S E L F - D E C E P T I O N

A Dissertation Submitted to the Faculty of Arts, University of the Witwatersrand, Johannesburg, in partial fulfiment of the Degree of Masters of Arts by Coursework in Philosophy

I declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Arts by Coursework at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other university.

Signed: EJ Gates

Date: 30 January 1991

A C K N O W L E D G E M E N T S

I would like to thank my supervisor, Mark Leon, for his much appreciated assistance, advice, analysis, and encouragement. All remaining errors I claim as my own.

PREFACE

The topic of this dissertation is self-deception. It involves a general analysis of the problem; however, the central discussion centres on an agency-type approach to the problem, and in particular on a Freudian-type approach.

CONTENTS

INTRODUCTION

Self-deception is just one of the kinds of irrational behaviour that we as human beings seem so apt to indulge in. We all act irrationally at times - do things that are patently not in our best interests, behave in ways that contradict our basic assumptions about what we ought to do, decide on a plan of action and then act against it for no good reason, believe for a moment that the world really is the way we would like it to be. These and other types of behaviour are so commonplace that their very strangeness often goes unnoticed. Self-deception is one of such types of behaviour: like its close cousin, wishful thinking, it is so readily indulged in that we often live in blissful ignorance of either its existence or its effects. In fact, unlike many other kinds of irrational behaviour, self-deception relies on our not being aware of many of our beliefs and desires. Self-deception has a stake in remaining covert. In this paper I will examine the problem of self - deception, noting just where the problem lies, and discussing various approaches to a solution of the problem and where these approaches are problematic. I will also attempt to offer some kind of solution to the problem by examining how a Freudian approach to the nature of the mental can subsume under it the particular problem of self-deception.

The central assumption that underlies all problems of irrationality, including that of self-deception, is that humans are essentially rational creatures. Without this central assumption, there would be no problem with failures of rationality; failure of rationality is possible - and problematic - only for creatures who operate according to the dictates

of the rational. The principle of rational operation need not be conscious - it may be that someone will act with perfect rationality but will be unable to articulate the principles of reason on which she operated. This would not make her behaviour less rational, or irrational. In order for us to understand a creature's behaviour, we must know the principles under which it implicitly operates. If we can ascertain no such principle, then attributing beliefs to such a creature will be impossible.

The very notion of belief also implies rationality. Beliefs operate and make sense within belief systems of which they form a part. These systems operate on a principle of coherence; beliefs within a system must cohere with other beliefs of this system. If they do not, this will necessitate the creation or emergence of a new belief system. If two beliefs are recognized as incompatible, then either one, or both, of the beliefs will be dropped, and either a new belief will be formed, or the evidence at hand will be reviewed. It is not possible that a person consciously and knowingly believes two contradictory propositions. Thus the unity of the mental is implied in this notion of the coherent nature of belief systems. Since beliefs influence behaviour - often by acting as reasons without which the behaviour would be inexplicable, we are able to gauge the beliefs of an individual by analyzing her behaviour, and thus a third person may have access to some of the agent's beliefs, even if the agent herself is not aware of the operation of the belief. Beliefs are central to the problem of self-deception, with paradoxes of belief being at the heart of the issue.

At the centre of the problem lies a paradox. If one models self- deception on ordinary "other-person" deception, we have the following problem: if A deceives B with respect to a certain proposition p, this usually implies that A knows or believes that p is true (false), and wants to convince B that p is false (true) by deliberately misleading B - either

by introducing false evidence, or by withholding evidence, or by causing B to misinterpret the evidence at hand. There are some issues here that need to be noted and distinguished. The first is that the knowledge claim is not crucial to the problem: A need not know that her belief that p is either false or true. A's belief that p is sufficient - that is, A must have a cognitive relationship, as it were, with the belief that p. A could also be wrong about the truth value of p, and still be said to deceive B about p. So even if A thinks that p is true, and persuades B to believe that p is false, and p is false, A still deceived B with respect to p. I would also claim that it is sufficient that the mechanism of deception be in operation in order for us to have deception. Sufficient also for deception is that A has the intention to deceive B. Suppose that the deceptive plan failed to affect B in the desired way - we would still label the event as one of deception if A intended to deceive B, and involves himself in the deceptive plan. Deception may also be an ongoing process that requires constant adjustment to situations; it is not exhausted by a single action but may require further actions on the part of the deceiver. So in this sense deception is teleological: not only is it purposive, but it is an intentional action that is performed by a rational agent and it is directed towards a certain end. The agent is able to make judgements and alter her plans in accordance with the desired aim.

The full force of the problem arises when we attempt to carry the analogy over to self-deception. Here we would have a situation where A believes that p and at the same time tries to persuade herself that -p. This could occur in a number of ways: A may believe that p despite having evidence that supports -p and no or little evidence that supports p; or, the thought that -p might be true may motivate A to cause him to believe p. For example: Brenda believes that Bill is a faithful husband; she would be devastated by any acts of infidelity on his part, and would consider

divorcing him were he to be unfaithful. Bill is being unfaithful - he is having an affair with his secretary Bridget. Brenda ignores all the evidence: Bill's many late nights at the office, his weekends away on "business trips", the late night 'phone calls, his inattention to herself. But because the thought of Bill's infidelity would be too painful for Brenda, she ignores all the evidence at hand, and continues to hold onto her faith in his fidelity. In fact, Brenda's fear that Bill might be unfaithful - and her unsaid suspicion that he may indeed be having an affair with Bridget - motivates her to hold onto the desired belief, and constantly "look the other way", and her excuses for many of his actions often outdo even Bill's ingenuity.

In another case we may have something different, where the motivated belief is undesirable. Here we could have a suspicious husband who believes against all the evidence that his wife is being unfaithful to him. Here he would ignore all the indications of her faithfulness and concentrate on finding and magnifying any possible clues of her adulterous behaviour; Othello would be a good example of this type of behaviour. This example shows that self-deception may not always be motivated by desire, but may be motivated by fear or anxiety; the process should in any case be the same.

One other example should be sufficient. Here we have the case of someone who begins by holding the belief that p, but attempts to persuade himself to believe that -p, despite (or because of) evidence that an impartial observer would see as supporting p. Carl buys Caspar's old Chevrolet at a reasonable price. He realises that it is not in great condition but he wishes to use the engine and some other parts to build a jeep - Carl likes tinkering with cars. Carl's friend Clifford sees the Chevrolet, and, needing a used but reliable vehicle for his business, offers to buy it from Carl for a good price. Carl is unsure what to do, because he does

not think that the car is in a good enough condition to sell to his friend. Carl, however, looks over the Chev, and tinkers with it a bit, and finally decides that the car is in a better condition than he first thought. He finally is able to sell it to Clifford with a clear conscience. Here we have a case of self-deception where the person does not hold both p and -p simultaneously, but where he comes to replace one belief with its contradictory.

These examples are merely illustrations of the basic theme: we have situations where, given certain evidence, the subject should form the belief that p. More than this, there is the assumption that, in other circumstances, the subject would form the belief that p. (There is no assumption here that the subject is incapable of forming the correct or justified belief.) However, in these circumstances, the subject is strongly motivated by a desire to believe -p, and, as a result of this desire, does in fact form - and act upon - the belief that -p. The belief that -p may replace the belief that p, or the belief that -p may be formed instead of the belief that p. What is important to note is that the subject motivates for the belief that -p, in the face of and despite having access to evidence that should lead her to believe that p. I hope that the similarity between the cases presented is sufficiently clear for it to be seen that they all belong to the same species - self-deception - of the genus of irrationality. The task is to resolve the paradox.

Self-deception has very often been allied to wishful thinking, and, although there are similarities between the states, there are also important differences. Suppose that Anne likes Andy and wants to believe that Andy likes her too. She may indulge in some wishful thinking at this point, and imagine that Andy does in fact like her. She interprets Andy's welcoming smile and assistance at work as evidence of his love for her, despite the fact that he hasn't invited her on a date, or declared his love

for her. At this stage Anne's state of mind is at the level of wishful thinking. Were Andy to be seen by Anne to present flowers to Annabel, and to invite Annabel out on a date, then, were Anne to continue to believe that Andy loved her, she would be entering the state of self-deception. The more she ignored evidence that disproved Andy's love for her - suppose Andy told her he didn't love her - the more serious would be her state of deception. The central difference, then, between wishful thinking and self-deception is that the former is carried out at the level of fantasy whereas in self-deception the person deliberately ignores or avoids taking into consideration evidence that would contradict the desired belief.

There are also many conflicts in belief states that do not amount to self-deception. Often there may simply be intellectual indecision, such as where a person has evidence both for a proposition and its contradictory, and cannot decide between them; here the rational thing to do would be to suspend judgement. Or a scientist may have evidence for p but may desire to prove that -p; here she would set up a series of experiments that would prove -p and would show the limitations of the evidence that supports p. If she failed to establish any justification for -p, then she would have to give up her belief in -p; should she not do this, and should she persist in her belief that -p in the face of mounting evidence to the contrary, then her state of mind would be bordering on that of self-deception. The stronger the desire that motivates the holding onto a belief that, in the circumstances, is irrational, or the stronger the motivation to hold onto a desired but irrational belief, the more likely is the person to fall into a state of self- deception. It may also be the case that a person believes both p and -p simultaneously and yet is not self-deceived. It may be the case that the beliefs are situated in very different cognitive domains within the person's mind, and he is simply unaware of their connection. Someone may also not realise the

import of the available evidence, or may fail to draw the correct or obvious conclusion because he is ignorant or foolish. Here we would not have an example of self-deception. A person may also consciously know that p but allow himself to act as though -p were true. For instance, we all "suspend disbelief" when we enter into the theatre, and we may be moved to tears by the final death scene, knowing all the while that it is "only a play". Or we may attempt to grasp a stick in the water, knowing that it only appears bent and is not "really" where it seems to be. These cases would not be classified as self-deception - or even as irrational - because the agent is able to realise the process that he is undergoing, and the suspension of the true belief is only temporary, and is performed with the agent's consent.

