as primary is mistaken and can, at best, provide only a rather superficial account of moral phenomena. My specific claim was, and remains, that a full understanding of moral concepts (such as guilt) cannot be gained from conceptual analysis alone. Moral concepts have complex developmental histories and to ignore, or deny the importance of, this aspect is, to that extent, to falsify them. In section nine, we saw David Jones using conceptual analysis to deny the moral psychologist access to moral philosophy and its concepts. To my mind, this is an illegitimate move. Hopefully I have shown that it is much more productive to allow the moral psychologist access to the domain of moral philosophy. In this way moral theory is enriched, the opposite approach can be very detrimental.

In sections ten and eleven I showed that the Kleinian model of the development of conscience in the individual provides us with a plausibly sophisticated naturalistic theory of moral motivation. Furthermore, psychoanalytic naturalism can be seen as both a development of, and an improvement on, Hume's theory. As such, psychoanalytic naturalism is an important and formidable position in moral theory. It simply would not do to exclude it on the grounds that it is 'not really philosophy'.

Sections twelve and thirteen also provide us with evidence of moral psychology's (and, in particular, psychoanalytic theory's) contribution to moral theory. If, as I have argued, the Kleinian model can be seen as a plausible contender for a correct theory of moral development, there are some important results in metaethics. The first is that rationalism, in any of the three senses identified, is false. The Kleinian model is incompatible with rationalism, and I think that there is good reason to accept the Kleinian model as true. Of course, that the two positions are incompatible will be seen by some philosophers as good reason for rejecting the Kleinian model. Which theory one rejects will be determined by independent factors. My reasons for accepting the Kleinian model (and therefore rejecting rationalism) constitute sections two through seven of the present paper. In section thirteen, I tackled the question of the status of ascriptions of value and concluded that psychoanalytic theory has a positive contribution to make. Psychoanalytic projectivism is a powerful and, to my mind, plausible, theory. If psychoanalytic projectivism is the correct account of value (and on this point I am less committal), realism is false and the cause of quasi-realism is furthered. The description of the projective mechanisms involved in ascriptions of value is an invaluable contribution psychoanalysis makes to moral theory.

In section one, I said that I will have succeeded in my aims if I answer two questions: What reasons do we have for supposing that the psychoanalytic account of moral motivation is true? And, if correct, what consequences does the psychoanalytic account of moral motivation have for moral philosophy? In sections two through seven I have provided an answer to the first, and in sections nine through fourteen I have provided an answer to the second of these questions.

Bibliography

Blackburn, S, *Spreading the Word: Groundings in the Philosophy of Language* (OUP 1984).

Blackburn, S, 'Errors and the Phenomenology of Value' in Honderich (ed), *Morality and Objectivity: A Tribute to J. L. Mackie* (Routledge 1985).

Brenner, C, *An Elementary Textbook of Psychoanalysis* (International Universities Press 1955).

Brown, J, *Freud and the Post-Freudians* (Pelican 1961).

Campos, J. Barrett, K. Lamb, M. Goldsmith, H. and Stenberg, C, 'Socioemotional Development' in Mussen (ed), *Handbook of Child Psychology*, vol. 2 (New York 1983).

Carey, S. and Spelke, E, 'Domain Specific Knowledge and Conceptual Change' in Hirschfeld and Gelman (eds), *Mapping the Mind: Domain Specificity in Cognition and Culture* (CUP 1994).

Cavell, M, 'Interpretation, Psychoanalysis, and the Philosophy of Mind' in *Journal of the American Psychoanalytic Association* Vol. 36, 1988.

Cavell, M, 'Knowing and Valuing: Some Questions of Genealogy' in Hopkins and Savile (eds), *Psychoanalysis, Mind and Art: Perspectives on Richard Wollheim* (Blackwell 1992).

Cavell, M, *The Psychoanalytic Mind: From Freud to Philosophy* (Harvard University Press 1993).

Church, J, 'Morality and the Internalised Other' in Neu (ed), *The Cambridge Companion to Freud* (CUP 1992).

Damon, W, 'The Moral Development of Children' in *Scientific American* Vol. 281, no. 2, 1999.

Darwall, S., Gibbard, A. and Railton, P, 'Toward *Fin de Siècle* Ethics: Some Trends' in Darwall, Gibbard and Railton (eds), *Moral Discourse and Practise: Some Philosophical Approaches* (OUP 1997).

Davidson, D, *Essays on Actions and Events* (Oxford 1980).