In this paper I shall attempt an examination of the problem of self-deception. There are two broad ways of approach; one is to argue that the "deception" part of the problem can be reanalysed and redefined so as to resolve - or dissolve - the paradox. Deception is seen as some other and less problematic kind of problem within the formation of belief. So the approaches here say that what we have in the case of self-deception is not really deception at all, but rather some other kind of activity. I will argue that this type of approach is not wholly satisfactory. The second line of analysis is to focus on the "self" part of self-deception: what kind of self is necessary in order for self-deception to be possible? This line of enquiry will occupy the remainder of the paper. Chapter two will examine the Davidsonian line, noting the problems involved, and chapters three and four will be concerned to analyse the problem within a Freudian examination of the self. Chapter three will focus on Freud's topographic theory, and chapter four on the structuralist theory. Both theories are sufficiently different and complex to warrant individual attention. It may be admitted that the Freudian approach - and any other approach that posits the self as fundamentally split - is not needed for

r cases of self-deception. That is , there are several degrees of self-deception, and not all of them are equally problematic. Mild cases may be more superficially delt with. This may be so. It thus may be the case that the Freudian approach is more suitable for an analysis of the stronger cases of self-deception. Perhaps an understanding gained at this level will filter down and allow for greater clarity at lower levels. If not - if it is suitable only for harder cases of irrationality - an understanding gained at this level alone will be interesting.

ONE

DECEPTION: SOME APPROACHES

In this chapter I shall examine some of the attempts to resolve the paradox inherent in the problem of self-deception. These attempts focus on the acquisition of the conflicting beliefs, and attempt to provide a solution by reformulating the issue of deception within the parameters of the conditions under which beliefs are acquired and/or maintained. The attempt is to reformulate the problem without the contradiction; that is, the majority of these approaches define self-deception as not necessarily involving the paradox that A believes both p and -p simultaneously. Their claim is that this is not a necessary feature of self-deception, and hence that self-deception may lie in other, less "strong" cases. Various approaches within this area will be examined and assessed. It is worthy of note that these approaches all work within the framework of the rationality of the mental; that is, they assume that mental functioning is essentially a rational process, and indeed the very notion of a belief-system both assumes and implies an internal coherence or rationality. The problem of self-deception and other problems of irrational behaviour is precisely that they seem to contradict the essential rational workings of the mind. The following approaches attempt to solve the problem of self-deception by saving the appearance of rationality: self-deception does not in actual fact involve an internal contradiction, since we can have self-deception without having the subject believe both p and -p at the same time. Some of the arguments for this proposal will be set forth and examined.

be possible. So the overt paradox raised by the holding of two contradictory beliefs is shown to be an apparent one - no underlying contradiction exists once the apparent paradox has been reformulated. Audi does not attempt a solution along the lines of creating divisions - systems or agencies - within the mind, although he does postulate unconscious beliefs. He does not, however, see their existence as a solution to the paradox. Audi's approach centres on his distinction between believing that p, and avowing that p, and it is this aspect of his thesis that requires further examination.

It appears that Othello believes that he loves Desdemona - he believes that he believes that he loves her. But Othello has in the past loved Emilia, to whom he still feels attracted. Othello wants to deny his attraction to Emilia - he is an honourable man, and wants to feel that Desdemona claims his full devotion and attention. So he cannot admit to himself that he is attracted to Emilia - something that a perceptive third person observer would be able to infer from Othello's behaviour. So the honourable Othello has an unconscious belief - namely that he is attracted to Emilia - which he disavows. Shorthand, O believes that p but avows that -p. At this point, we could attribute to Othello both the belief that p and the belief that -p, and attempt to resolve the paradox by making B[O(p)] unconscious, and B[O(-p)] conscious. Audi's move is not to do this, but rather to claim that O's avowal that -p does not amount to a belief that -p, and hence no contradiction ensues.

In Audi's account, avowal may normally be sufficient for, but does not entail, belief, and hence we can have O's avowal that -p without this implying that O believes that -p. Audi also claims that O's avowal is sincere; what this means is that O is, firstly, aware of his avowal that -p, and he is not lying or otherwise trying to deceive a third party. Secondly, the sincerity of the avowal implies that Othello believes that

Robert Audi[1] argues that self-deception does not involve a central paradox or multiple agents, but does involve unconscious beliefs. An outline of his position can be summarized as follows: subject A is in a state of self-deception iff:

(i) A has unconscious knowledge that -p.

(ii) A avows sincerely that p.

(iii) A has at least one want which explains, in part, both why the unconscious knowledge that -p is unconscious, and why A is disposed to disavow her belief that -p, and to avow p, even when presented with what she sees as evidence against p.

Audi hopes that his proposal will successfully explain the following:

(1) The analogy with other person deception: Audi intends to preserve the analogy.

(2) The notion of refusing to admit to oneself: A really does believe the proposition she disavows. Central to Audi's argument is the distinction between avowal and belief.

(3) The self-deceiver's capacity to respond to evidence: A's belief states act in the way ordinary belief states act -they are evidence sensitive. It is only A's states of avowal that do not, but as they are not belief states, this does not present a contradiction. The only strange notion in this regard is that the belief state formed in response to the evidence is unconscious.

The central problem in Audi's account is his distinction between states of belief and avowal, a distinction central to his argument. This will need to be examined to determine whether such a distinction can successfully be made.

Audi's argument wants to preserve the rationality of the system. This would mean that no internal inconsistency within the unified system would

---

[1] Audi, R "Self-Deception and Rationality" in Martin, M. W (Ed) pp 169-194

he believes that he is not attracted to Emilia - he has a second order belief about his avowal, but, since it is merely avowal, there is no corresponding lower order belief. We may assume that this is characteristic of states of avowal; we may also assume that the higher order belief would be necessary to states of avowal - the notion of unconscious avowal seems contradictory in the light of Audi's account. We could thus say that, in this case, a state of avowal is equivalent to that of a conscious second order belief with no necessary corresponding first order belief. The state of self-deception is characterized by the simultaneous existence of B[O(p)] and A[O(~p)], or by B[O(p)] and by B<B[O(~p)]>. In fact, we could go on to say that, in the case of self-deception, B[O(p)] ==> A[O(~p)]. That is, the unwelcome unconscious belief motivates S to consciously avow the contradictory of this belief. So the unconscious belief that p, together with the desire that ~p, motivates the agent to form the second order belief that he believes that ~p.

Audi does not attempt a resolution of the problem by appealing to the conscious / unconscious distinction. Here we would accommodate the contradictory beliefs that p and that ~p by positing one as conscious and the other as unconscious. But we would still have the situation of two contradictory belief states being held by the agent. By introducing the mental attitude of avowal, Audi hopes to avoid the paradox posed by conflicting beliefs. In his paper "Self-Deception, Action, and Will", Audi has the following to say on the matter:

> [the subject] operates, as it were, at two levels. On the unconscious level, at which he knows p is false ... he gets himself to "believe" p ... and since he sincerely avows p and is not conscious that he believes ~p, it looks quite as if he did believe p. I think we thus preserve enough of the analogy to account for the term "deception" in both cases; yet we avoid the paradox implied in saying the self-deceiver believes something he knows (or believes) is not true. My point is not that such a belief is impossible; but how it is possible, if it is, is not clear, and we can account for self-deception without positing such beliefs.

(p 141)

One characteristic that Audi asserts belongs to belief and not to avowal is the influence on behaviour. On this account, the evidence for the existence of belief states and states of avowal would differ. The overall weight of behavioural evidence would support states of belief rather than states of avowal. Also, belief states would be evidence sensitive, whereas states of avowal would not be. This would be one of the reasons that a person could avow that p even in the face of evidence to the contrary. So, for instance, we attribute the belief that he loves Emilia to Othello because Othello "not only exhibits embarrassment around Emilia, but lavishes unusual attention on Desdemona ... and protests too much both regarding his attraction to Desdemona and concerning his immunity to the charms of Emilia." ("Self-Deception and Rationality" pp 173/4). Still, it may be the case that much of Othello's behaviour could be taken as evidence for his states of avowals instead of his belief states. For instance, his lavishment of attention on Desdemona would count as behavioural evidence of the love that he avows. Certainly Othello would see his behaviour in this way, and would furthermore act lovingly towards Desdemona in an attempt to convince both himself and others that his avowals of love are genuine. Audi would, however, see the difference as residing in the fact that most of the behavioural evidence would be evidence for Othello's belief state that he loves Emilia rather than for his state of avowal of his love for Desdemona. The argument is that we would always be able to find additional evidence that would confirm the existence of the belief state that p, despite that avowal that ~p.

The main problem with Audi's account is that it may not avoid the paradox it attempts to circumvent. Audi's states of avowal can be translated into belief-talk. When O avows that p, he forms a higher order belief -

B<B[O(p)]>. Since O is self- deceived, we have B[O(-p)] but not B[O(p)]. So, the agent must not come to form either of the following beliefs: either B[O(p)] or B<B[O(-p)]>, as this would simply reintroduce the paradox. In the latter case, the subject must not come to form the higher order belief that he believes that -p. So the belief that -p must not be allowed to become conscious, nor must there be any awareness of this belief that may cause the agent to form a second order belief about it. So there must be a motivated block to forming this higher order belief. We thus have the problem of explaining how the motivation for the prevention of this second order belief arises. The point is that the subject does not have to avow that -p precisely because he has an undesirable belief that -p. So the avowal that p works as a deterrent against the avowal that -p only insofar as there already exists the belief that -p. Here the paradox emerges at a higher level - the subject must know what it is that he wishes to conceal in order to conceal it more effectively. It is not simply that this higher order belief that -p does not emerge, but its non-emergence is motivated, and it is this motivational aspect here that would need to be explained. It is not clear from Audi's account how the higher order belief that -p is prevented from emerging in the subject. We would need to know, for instance, if there were any mechanism that existed to mitigate against its formation, and how such a mechanism would work.

On the other side of the coin we would also have to have the non- existence of the lower order belief that p, since, were this belief to be formed, there would arise conflict between it and the already existing belief that -p. So, for instance, Othello's avowal of his love for Desdemona does not have a corresponding belief, because Othello believes that he loves, not Desdemona, but Emilia. Were Othello to come to believe that he loves Desdemona, we would have a paradox of belief existing at this level. The contrary belief - that he loves Emilia and not Desdemona (L[O(E)] &

-L[O(D)]) - must prevent the formation of the lower order belief that he loves Desdemona and not Emilia (-L[O(E)] & L[O(D)]) without itself becoming conscious. Audi's account must be able to explain how the belief that -p, say, can act as a reason against the formation of the belief that p without itself becoming conscious. There is the danger of the paradox emerging once again over here - there is motivation against the undesirable belief that -p becoming conscious. So, in sum, there are several problems or paradoxes that Audi's account still needs to resolve. The first is the basic one concerning the relationship between states of avowal and states of belief held by the agent: given that O believes that -p, how can he avow that p? The second arises at the level of the inter-relationship of belief states: the unconscious belief that -p must block the formation of the belief that p without itself becoming conscious. The question is - how is this possible? This second problem is closely related to another: there is a motivational block against the formation of a higher order belief about the unconscious belief that -p - i.e. there is a motivational block against the formation of the avowal that -p. How does the existence of the belief that -p motivate against the formation of the avowal that -p? So the problems of self-deception that Audi wanted to avoid seem to re-emerge in a different form and at another level in his argument.

Bela Szabados[2] gives us an analysis of what it means to be in a self-deceived state. It is, basically, to be disposed to ignore or subvert evidence that would support an unwelcome belief, and perhaps, in addition, pursue evidence that would support the desired belief. It is important to note that to be in a state of self-deception is to be motivated to so

---

[2] Szabados, B "Self-Deception" Canadian Journal of Philosophye 4 1974 pp 51-68

corrupt the belief-forming procedure, and it thus involves a constant adjustment by the system to cope with changes of evidence. This distinguishes self-deception from something like wishful thinking where we have a state that is less involved with reality and that is less tenacious in the face of contradictory or unwelcome evidence. The wishful thinker might dream about the new raise or the love the new woman at the office feels for him, but will relinquish these thoughts if snubbed constantly by his boss or the woman; he may continue to daydream, but will confine his thoughts to the world of phantasy and will not act as if they were true.

The self-deceiver on the other hand attempts to shape reality to fit his desires, either by ignoring conflicting evidence, or by misinterpreting it. Again, the tenacity of the self-deceiver in the face of overwhelming evidence indicates a motivational component of self-deception. It is not necessarily - or even usually - the case that there be a deliberate plan of action that the person systematically tries to follow. This may not even be possible, given the fact that the self-deceptive process is usually unconscious, and hence remains unarticulated by the agent. A deliberate plan of action usually requires that the agent recognize the desired end result, as well as the procedures required for arriving there, and act accordingly. Still, the act of self-deception is intentional or deliberate in the sense that it is motivated by the agent; the agent's actions are directed towards the holding of the desired belief, or against the realization or acceptance of the undesired one. In the case of self-deception we have the situation of the agent persuading herself to believe that a certain state of affairs exists despite her prior belief or suspicion that this desired state of affairs does not exist. The procedures for such self persuasion may be complicated, and would not simply be the result of a mechanical procedure inherent in the agent, although such mechanical processes might come into play once the agent

entered into the self-deceptive procedure. Also, the agent may often not opt for self-deception, but may decide to face the unpleasant truth despite a strong desire to believe the opposite. The self- deceptive course may not be overtly chosen - the very notion of self-deception may not allow this type of choice - but it is a motivated course of action that the agent may, at a later stage, come to change. Although self-deception may be an activity that humans are predisposed to indulge in - and so is often entered into automatically - one may also choose not to be self-deceived. Were the procedure of self-deception simply mechanistic, it is unlikely that this breakdown in the process would regularly occur; instead there would be a regular turning away from belief if there were a sufficiently strong desire to act as a cause. Furthermore, the agent often is aware of the outcome of the deceptive process; she indulges in it precisely in order to achieve a certain state of mind. This goal-directed, teleological-type behaviour also suggests that something other than a mechanical process is at work. This motivational action is usually unconscious, or the real motivation for a particular action will not be acknowledged by the agent. The question as to how the agent can be involved in a motivational plan that is not at the same time conscious and intentional is, I think, at the centre of the problem; perhaps the solution to this problem will provide a solution to the problem of self-deception.

Szabados argues that in order for us to have self-deception, it is sufficient for the subject to change her beliefs over a period of time. Now of course not just any change of belief states will do - it is very often the rational plan to alter beliefs according to changes in evidence, and it is something we do all the time. Something must be added to this description to give us self-deception. Szabados gives us the following story: suppose there is a merchant sea-man who wishes to hire out his vessel to transport cargo on a long journey. He has some misgivings about

the condition of his ship, but alleviates them by telling himself that he is over-reacting - that he is being too cautious. He may also ask one of his friends to examine the ship for him, knowing that the outcome will be a favourable one. So his suspicion that the ship is not seaworthy is transformed into a trust that all will go well on the journey, and his prior fears are allayed. When he hears months later that the ship has gone down at sea, he is able to collect the insurance money with a clear conscience, absolving himself of all blame.

Here we have a case of self-deception without the simultaneous holding of both p and -p by A, and hence the paradox is resolved. The sea-man's belief that the ship is indeed seaworthy is held by him only after he has relinquished his belief that it is not. He deceives himself all right, because he avoids gathering evidence that he suspects might indicate that his vessel is falling apart, and he bases his latter belief on evidence and judgement that he suspects might be faulty. (We shall avoid, for the time being, the problems inherent in the use of such words as "suspects"). But the state at the end of the whole process - the deceived state, as it were - is not marked by the struggle of one belief over another; indeed, the merchant appears to be conflict free.

Szabados's position also raises other problems, which, due to their similarity, will be discussed in conjunction with Mele and other accounts. Mele too[3] gives an account of self-deception which involves conditions other than that of holding two contradictory beliefs simultaneously. He cites four ways of self-deception:

---

[3] Mele, A. R Irrationality: An Essay on Akrasia, Self-Deception and Self-Control

(i) negative misinterpretation, where A avoids or misinterprets evidence that would count against the desired belief that p.

(ii) positive misinterpretation, which involves A interpreting data in a more positive way with respect to p than would otherwise be the case if A were an unbiased observer.

(iii) selective focusing / attending, where A "fail[s] to focus his attention on evidence that counts against p and ... focus[es] instead on evidence suggestive of p."

(iv) selective evidence gathering, where A overlooks easily obtainable evidence in support of p and goes in search of less accessible evidence in support of p.

Now according to Mele, none of these fairly common ways of deceiving oneself involves the holding of two contradictory beliefs by the agent. For Mele, self-deception is characterized by the holding of a desired belief by an agent, who, in other circumstances* would not come to hold the belief. The holding of two contradictory beliefs is not a necessary characteristic of self- deception. Mele also argues that the agent need not have, in the past, believed the negation of the proposition that he now desires to believe; part of the self-deceptive process may involve the prevention of the formation of the undesired belief. Mele sets out four conditions that he claims are sufficient for self-deception. They are:

(i) The belief that p which S acquires is false.

(ii) S's desiring that p leads S to manipulate (i. e. to treat inappropriately) a datum or data relevant, or at least seemingly relevant, to the truth value of p.

(iii) This manipulation is a cause of S's acquiring the belief that p.

(iv) If, in the causal chain between desire and manipulation or in that between manipulation and belief-acquisition, there are any accidental inter-

---

* i. e. in circumstances similar to this one in all regards except that here the truth of that same belief is not desired by the agent.

> mediaries (links), or intermediaries, intentionally introduced by another agent, these intermediaries do not make S (significantly) less responsible for acquiring the belief that p than he would otherwise have been.

Condition (iv) refers to any external conditions that may enter into the decision making process, and influence S's belief acquisitions. The more that external factors influence S's belief acquisition, the less central a case of self-deception we would have. However, the point here is that the mere existence of accidental factors that may contribute to the agent's becoming self-deceived do not, in themselves, make him less responsible for becoming so deceived.

So the central point of Mele's account is that the state of self-deception is characterized by the acquisition of a false but desired belief that, in other circumstances, the agent would not hold. The desire for the truth of the belief causes the agent to manipulate data in such a way so as to build up a case for the desired belief, and ignore or otherwise misinterpret evidence for the undesired belief. So, for instance, a man may misinterpret evidence of his wife's infidelity by thinking about her virtues - she is a devoted mother who is very fond of her children, say - and by interpreting her current strange behaviour as indicative simply of his wife's new interest in her hobbies, or some other innocuous activities. Here, according to Mele, there is no holding of two contradictory beliefs by the agent.