Deigh, J, *The Sources of Moral Agency: Essays in Moral Psychology and Freudian Theory* (CUP 1996).

Farrell, B, *The Standing of Psychoanalysis* (OUP 1981).

Freud, S, *Standard edition of the complete psychological works of Sigmund Freud*, 24 vols. (London 1981). Referred to as *S.E.*

- (1900) *The Interpretation of Dreams*, S.E. 4-5.
- (1905) *Three Essays on the Theory of Sexuality*, S.E. 7.
- (1908a) 'Character and Anal Erotism', S.E. 9.
- (1908b) "'Civilized" Sexual Morality', S.E. 9.
- (1911) 'Formulations on the Two Principles of Mental Functioning', S.E. 12.
- (1912-13) *Totem and Taboo*, S.E. 13.
- (1914) 'On Narcissism: An Introduction', S.E. 14.
- (1915a) 'Thoughts for the Times on War and Death', S.E. 14.
- (1915b) 'Instincts and their Vicissitudes', S.E. 14.
- (1915c) 'Repression', S.E. 14.
- (1915d) 'The Unconscious', S.E. 14.
- (1916) 'Some Character Types Met with in Psycho-Analytic Work', S.E. 14.
- (1916-17) *Introductory Lectures on Psycho-Analysis*, S.E. 15-16.
- (1917a) 'A Metaphychological Supplement to the Theory of Dreams', S.E. 14.
- (1917b) 'Mourning and Melancholia', S.E. 14.
- (1920) *Beyond the Pleasure Principle*, S.E. 18.
- (1921) *Group Psychology and the Analysis of the Ego*, S.E. 18.
- (1923) *The Ego and the Id*, S.E. 19.
- (1924a) 'The Economic Problem of Masochism', S.E. 19.
- (1924b) 'The Dissolution of the Oedipus Complex', S.E. 19.
- (1925a) 'Negation', S.E. 19.
- (1925b) 'Some Psychological Consequences of the Anatomical Distinction between the Sexes', S.E. 19.
- (1926) *Inhibitions, Symptoms and Anxiety*, S.E. 20.
- (1927a) *The Future of an Illusion*, S.E. 21.

- (1927b) 'Humour', *S.E.* 21.
- (1930) *Civilization and its Discontents*, *S.E.* 21.
- (1933a) *New Introductory Lectures on Psycho-analysis*, *S.E.* 22.
- (1933b) *Why War?*, *S.E.* 22. (With Einstein).
- (1939) *Moses and Monotheism*, *S.E.* 23.
- (1940a) *An Outline of Psychoanalysis*, *S.E.* 23.
- (1940b) 'Splitting of the Ego in the Process of Defence', *S.E.* 23.

Friedman, L, 'Psychoanalysis and the Foundations of Ethics' in *The Journal of Philosophy* 53, 1956.

Furer, M, 'The History of the Superego Concept in Psychoanalysis: A Review of the Literature' in Post (ed), *Moral Values and the Superego Concept in Psychoanalysis* (International Universities Press 1972).

Gardner, S, 'The Unconscious' in Neu (ed), *The Cambridge Companion to Freud* (CUP 1992a).

Gardner, S, 'The Nature and Source of Emotion' in Hopkins and Savile (eds), *Psychoanalysis, Mind and Art: Perspectives on Richard Wollheim* (Blackwell 1992b).

Gardner, S, *Irrationality and the philosophy of psychoanalysis* (CUP 1993).

Gardner, S, 'Psychoanalysis, Science, and Commonsense' in *Philosophy, Psychiatry, & Psychology* 2, 1995.

Gilligan, C, *In a Different Voice: Psychological Theory and Woman's Development* (Harvard 1982).

Gillman, R, 'Preoedipal and Early Oedipal Components of the Superego' in *The Psychoanalytic Study of the Child* 37, 1982.

Glymour, C, 'Freud, Kepler, and the Clinical Evidence' in Wollheim and Hopkins (eds), *Philosophical Essays on Freud* (CUP 1982).

Grünbaum, A, 'Is Freudian Psychoanalytic Theory Pseudo-Scientific by Karl Popper's Criterion of Demarcation?' in *American Philosophical Quarterly* 16, 1979.

Grünbaum, A, *The Foundations of Psychoanalysis: A Philosophical Critique* (University of California Press 1984).

Habermas, J, *Knowledge and Human interests*, trans. Shapiro (Boston 1971).