Mele's account of self-deception does still not make it clear how the state is arrived at. The conditions given seem to describe an end state, and the paradox of self-deception may arise at the level of the process. The conditions described by Mele above are irrational and indicative of self-deception only if the agent is aware of the existence (or the possibility of the existence) of evidence against the desired belief that

p. If an agent desires that a certain proposition p be true, he may concentrate his attentions on acquiring evidence in favour of p without thereby acting irrationally. The action becomes irrational only if the agent suspects that there may be evidence in favour of -p, and then refuses to consider it. If this is the case, then it appears as if a form of the paradox emerges lower down. Here we would have the situation where S desires p, but realizes or suspects that -p is more likely to be true, and yet persuades himself - by whatever means, including those outlined by Mele - to hold p. It may be that here the form of the paradox cannot be translated into the form of "S believes both p and -p simultaneously", because it may be that S suspects that -p is more likely to be true, and thus prevents the formation of the belief that -p. What is still puzzling, though, is the procedure that the agent must undergo in order for the evidence for -p - and that against p - not to enter into the decision making process, as it were. We have, on the one hand, the belief that p and, on the other hand, either the belief that -p or the evidence that, in other circumstances, would lead the agent to form the belief -p. What is of interest here is: what enables the agent to form the belief that p despite the evidence to the contrary? How is it possible that the evidence for -p can be so ignored? What are the underlying conditions that allow for the possibility of such a phenomenon? Even if Mele has given us the minimal conditions for self-deception, it seems to me that characterizing the problem as one merely of belief acquisition does not dissolve the paradox, but forces us to look further down the process and confront the paradox, or the tension, that arises <u>there</u>. The criteria what he gives us are used within the notion of self-deception rather than being explanatory. For instance, in order for S to ignore evidence and focus selectively, he must already have a notion of what belief state he wishes to avoid being in, and must thus work against acquiring the relevant belief. Mele's analysis does also not tell us <u>how</u> for instance, S's desire that p is able to cause S to ignore evidence that S recognizes as

supporting -p. The interesting aspect of the problem - the mechanism of self-deception - is left out of the picture. I think that a Mele-type approach needs to be subsumed under another type of approach that examines the nature of the self that is involved in the deceptive process.

Another problem with Mele's accounts of self-deceptive behaviour outlined above is that they may be characteristic of types of irrational behaviour other than that of self-deception. For instance, an obsessive researcher intent on showing the world that his theories are true, and who wants to receive fame and accolades for his work, may obsessively pursue evidence in favour of his theory and, from sheer arrogance, may ignore or otherwise dismiss data that could support a rival theory. The researcher vehemently believes that he is right - he does not regard the rival theory as being well-supported by the evidence - and although his actions are certainly irrational, he could not be characterized as deceiving himself. This again shows, I think, that the agent must be in possession of - and be aware of the significance of - the evidence supporting the belief contrary to the one he forms. If this is so, then we encounter the problems discussed above.

Both Szabados and Mele's accounts seem unable to account for the problem of self-deception. It may be true that, overtly, A may not appear to hold two contradictory beliefs simultaneously, but I think that there is a paradox further down that is not accounted for. The problem is that there is not simply a change from p to -p, say, but that the beliefs are not altered without conflict, as it were. Let us take the case of the sea-man and quote Szabados here:

> There is a stage ... where the cherished and favoured belief is challenged - evidence is brought forth against it. The person in question sees where the evidence points and he experiences conflict; he is plagued by doubt. It is at this stage that he can naturally be said to believe p and not to believe p.

It seems to me that this is the crucial point that the above theses do not address. Before the self-deceiver can arrive at the desired belief state, he has to go through a stage of deceiving himself; here he must be "aware" of his desire that p be true, and he must at the same time be "aware" of the evidence that suggests the contrary. In the words of Sartre, the agent "must know the truth very exactly in order to conceal it more carefully", and the above approaches do not seem to solve this problem. Hence the need for an explanation or description of the process that would act as the underlying condition of possibility of self-deception. Such a process would explain how the agent is able to move from his holding of the belief that ~p, to his holding of the desired belief that p, all the while being aware that ~p is more likely than p to be true. In the above approaches, the problem of the process that the agent must go through is not elaborated upon. We are not told, for instance, if the deception is made possible by the separation of the beliefs in some way, or how the desire for p is able to cause the agent to ignore or otherwise manipulate evidence for ~ p. Such an avoidance of this issue enables the paradox to emerge at a lower level, and the sting of the problem is not removed.

I would like to mention here briefly one or two other proposals as solutions to the problem. These proposals, have, I think, been adequately dealt with in the literature, but as they first attempted a resolution of the problem, they deserve some mention. Also, by their failures, they indicate what self-deception is not and in this way make their contribution. The first of these proposals is that of Canfield and Gustavson.[6] Their suggestion is that self-deception is simply the acquisition of a

---

[6] Canfield, J. V and Gustavson, D. F "Self-Deception" Analysis 23 1962/3 pp 32-36

belief that p in circumstances "adverse" to the formation of such a belief. Such belief-adverse circumstances would be ones in which the evidence against p appears to be "overwhelming" - for instance, if a man believes in the innocence of his son despite evidence that points to his son's guilt. The problem with this proposal is that it does not distinguish sufficiently between self-deception and other cases of ignorance, stupidity, failure to see the import of evidence, as well as cases of blind faith or other kinds of irrational behaviour. For instance, in the above example, the man may believe in his son's innocence not because he is deceiving himself, but because he has faith in his son's character, a fact about himself which he recognizes. Indeed, his faith about his son may turn out to be justified if indeed his son is innocent. Canfield and Gustavson also don't seem to deal adequately with the case of someone who knows that p but deceives himself that -p. "What is odd about [these cases]" they say " ... is that, although the person knows a proposition p, it would take drugs, or drink, or hypnotism, or psychoanalysis to get him to admit it ... even to ... himself." They attempt to explain this by saying that we have two different senses of "know" here -"the sense in which he knows he is not a great artist ... is quite different from the sense in which he would be said to know this if he were trying - deliberately - to deceive us into thinking that he is a great artist ..." But they postulate no theory as to what these two different types of knowledge claims are, or how they differ. They do not postulate an unconscious as a solution to the problem of self-deception, and so it is difficult to see how postulating two different types of knowing, by itself, solves the problem.

Another objection to the Canfield and Gustavson thesis is that two people could form the same belief in the same "belief adverse circumstances", one of whom we might say was deceiving herself and one of whom we might say the opposite, depending on the outcome. For instance, suppose two

robbers are fleeing the police; they come to a high wall, and attempt to jump over it. Each robber is small and unfit, and in normal circumstances would not be able to jump such a wall. It is also dark and the wall is slippery. Now one robber manages to scale the wall, urged on by his fear - a factor which he had not taken into account when assessing his chances of escape. The second robber slips on a piece of grass as he tries to jump and is nabbed by the police. Now since the robbers both formed the belief in the same circumstances, either both are self-deceivers or neither is. On Canfield and Gustavson's account, both should be equally guilty of self-deception. Since the one robber did in fact manage to jump over the wall, he was not deceiving himself when he thought he could do so. He formed a belief, acted on it, and was proved right. It seems inappropriate to attribute self-deception to either party; both believed p in circumstances adverse to the belief that p; one person was right, and one was wrong, but neither seems to have been deceived. Forming a belief in belief- adverse circumstances does not seem sufficient to account for self-deception. What we need for self-deception is that the belief be formed by the agent despite the fact that he believes the contrary of the belief, or that he realizes that the contrary of the belief is more likely to be true. This may occur in "belief-adverse circumstances", but the phrase as used by Canfield and Gustavson does not capture the paradox or the tension specific to self-deception.

Another proposal by Raphael Demos[6] suggests a different kind of approach. Demos rejects both the thesis that in self-deception "believing p and disbelieving p occur at different and successive times" and the thesis that "the agreeable belief occupies the conscious mind while the unpleasant one is repressed into the unconscious." Demos instead proposes

---

[6] Demos, R "Lying to Oneself" Journal of Philosophy 57 1960 pp 588-595

that "the belief and the disbelief are simultaneous and both exist in the consciousness of the person."[7] Demos attempts to achieve this by dividing our mental activities into "noticing" and "ignoring". This distinction can be highlighted in the following way: if I have a headache, say, and go the movies in an attempt to "forget" about the pain, then while I am engrossed in the film, I fail to notice that I have a headache, and notice it again once the movie is over. It is not that the headache has gone away, as it would have if I had taken medicine, but that my attention has been directed away from it and onto something else. Similarly, says Demos, with self-deception; the deceiver "fails to notice or ignores what he knows to be the case." Demos's account has the problem of explaining what is meant by a conscious belief that is "ignored" or "unnoticed". In Demos's account we are presented with the notion of an "unfelt pain". This seems a very strange concept indeed, and I am not sure if it is intelligible. It seems that our understanding of pain depends not so much on whether there is something physical going on in the brain - c- fibres firing, or whatever - but rather on the phenomenological "feel" of the experience. If there is no feeling of pain, then it seems safe to say that the subject is not in pain, regardless of whether the c-fibres are firing. (Similarly, a feeling of pain is sufficient for the subject to be in a pain-state, even if no c-fibres are firing.) So Demos's example seems to involve strange notions of "unfelt pains", and as this is his central case for his account, it casts doubt on the entire account.

More generally, the notion of two conscious beliefs that are unaware of one another, or rather of which the subject is unaware, seems very strange

---

[7] Demos, R ibid p 592

indeed,* unless we propose a radical split in the subject, such as one caused by a commissurotomy. At any rate, this type of exploration is beyond the limits of this paper. Suffice it to say that usually we do not encounter two consciousnesses within a subject, and the proposal appears to be fundamentally flawed. So Demos's proposal - and one's similar to it - does not solve the problem at hand.