- Harman, G, 'Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error' in *Proceedings of the Aristotelian Society* 99, 1999.
- Hinshelwood, R, *A Dictionary of Kleinian Thought* (London 1991).
- Hinshelwood, R, 'Commentary on "Psychoanalysis, Science, and Commonsense"' in *Philosophy, Psychiatry, & Psychology* 2, 1995.
- Hopkins, J, 'Introduction: Philosophy and Psychoanalysis' in Wollheim and Hopkins (eds), *Philosophical Essays on Freud* (CUP 1982).
- Hopkins, J, 'Synthesis in the Imagination: Psychoanalysis, Infantile Experience and the Concept of an Object' in Russell (ed), *Philosophical Perspectives on Developmental Psychology* (Blackwell 1987).
- Hopkins, J, 'Epistemology and Depth Psychology: Critical Notes on *The Foundations of Psychoanalysis*' in Clark and Wright (eds), *Mind, Psychoanalysis and Science* (Blackwell 1988).
- Hopkins, J, 'The Interpretation of Dreams' in Neu (ed), *The Cambridge Companion to Freud* (CUP 1992a).
- Hopkins, J, 'Psychoanalysis, Interpretation, and Science' in Hopkins and Savile (eds), *Psychoanalysis, Mind and Art: Perspectives on Richard Wollheim* (Blackwell 1992b).
- Hume, D, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals* (OUP 1975).
- Hume, D, *A Treatise of Human Nature* (OUP 1978).
- Isaacs, S, 'The Nature and Function of Phantasy' in Riviere (ed), *Developments in Psychoanalysis* (London 1952).
- Jones, E, *The Life and Work of Sigmund Freud* (edited and abridged by Trilling and Marcus) (Penguin 1961).
- Jones, D, 'Freud's Theory of Moral Conscience' in *Philosophy* 51, 1966.
- Kant, I, *Groundwork of the Metaphysics of Morals* trans. M. Gregor (CUP 1997).
- Klein, M, 'Symposium on Child Analysis' (1927a), in *Love, Guilt and Reparation and other works 1921-1945* (Vintage 1998).
- Klein, M, 'Criminal Tendencies in Normal Children' (1927b), in *Love Guilt and Reparation and other works 1921-1945* (Vintage 1998).
- Klein, M, 'Early Stages of the Oedipus Conflict' (1928), in *Love, Guilt and Reparation and other works 1921-1945* (Vintage 1998).

- Klein, M, 'Personification in the Play of Children' (1929), in *Love, Guilt and Reparation and other works 1921-1945* (Vintage 1998).
- Klein, M, 'The Early Development Of Conscience In The Child' (1933), in *Love, Guilt and Reparation and other works 1921-1945* (Vintage 1998).
- Klein, M, 'A Contribution To The Psychogenesis Of Manic Depressive States' (1935), in *Love, Guilt and Reparation and other works 1921-1945* (Vintage 1998).
- Klein, M, 'Notes On Some Schizoid Mechanisms' (1946), in *Envy and Gratitude and other works 1946-1963* (Vintage 1997).
- Klein, M, 'On Observing the Behaviour of Young Children' (1952), in *Envy and Gratitude and other works 1946-1963* (Vintage 1997).
- Klein, M, 'On the Development of Mental Functioning' (1958) in *Envy and Gratitude and other works 1946-1963* (Vintage 1997).
- Klein, M, 'Our Adult World and its Roots in Infancy' (1959) in *Envy and Gratitude and other works 1946-1963* (Vintage 1997).
- Kohlberg, L, *Essays on Moral Development, Volume II, The Psychology of Moral Development: The Nature and Validity of Moral Stages* (New York 1984).
- Laplanche, J, and Pontalis, J, *The Language of Psychoanalysis*, trans. Nicholson-Smith (London 1973).
- Mackie, J, *Ethics: Inventing Right and Wrong* (Penguin 1977).
- Mackie, J, *Hume's Moral Theory* (Routledge 1980).
- Matthews, G, 'Concept Formation and Moral Development' in Russell (ed), *Philosophical Perspectives on Developmental Psychology* (Blackwell 1987).
- McDowell, J, 'Values and Secondary Qualities' in Honderich (ed), *Morality and Objectivity: A Tribute to J. L. Mackie* (Routledge 1985).
- Meltzoff, A. and Moore, M, 'A New Foundation for Cognitive Development in Infancy: The Birth of the Representational Infant' in Scholnick, Nelson, Gelman and Miller (eds), *Conceptual Development: Piaget's Legacy* (LEA 1999).
- Moore, G. E., *Principia Ethica* (CUP 1959).
- Mounce, H, *Hume's Naturalism* (Routledge 1999).
- Nietzsche, F, *Daybreak* (CUP 1997).
- Norton, D-F, 'Hume, Human Nature, and the Foundations of Morality' in Norton (ed), *The Cambridge Companion to Hume* (CUP 1993).