It seems that if the notion of self-deception is to be at all intelligible, we must have the situation of either having the opposing beliefs occurring at different times, or we must have that the beliefs are not simultaneously conscious. The first of these proposals does not by itself do the trick, for, although it illustrates how the beliefs can be held by the agent, it does not clarify the process used within the deceptive procedure that brings about the desired mental state. We need to know how the agent is able to persuade herself to come to believe the contrary of the belief that she now holds, given that this contrary belief is known or believed by her to be either false, or to have less probability of being true than the belief she now holds. I think that there must be some kind of awareness by the agent of the tension involved in the acquisition of the desired belief -p in place of the belief that p. The exact nature of this "awareness" may be difficult to characterize, but an attempt will be made later on in the paper. At any rate, it is connected to the motivational aspect of self-deception. The point is that the thesis that the contradictory beliefs are not held simultaneously by the agent does not in itself solve the problem; it must, I think, be

---

* There is one way in which A could hold two contradictory conscious beliefs simultaneously, and that is if, for some reason, A does not realize the connection between these beliefs. This may occur if the beliefs arose in different contexts, or are connected to different areas within A's belief system. In this case, however, we have ignorance or misunderstanding and not self-deception.

subsumed under the investigation of the nature of the agent. This shall be attempted in later chapters. The latter proposal - the one concerned with the conscious nature of beliefs - will also be examined in greater detail at a later stage.

T W O

THE SELF: DAVIDSON

Solving the problem of irrationality may involve postulating a division or split within the mind, but there are various ways that such a division may be postulated. We can have the division along the lines of a conscious/unconscious distinction, or along the lines of agencies that operate according to certain principles; both of these methods are central to Freudian theory, and will be analyzed in the following chapters. In this section, I would like to explore another way of dividing the mind. This procedure is one developed by Davidson, but it has been examined by other philosophers of mind.* The Davidson Thesis (also called the Functional Theory by Pears) does not divide the mind along the divisions of consciousness, nor does it postulate separate permanent agencies within the mind whose workings define and constitute the very nature of the agent. The Davidson partitions are more fluid and less permanent than their Freudian counterparts, and their mode of functioning is different. It will be necessary to examine some of the background arguments to Davidson's thought before the nature of the partitions can be examined.

The theory is centred around the finding of an adequate reason for a belief. According to Davidson, an irrational belief is characterized by

---

* See Davidson, D "Paradoxes of Irrationality" in Wollheim and Hopkins (Eds), and "Deception and Division" in Elster, J (Ed). Also see Pears, D Motivated Irrationality and Mele, A Irrationality

having been caused by a mental event that is not at the same time a sufficient reason - in terms of providing sufficient evidence - for the belief. For instance, my desire that I have passed the test is not in itself a sufficient reason for my belief that I have indeed passed it. Some other criteria are needed - such as evidence relating to how I performed on the test. One can distinguish between having a reason to believe in a proposition, and "having evidence in the light of which it is reasonable to think the proposition true."[18] It may be argued that my desire that I passed the test is indeed a reason for my believing it to be true. That is, I may have a good reason for believing in the truth of pleasant beliefs, namely my desire that such beliefs be true. John Heil takes this view, and hence expands somewhat Davidson's account of what we should take as a reason. We should therefore modify the analysis, and say that an irrational belief is one that is caused by a reason that does not count as evidence for the truth of the belief.

Defining exactly what is meant by "evidence" is tricky, and much of the discussion could be pursued within the realms of epistemology. Evidence could mean the totality of evidence that is available with respect to a proposition p. In this case, most beliefs concerning p will be irrational, because such totality of evidence is not available to most enquirers. The term could also refer to all the evidence available to a person, but even here the demands seem to be too stringent, because an enquirer may miss some relevant details and yet not form beliefs that are irrational in the circumstances. We might limit the term to refer to all the evidence that is available to the enquirer, the importance of which is understood by her. So if Suzy, a scientist skilled in weather prediction, believes that it will be warm and sunny for her party the next

---

[18] Davidson, D "Deception and Division" in Elster, J (Ed) p 85

day, despite scientific evidence of which she is aware that points to a rainy and cold day, we would classify her belief as irrational. Her belief that it will be warm and sunny is caused by the wrong sort of thing - her desire - in the face of a cause of the right sort - the evidence - for the opposing belief. The minimum basis for irrationality of any kind would be the formation of a belief in a way such that it subverts the normal process of rationality, the agent believes that p in the face of sufficient evidence that -p, however we decide to define "sufficient evidence". Davidson's aim is to show how the agent can come to form such an irrational belief, given that the general mode of functioning of the agent is one of rationality. This feature is important, because, as Davidson insists, irrationality is a failure within the realms of rationality, and presupposes it. A non-rational creature cannot be accused of being irrational, because the term simply does not apply. A creature who is rational must operate within or under a principle of total evidence requirement. This principle governs the formation of beliefs in the light of evidence, and says that, all things considered, it is better to act in accordance with the total evidence that is available to the agent. Acting in accordance with the total evidence principle does not imply an awareness of the principle; an agent may conform to the dictates of the principle despite an inability to formulate it. It may simply be the case that we just do operate under the dictates of this principle, even although its workings remain implicit. So an irrational belief is one held by the agent in the face of the available evidence, and against the total evidence principle.

The formation of a belief that is caused by an event that is not a reason for it presents no paradox in the case of two individuals. Here we would have the desire that p be true held by one person (A), and the effect of that desire could be an action on the part of another person (B). Suppose that A desires to have a new wrist- watch for her birthday. A knows that

B will buy her a new one if he thinks that A's wrist-watch no longer works correctly, or that she has lost it. A pretends to lose her wrist-watch, and hints to B that she would like a new one for her birthday. B buys the wrist- watch. Here we have the desire that p be true acting as a cause for the resulting action. The problem presents no puzzle because the desire - and accompanying plot - were formed by A, whereas the action was performed by B. Here A's desire acted as a cause but not as a reason for B's behaviour - B may have been blissfully unaware of A's desire, and may have thought that his gift of a wristwatch would come as a great surprise to A. So if we postulate the relation between desire and event as occurring between two different minds, then we have no paradox. Davidson wishes to take this analogy and transfer it to the case of a single individual - hence his postulating the existence of a division within the mind. There are three characteristic features that underlie this picture: the fact that the mind can be partitioned, the existence of "a considerable structure in each quasi-autonomous part, and the non- logical causal relations between the parts."[11] We shall now turn to examine the nature of these mental partitions, and the principles under which they operate.

Davidson does not specify the number of divisions that exist within the mind, nor does he say anything about their permanence, or how they arise. In this respect he differs fundamentally from Freud. The compartments into which the mind is divided are also not separate and indivisible - partitions may overlap, and an element may belong to one or more partition simultaneously. The principle according to which partitions exist or are created is the following: a belief is assigned to a separate partition if it cannot interact rationally with the belief system of another par-

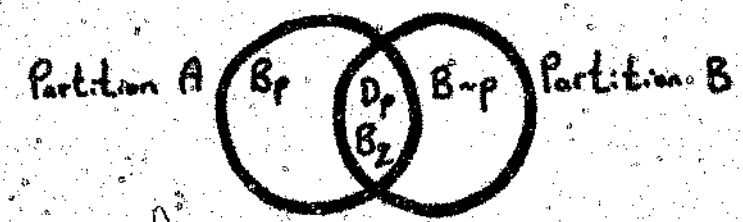
---

[11] Davidson, D "Paradoxes of Irrationality" in Wollheim and Hopkins (Eds) p 304

tition. The borders of a subdivision are defined by the breakdown of the rational relationship between beliefs, or between desires and beliefs. A line is drawn in order to prevent a belief or desire from producing its effect in the main system; the irrationality of the system as a whole is accommodated by the creation of sub-divisions that are themselves rational. The irrationality of the system occurs precisely because the evidence that would lead to the conclusion that -p, say, does not intervene, and the opposite (and desired) belief is instead formed. The forming of p is not internally irrational - that is, within its partition, there are no other beliefs that conflict with the belief that p. The irrationality occurs at the level of the entire system or agent. The divisions occur in order to both ensure an internal rationality within the sub-divisions, but they also allow for the irrationality of the system as a whole. This must be so, as the partition occurs precisely where there is, and in order to accommodate, irrationality. Wherever there is irrationality in the system as a whole there must exist rational subdivisions or partitions that allow for this.

One way in which an irrational relationship between beliefs and desires may exist is if a desire acts as a cause for a belief without at the same time being a sufficient reason for that belief. For example: Suzy wishes to believe that Sam is in love with her and does not love Sandra. This wish causes her to believe that Sam loves her, and it causes her to interpret some of the things he says or does as proof of his love. Sam's inviting Sandra to the dance and his giving her flowers are ignored by Suzy, or interpreted as acts to make her jealous. What is irrational here - and what causes the creation of the partition - is not only the wish that p, but the wish that p in the face of the evidence that -p. A wish is seen as an irrational cause of a belief by Davidson because a wish that something be the case is not in itself sufficient reason for believing it to be the case. But the wish that p and the belief that p can live

happily side by side. What would necessitate the formation of the division would be the holding of the belief that -p, or the holding of evidence that would lead to the belief that -p. The division would come about to accommodate this other belief in order that the irrational but desired belief may be held by the agent. In this case we would have the belief that Sam loves Suzy (call the propositional content of the wish p) assigned to one compartment, and the belief that -p assigned to another. The evidence that supports -p is kept apart from the belief that p, and in turn the irrational belief that p cannot be accommodated within the same system that supports -p. Now there will also be many beliefs that will be shared by the compartments - for instance beliefs relating to Sam and to Sandra that will be necessary for the formation of either p or -p. Such beliefs will belong to the overlap of the divisions. The picture will look something like this:



The divisions operate along the lines of a functional thesis; beliefs are assigned to compartments according to the way they function within the belief system. The thesis makes no direct use of the conscious / unconscious distinction in its anadysis of partitions. A sub-system harbouring the undesired or conflicting belief need not be unconscious. The agent could be in a position where the connection between the opposing beliefs is not apparent, and so their overt contradiction need not be addressed by the agent, despite the fact that both beliefs are conscious. Some other kinds of explanations may be given. The thesis is neutral, and not hostile, to the postulation of unconscious mental states, but argues that it is sufficiently explanatory without entering into the conscious / unconscious debate. If such a move is necessary, it is probable that the thesis would be able to accommodate it without relinquishing too much.

At any rate, it is a problem that lies outside the main domain of the thesis itself.

Now the assigning of a belief to a sub-system isolates it from the main system, but does not necessarily prevent the belief from affecting the agent. What this means is that the assigning of beliefs to different systems may not prevent the so-called irrational belief from having the desired effect. So, for instance, the rational belief that -p does not intervene and prevent Suzy from forming the irrational belief that Sam loves her. What merely happens is that the beliefs are assigned to different systems within the mind, and they can co-exist quite happily as long as they are not forced to interact. The principle that is at work here is one of internal rationality - each system is governed by its principle of rational coherence. Furthermore, the functional theory asserts that whenever there is an irrationality that the person is able to avoid, a split is created, and a sub-system appears to harbour the otherwise incompatible belief. Either the person is able to avoid the formation of the irrational belief, or it will be assigned to a sub- system. What this means is that the nature of the split is defined in terms of the irrationality of the belief that p, relevant to the main (or another sub-) system: if p cannot be accommodated into the main system, subsystem P is created to accommodate p.// So the nature of schisms or systems within the functional thesis gets its explanatory power by definition. In this way the functional thesis is problematic; it is unable to offer any independent corroboration of its thesis, unlike, say, the Freudian thesis, which was able to get independent evidence in support of its postulation of an unconscious area within the mind. In the functional thesis, a belief is assigned to a sub-system if it is irrational - that is, if it cannot be accommodated into the main belief system - but the sub-system came into existence only so as to be able to accommodate the belief. To quote Pears here: " So the thesis, that no degree of avoidable

irrationality is possible without a schism, will be true by the definition of the word 'schism'."[12] The way to avoid the circularity would be to define a system or a schism independently of its ability to accommodate an otherwise irrational belief - that is, to find independent corroboration for the existence of the partitions.

The question to ask is - how serious is this apparent circularity? One disadvantage for the thesis is that if we are able to find a rival thesis that explains all that the functional thesis explains and in addition is able to offer additional reasons that support its account of mental partitions, we would be inclined to forgo the functional thesis in favour of the new one. Another problem is that the lack of an account of the nature of the partitions - other than a definitional one - prevents us from asking questions about the nature and origins of the partitions. This would in turn not make it possible to ask questions about the operations of the partitions - we are not given any information regarding the mechanism of operation of the partitions. For instance, is there any method of censorship between the partitions? This may be an important consideration in, for instance, the area of self-deception. For example: A may deceive herself with respect to a certain proposition p. She wants to believe that p is true, while all the while suspecting that -p is probably true. We have the creation of a partition at the point of irrationality: the belief or suspicion that -p, together with its supporting evidence is relegated to one partition, while the wish that p - and the subsequent belief that is formed - are relegated to another system. So far we have a picture of how the irrationality can be accommodated. But A still feels some unease: she believes that p - the desired state of affairs - is true, but she still harbours the suspicion that

---

[12] Pears, D Motivated Irrationality p 70

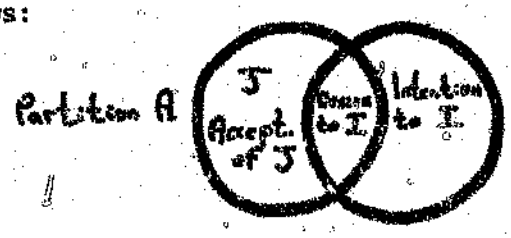
maybe -p. In order to hold onto the desired belief, motivation for the prevention of the undesired belief -p in the main system must exist. That is, the functional thesis is able to paint a picture of the accommodation of opposing beliefs in separate partitions, but it is not able to explain how, if at all, the partitions arise or interact (or fail to interact). In this example, we are not told how the belief that -p is prevented from making an appearance in the main system, only that it does not do so. Surely the role of the division is to ensure non-interaction of partitions - that is what a division means. But in a situation such as self-deception, we appear to have a motivational aspect working against forming the undesirable belief in the main system. It is not clear whether this can be accounted for by the functional thesis, because here the mere existence of the partitions should be able to accommodate and account for the contradiction of the beliefs. But if only this were the case, there should be no element of tension within the agent once the partitions have been created, and no further motivation for keeping the two beliefs separate should exist. Since it appears that such motivations do indeed exist, this may present problems for the functional account. So this lack of a fuller picture of the working mechanism of the partitions prevents us from asking questions that seem important, especially with regard to self-deception.

There are further objections raised against Davidson's thesis. John Heil argues[13] that the divisions postulated by Davidson do not explain behaviour adequately enough, and that we often must break down the compartmentalizations in order to get a true picture. Consider, for instance, the following example: A desires to insult B. A judges that it would be best not to insult B, but nevertheless insults him. This would

---

[13] Heil, J "Minds Divided" Mind Vol XLVII Oct'89

be a clear case of irrational intention or action. The picture would be as follows:



Now, in order to identify the irrationality, we must identify a cause of the intention to insult that is not a sufficient reason for the intention. In the above example, the desire to insult is both a cause and a reason for the intention to insult, and so the intention is not irrational. The irrationality occurs at the level of the interaction of the various partitions: the intention to insult is irrational only when taken in conjunction with the judgement that it would be best not to insult B, together with the operation of the principle that one should act according to one's better judgement. Heil argues that in order for the irrationality to appear we may have to look across the sub-systems to see the full picture. So in the above example, the irrationality is the result of the forming of the intention to insult, on the basis of the desire to insult, and in conjunction with both A's judgement that it would be best not to insult B, and the principle that one should act in accordance with one's better judgement. The point is that in order to appreciate the irrationality of the system, we must look at the entire picture - we must look across the divisions. This brings us back once again to the problem raised earlier - we are not given a sufficiently complex picture that enables us to see the interactions of the systems. Davidson's picture is concerned with the breakdown of a potentially irrational state of affairs into minor compartments that are themselves rational, and so not problematic. What is not explained is how these systems then interact in order for the agent to form an irrational belief or perform an irrational action. The point is this: if the agent acts in accordance with the beliefs and desires of the main system alone, her action will not be irrational. Irrationality enters only when we combine the beliefs of the

main system with the sub-divisions created precisely to harbour the conflicting belief. If the agent is not able to see across the boundaries of the partitions - if this viewpoint is available only to an outsider - then the action or belief of the agent cannot be irrational. If the agent can see across the partitions, then her action would be irrational; such a perspective would, in effect, collapse the partitions. The question is: if the agent is not able to see that the belief formed contradicts a previously formed judgement and an underlying principle to act in accordance with one's better judgement, in what sense is the agent self-deceived? On the other hand, if the agent is aware, at some level, of the tension between various belief systems, how can this be accommodated by the functional thesis? It is important to note that, in the case of self-deception, the agent acts on a desired belief in the face of evidence to the contrary, and it must be evidence that the agent herself knows would count against performing the action or forming the belief. Neil's example once more highlights problems with the functional thesis raised earlier. If we must look at the interaction between the sub-systems in order to fully appreciate the irrationality of the belief formed, or of the action performed, then we must ask how the various systems are able to interact.

Another problem is that there must exist quite a large overlap between some of the systems, especially those that relate to the same belief-content. Thus the partition that holds Suzy's belief that Sam loves her will have many beliefs in common with her belief that Sam loves Sandra. There may arise here tensions in the main belief system, such as when one of the shared beliefs has more connections with the belief that ~p in the sub-system than with the belief that p. More and more beliefs in the person's main system may be relegated to the sub-system, thereby creating large gaps in the person's belief system.

Alfred Mele[14] wants to argue against Davidson's thesis by showing that we may not have irrationality when we have a belief or an action caused by a mental event that is not also a reason for the belief or action. So Mele argues against Davidson's conditions for irrationality. Among the examples given by Mele are those that illustrate types of behaviourist techniques, where the person wants to avoid doing something that is harmful, and so engages in activity that prevents him form performing the action, not by any method of logical persuasion, but by some process of association. For instance, a person intent on giving up smoking may persuade himself to think of something pleasant - way of floating in the Caribbean - whenever he develops a craving for a cigarette. This relaxes him sufficiently for him to be able to control his desire for a cigarette. Here we have a mental cause of another mental item that is not itself a reason for the item. Mele's argument against Davidson is that there is often no need to postulate divisions within the mind to accommodate cause-effect pairs that are not at the same time reason-result(effect) pairs; many such cause-effect pairs have an intelligibility of their own when they are seen in the context of the larger goals or strategies of the agent, and postulating mental divisions is often an unnecessary complication. Davidson would need to add another condition that would indicate what kinds of causes that are not at the same time reasons for their effects would cause a division to occur in the mind of the agent. Perhaps this criterion might be that the agent ignores her principle that she ought to act in accordance with her judgement of what would be best to do, all things considered. If a cause-effect pair that is not at the same time a reason-result pair does not go against the agent's principle that it would be best to perform another action, then there will be no need for a division within the agent's mind. Similarly if the cause-

---

[14] Mele, A Irrationality esp. Chapter 6

effect pair is adopted by the agent for some other purpose, and is hence motivated. This would be the case in, say, the example of the smoker. That is, as long as we do not have subversion of the normal belief-forming process we do not require the divisions. That is - a belief caused by a mental event that is not at the same time a sufficient reason for the belief need not be irrational, provided that it does not go against some other belief of the agent. Divisions arise in the face of two factors: the existence of a mental cause for an item (a mental item) that is at the same time not a sufficient or evidential reason for the item, and the overriding of the principle that one ought to act on what is best, under the circumstances.