- Piaget, J. and Inhelder, B, *The Psychology of the Child* (New York 1969).
- Pigden, 'Naturalism', in Singer (ed), *A Companion to Ethics* (Blackwell 1991).
- Popper, K, *Conjectures and Refutations* (Routledge 1989).
- Rich, J. and DeVitis, J, *Theories of Moral Development* (Charles C Thomas Publisher 1985).
- Rycroft, C, *A Critical Dictionary of Psychoanalysis* 2nd edn. (Penguin 1995).
- Ryle, G, *The Concept of Mind* (New York 1949).
- Sachs, D, 'In Fairness to Freud: A critical notice of *The Foundations of Psychoanalysis*, by Adolph Grünbaum' in Neu (ed), *The Cambridge Companion to Freud* (CUP 1992).
- Sandler, J, 'The Concept of Projective Identification' in Sandler (ed), *Projection, Identification, Projective Identification* (London 1988).
- Schafer, R, 'The Loving and Beloved Superego in Freud's Structural Theory' in *The Psychoanalytic Study of the Child* 15, 1960.
- Scheffler, S, 'Naturalism, Psychoanalysis, and Moral Motivation' in Hopkins and Savile (eds), *Psychoanalysis, Mind and Art: Perspectives on Richard Wollheim* (Blackwell 1992a).
- Scheffler, S, *Human Morality* (OUP 1992b).
- Segal, H, *Introduction to the work of Melanie Klein* (London 1978).
- Snare, F, *Morals, Motivation and Convention: Hume's Influential Doctrines* (CUP 1991).
- Spelke, E. and Van de Walle, G, 'Perceiving and Reasoning about Objects: Insights from Infants' in Eilan, McCarthy and Brewer (eds), *Spatial Representation: Problems in Philosophy and Psychology* (Blackwell 1993).
- Thomas, L, 'Morality and Psychological Development' in Singer (ed), *A Companion to Ethics* (Blackwell 1991).
- Walker, L, 'Sex Differences in the Development of Moral Reasoning: A Critical Review' in *Child Development* 55, 1984.
- Walker, L, 'Sex Differences in Moral Reasoning' in Kurtines and Gewirtz (eds), *Handbook of Moral Behaviour and Development*, Volume 2, (Lawrence Erlbaum Associates Publishers 1991).

Walker, L. Pitts, R. Henning, K. and Matsuba, M, 'Reasoning about Morality and Real Life Problems' in Killen and Hart (eds), *Morality in Everyday Life: Developmental Perspectives* (CUP 1995).

Williams, B, *Ethics and the Limits of Philosophy* (Fontana 1993).

Wollheim, R, 'The Mind and the Mind's Image of Itself' in *On Art and the Mind* (London 1973).

Wollheim, R, 'Art and Evaluation' in *Art and its Objects* (2nd Edition with Six Supplementary essays) (CUP 1980).

Wollheim, R, *The Thread of Life* (CUP 1984).

Wollheim, R, *Painting as an Art* (Thames and Hudson 1987).

Wollheim, R, *Freud*, 2nd edn, (Fontana 1991).

Wollheim, R, 'The Sheep and the Ceremony' in *The Mind and Its Depths* (Harvard University Press 1993a).

Wollheim, R, 'The Good Self and the Bad Self: The Moral Psychology of British Idealism and the English School of Psychoanalysis Compared' in *The Mind and its Depths* (Harvard University Press 1993b).

Wollheim, R, 'The Bodily Ego' in *The Mind and its Depths* (Harvard University Press 1993c).

Wollheim, R, 'Desire, Belief and Professor Grünbaum's Freud' in *The Mind and its Depths* (Harvard University Press 1993d).

Wollheim, R, 'Crime, Punishment and "Pale Criminality"' in *The Mind and its Depths* (Harvard University Press 1993e).

Wollheim, R, 'Correspondence, Projective Properties, and Expression in the Arts' in *The Mind and its Depths* (Harvard University Press 1993f).