Suppose that Anna wants to stop smoking. She decides that the best way to do it would be to engage in some kind of aversion-therapy. She decides that whenever she desires to smoke, she will force herself to watch a very violent movie. Anna hopes that the disgust she feels towards the film will come to be associated with smoking, and she will eventually lose her desire to smoke. However, suppose that Anna objects to the making and showing of very violent movies - she finds them morally reprehensible. She feels that watching such films is more damaging to herself - albeit in a moral and non-physical sense - than is smoking. So now if Anna were to continue to watch such movies as a means of stopping smoking, her behaviour would be irrational. She would be acting against another stronger belief. On the other hand, in Mele's example of the smoker, the action undertaken was performed in the agent's best interests. We could perhaps say that, although the daydreaming about floating in the Caribbean was not in itself a reason for giving up smoking, it was employed by the agent as part of the strategy for giving up smoking. That is, it was subsumed under the larger intentional plan of the agent, and, although it may have acted as a non-reason cause of the particular action of the agent, it did not undermine the rational functioning of the agent. There

was no other reason that mitigated against the daydreaming plan. So I think that Mele's objection to Davidson can be accommodated within Davidson's account, and does not, I think, pose a great threat to it.

The two main problems with Davidson's account are: the partitioning of the mind may not in itself solve the problem of self-deception, for the reason that we often have to look <u>across</u> divisions in order to appreciate the force of the paradox. This is something the agent cannot do, and so the notion of agency may be at stake here. Heil's objection above relates to this point. Another problem is that we are not given any explanation or description of the genesis or mechanism of operation of the partitions - we are not told <u>how</u> we arrive at these divisions. In the following chapters, I shall examine the Freudian account of the nature of mental divisions where criteria for their genesis and functioning are postulated, and hence the problems associated with the Davidsonian case may be avoided.

THREE

FREUD: THE TOPOGRAPHIC THEORY

Before I begin an analysis of the Freudian thesis - or rather theses - I would like to mention one or two criticisms of this type of approach in order that the reader may become aware of the problems before approaching the proposed thesis. I think that this will be useful insofar as it will allow the reader to assess whether the proposed thesis does in fact encounter the problems mentioned, or whether it either avoids them or does not encounter them at all. Furthermore, some of the problems to be discussed are fairly general in their range - that is, they would apply to any theory within this broad approach (the approach being that of the "divided mind") - and so for this reason at least need to be mentioned. There will be no attempt at solutions at this stage; rather the issues will be addressed either in the body of the discussion should it seem relevant, or at the end of the discussions, when the theses under scrutiny will have been discussed. One more word of caution: in the following discussions I shall be referring to the "minds" of individuals: what I mean by this term may require much discussion to clarify, but allow me to say that I do not mean by this term the brain of the individual. That is, the term "mind" does not necessarily refer to any physical entity. The relationship between mind and brain is a large and complex issue in the philosophy of psychology, and cannot be reviewed here. Suffice it to say that the term "mind" as used here is essentially non-physical - what its relation to the physical brain may be I shall not explore.

Another issue that needs to be raised at the start of my discussion on Freud is my approach taken in the handling of the Freudian material. I have tried to be as historically correct as possible in terms of representing Freud's ideas accurately. However, I am also interested in using the Freudian material as a source for creative ideas in looking for a solution to the problem of self-deception; this may appear to lead me astray from traditional Freud, but I trust that these explorations of Freud's ideas do not distort the theory in any way. I also believe that Freud's ideas lend themselves to further investigation within the philosophy of psychology and related areas, and that such creative use of his ideas can be both interesting in themselves, and useful in terms of analysing possible developments within the theory itself.

There arises the problem of whether the notion of a divided mind is at all intelligible. To quote Davidson here - "[the theory of mental partitioning] seems to require that thoughts and desires and even actions be attributed to something less than, and therefore distinct from, the whole person. But can we make sense of acts and attitudes that are not those of an agent?"[15] A "divided mind" thesis raises problems within the area of agency because it seems to splinter the essential unity of self necessary for the notion of agency at all. If we attribute the behaviour or belief-state of an individual to the interaction of various partitions within the self, then the locus of identity of this very "self" may become problematic. Addressing the agency problem will be one of the requirements of a divided mind thesis. Another issue to be addressed is, of course, the procedures for dividing the mind. As will be seen, one can divide the mind in a variety of ways, and the arguments for the divisions

---

[15] Davidson, D "Paradoxes of Irrationality" in Wollheim, R (Ed) p 291

also require examination. In this section I shall examine one of the theses put forward by Freud.

Freudian theory of the mind can be separated into two broad theories. The first, the topographic theory, is the earlier of the two, and was first presented by Freud in Chapter VII of The Interpretation of Dreams. The second proposal, the structural theory, was developed in 1923 in The Ego and the Id, primarily to account for inadequacies in the topographic theory to account for the formation of neurotic symptoms, but also to give a new and different account of the role of repression. It has been argued[16] that the two accounts are incompatible, and that the structural theory supercedes the topographic theory; that is, it has been argued that, instead of working with both theories, and switching from one to the other depending on the phenomena to be explained, we should use only the ideas and terminology developed in the structural theory. In this paper, however, I shall examine both theories separately, and not concentrate only on the structural theory. The reason for this is that the theories postulate the divisions of the mind as occurring in different ways, and it will be useful to explore both of these pictures. Furthermore, the divisions postulated by the topographic theory have been used by philosophers in an attempt to solve the problem of self-deception, and thus the ideas postulated by the theory have a respected place in the literature. Also, although the topographic theory was superceded by the structural theory, for reasons mentioned above and to be explored further, it may be the case that the topographic theory may adequately explain the notion of self-deception, even if it cannot account for other kinds of mental functioning. However, even if the accounts are incompatible, and

---

[16] Most notably by Arlow and Brenner. See their Psychoanalytic Concepts and the Structural Theory

the topographic theory turns out, upon examination, not to be fully explanatory, the fact that these two theories are still both used today within psychoanalytic and psychological discussion seems to warrant a separate treatment of each theory. It is to such an examination that I shall now turn.

The topographic theory divides the mind into what is conscious and unconscious; this distinction is exhaustive - a mental content must be either conscious or unconscious. The question of whether these divisions are mutually exclusive is not as clear. Certainly a mental content may change qualities: prior unconscious beliefs may be brought into conscious, while what is now conscious may, at a later stage, be repressed. There is uncertainty as to whether a belief or idea can be simultaneously conscious and unconscious. There is discussion in Freud concerning a belief that is repressed. In his paper "Negation" (1925), Freud discusses the situation where a patient makes a suggestion or association - thus indicating consciousness of the subject matter or content of an idea - only to deny it - that is, deny that the association is a true indication of the subject's belief state. Freud gives the following example: "The patient says [to the analyst]: 'You ask who this person in the dream can be. It's not my mother.' We [the analyst] emend this to: 'So it is his mother.' In our interpretation, we take the liberty of disregarding the negation and of picking out the subject-matter alone of the association. It is as though the patient had said: 'It's true that my mother came into my mind as I thought of this person, but I don't feel inclined to let the association count.'"[17] Thus it seems that a mental content is able to be conscious and repressed at the same time. Freud attempts to explain this

---

[17] Freud, S "Negation" (1925) Freud Pelican Library Volume 11 On Metapsychology p437

by dividing a belief or mental event into an ideational part and an emotive part. The patient accepts the ideational part of the idea - hence there may be an intellectual acceptance of the idea - while the emotional or affective aspect of the idea still undergoes repression. By negating something, the patient is in effect saying - "'This is something which I should prefer to repress.'" There is awareness of the matter as a possibility - the thought is in the mind - but the proposition, should it be entertained - is denied; a level of awareness is directed towards some aspect of a mental state that remains, on the whole, unconscious. The question of whether the same belief can - in its entirety - be both conscious and unconscious is uncertain. However, I think that it is clear that the partitions of the mental that are characterized by consciousness and unconsciousness do not overlap. That is - an area of the mind that is characterized by consciousness, say, will not overlap with an unconscious area of the mind. This is, I think, implicit in the notion of the Freudian agencies. So although an individual mental content may be characterized by both consciousness and unconsciousness (perhaps in the way outlined above), the larger partitions of the mind will be characterized by one quality only.

Freud further divides the mental apparatus into three systems; the Conscious (CS), the Preconscious (PCS), and the Unconscious (UCS). The PCS stands between Consciousness and Unconsciousness, and before a thought can become conscious, it must pass through the system PCS. What is Preconscious is not conscious (hence what is preconscious is also unconscious), but unlike the Unconscious (the system UCS), it can easily become conscious, whereas what is Unconscious may become conscious only with extreme difficulty, if at all. The terms "conscious" and "unconscious" are hence used in two different ways, and it is important to distinguish between the different usages in order to avoid confusion. Firstly, the terms are used by Freud simply to characterize the nature of the mental

- in this sense a thought or belief is either conscious or unconscious - that is it is apprehended by the perceptual system (if it is conscious) or it is held unawares by the agent (unconscious). On the other hand, the terms are used to characterize different systems, and when this is done, capitals shall be used. Clearly, the system CS is characterized by consciousness, and the system UCS by unconsciousness. As stated above, the system PCS is also unconscious, but the thoughts and beliefs belonging to it are readily accessible to consciousness. Hence the systems UCS, PCS, and CS do divide the mind along the lines of what is accessible to consciousness and what is not.

The systems differ also with respect to how they operate within these parameters; other features also distinguish them. For instance, another characteristic of the PCS is that it is connected to language, whereas the unconscious is not linguistic at all. The connection with language facilitates the accessibility of Preconscious thoughts to consciousness, although the reason for this accessibility is the lack of repression exerted on Preconscious ideas. The system PCS develops later in the individual than the UCS, and its behaviour is more "adult-like" than the UCS. The instincts of the Unconscious strive for immediate gratification, while the work of the Preconscious is to delay gratification of a wish in the face of reality. In order for an Unconscious wish to become accessible to consciousness it must pass through, and be regulated by, the Preconscious. So there is more to the various systems than simply the fact that they are conscious or unconscious - they are characterized by other qualities as well. Whether these other distinguishing features are problematic for the topographic theory will be explored at a later stage.

Central to the Freudian thesis, and also to the problem of self- deception (as explored within this model) is the concept of repression; it will be necessary to examine this notion in some detail, as one of the ways an

it must hide the information from itself. This means either that the censor is divided against itself, or else the censor required the existence of another censor to hide its operations from itself. In the latter instance we have the problem of an infinite regress of censors; in the former, we would need to explain the nature of this division as well as postulate another censor that would operate between the divisions. We can avoid these difficulties by postulating a censor that need not hide its own operations from itself. Instead of complicating matters, all we need insist upon is that the operations of the censor remain covert - that is, that there is no access to consciousness.

The latter point raises a problem regarding the activity of the censor: how does the censor hide its own activities from consciousness? If the censor belongs to the system PCS, then its operations should be easily accessible to consciousness. But it is clear, both from experience and from the very notion of a censor, that should this occur, the censor would work very ineffectively, if at all. This is one of the problems that Freud discovered in the topographic theory, and which precipitated the search for and discovery of the structural theory. Now, we could postulate that the censor operate between the UCS and the PCS; this way it would have access to the contents of the Unconscious and of the Preconscious without belonging to either system. It may or may not have access to the PCS; since the PCS is itself a filtering agent, it is most likely that the PCS would be aware of the censoring activity. However, since there is additional censorship between the PCS and Cs, this knowledge would not necessarily be accessible to consciousness. Does this covert operation of the censor imply that the censor is repressed? If we take this line, we must then go on to postulate a repressing agency, and such a move would, I think, complicate matters unnecessarily, and in such a way as to make the topographic theory extremely cumbersome. There is no good reason that would motivate the censor to strive for accessibility

must be able to supply us with reasons for her behaviour - and this she does - she no longer likes the food at the restaurant. But the "real" cause of her aversion to the restaurant is her fear of finding her boyfriend there with the waitress. So although the belief is repressed, it still influences the woman's behaviour. This example seems to illustrate that self-deception is an unstable predicament in which to be: the woman's belief that her boyfriend is not having an affair (despite the evidence which we assume she has "chosen" to ignore) is not assured, for if it were the woman should be able to frequent the restaurant free of suspicions.

So we see that the deceptive plan is often ongoing, and part of its function is to keep the deceived party deceived in order that the operation can continue to influence the subject's behaviour. Also, since the repressed unconscious belief still has the power to influence behaviour, it seems that dividing the mind into conscious/unconscious partitions is not sufficient an explanation of self-deception. The unconscious belief is not altogether "shut off" from the main system, but it continues to influence it - even in ways of which the conscious system is itself unaware. The unconscious nature of the cautionary belief is not sufficient an explanation of its "non-intervention" in the main system - in fact, the very unconscious nature of the cautionary belief is the very condition that allows for its having an influence in the main system at all.

The structural theory divides the mind into three agencies: Id, Ego and Superego. The Id corresponds closely with the system UCS of the topographic theory, and consists of the "mental representatives of the instinctual drives."[22] Thus the Id consists of wishes and desires that press for immediate gratification, and it is concerned with the gratification of its desires rather than with the demands of reality. The Ego is that part of the system that is regulated by the demands of reality. It is connected to the perceptual and motor parts of the self, thus allowing it to act in the world in such a way so as to best allow the fulfilment of Id desires in conformity with the strictures of reality and with least harm to the individual. It is also the job of the Ego to censor and control those contents of the Id that it considers either impossible or too threatening to fulfil in reality. The relations between Id and Ego may often be characterized by a state of conflict. This occurs precisely when the Ego represses aspects of Id instinctual desires that are too threatening. Often this act of repression by the Ego is triggered by a feeling of anxiety on the part of the Ego: the Ego cannot cope with or accept the demands of the Id, and, as a result of feeling anxious because of these demands, represses the Id-wishes. In this sense, the Ego acts like the PCS in carrying out censoring activities, and it is this aspect of the Ego that shall concern us in the question of self-deception. Other aspects of Ego functioning include sense perception, control of motor action, language and reality testing.

The Superego is a new agency, and does not have any correlate in the topographic theory. The Superego is concerned with the moral aspirations and laws of the subject, laws derived from parents and other authority

---

[22] Arlow, J. A and Brenner, C Psychoanalytic Concepts and the Structural Theory p 32

idea becomes unconscious is via repression. Since one of the outcomes of self-deception may be that one belief is held consciously by the agent while the other (unwelcome) belief is held unconsciously, the way this state is reached will be of importance. If it is found that there are problems or inconsistencies within the theory of repression, this may have an effect on our view of self-deception, and the adequacy of this approach to deal with the problem will have to be assessed.

The concept of repression is used to characterize the difference between the UCS and the PCS. Both systems are unconscious, since their contents are not conscious, but what characterizes the Unconscious and makes it inaccessible (or accessible only with difficulty) to consciousness is the fact that it is repressed. We can, if we like, say that the Unconscious is the unconscious-repressed. Contrary to this, the system PCS is latent, in the sense that it is unconscious, but capable of becoming conscious without further ado. The system PCS may be called descriptively unconscious, while the UCS is dynamically unconscious; the latter is subject to constant repression. What, then, is repression? The central role of repression is to keep out of consciousness thoughts or ideas that the Preconscious considers too threatening to the subject to be allowed fulfilment in the real world. The Preconscious acts as a kind of screen between the wishes of the Unconscious and the demands of reality, and if it is likely that the becoming conscious of a particular Unconscious wish would be destructive for the subject, the PCS withdraws cathexis (energy) from the wish and turns it back into the Unconscious. It also happens that the Unconscious wish is not denied access to Consciousness completely, but the Preconscious may decide to delay the fulfilment of a wish until the appropriate time for its discharge arises. In the topographic theory, the result of the failure of repression is the feeling of anxiety - the experience of unpleasure on the part of the subject is an indication that repression has not succeeded completely in keeping these thoughts

from reaching consciousness. For the topographic theory, repression (or its failure) precedes anxiety. This condition is reversed in the structural theory, a factor we will examine later. Suffice it to say for the moment that it was a change in his views on repression that led Freud to alter his theory of the mental.

The important question to ask at this stage is - how does repression work - and central to this - what does the repressing? In the topographic theory, since the central role of repression is to keep ideas or thoughts from consciousness, it cannot be the case that it is the conscious system - CS - that does the repressing; if this were the case then consciousness would be aware of what it was keeping from itself - a paradoxical notion, and a problem raised by Sartre.[18] The point to repression is that it keeps ideas out of consciousness. In the topographic theory, Freud locates censoring agencies between the systems UCS and PCS, and also between the systems PCS and CS. We are, says Freud, "to assume that to every transition from one system to that immediately above it (...) there corresponds a new censorship".[19] However, Freud goes on to say that the more important distinction lies between the Unconscious and the Preconscious, and not between the Preconscious and Consciousness. Here, the repressing agency would be located within the PCS, or the PCS itself would act as a

[18] See for example his article "Mauvaise foi and the Unconscious" in Woolheim and Hopkins (Eds). See esp. p 205 - "[when I am self-deceived] I must know in my capacity as deceiver the truth which is hidden from me in my capacity as the one deceived. Better yet I must know the truth very exactly in order to conceal it more carefully ..." This captures very well the central problem of self-deception, and the paradox that is involved. Sartre's attack would also, though, be directed against the Freudian theory in general, and not only in the specific area that I discuss. I would, of course, not agree with Sartre in his general attack, for reasons to be discussed.

[19] Freud, S "The Unconscious", (1915) Freud Pelican Library op. cit. p 196

censor; this is possible because the PCS, unlike the system CS, is unconscious. Motivation for repression of Preconscious elements from entering CS would, however, come from Unconscious derivatives that have managed to gain access to the Preconscious, but whose strivings towards consciousness are regarded with suspicion, and hence subject to further repression. The point here is that the major repressive activity lies between the UCS and PCS, and in most cases the existence of a repressing agency between the PCS and CS is the result of some kind of failure of repression by the PCS regarding UCS material. It is this material that would require additional censorship. It is still the case, though, that the primary repression occurs between the UCS and PCS; that is, the existence of a censoring agency between the PCS and CS should not make us redefine the PCS as dynamically unconscious. Contents in the PCS would still be described as latent, in the sense used above.

The censor within the system PCS has access to the contents of the UCS, and prevents these ideas from forcing their way into consciousness by withdrawing cathexis, usually word-cathexis, and forcing the idea back into the Unconscious. The censor would also require access to consciousness - it could not allow unconscious material into consciousness if there were a contradiction or conflict between the contents of either system and it would need to be aware of what Unconscious ideas would be threatening to the subject were they to become conscious. It need not be the case that this censor share any or either of the beliefs of either system. Furthermore, since the censor need not have a belief system of its own - other than the cautionary one that allows it to ascertain what it should censor - the censor need not be involved in any attempts to hide its activities from itself. Hence we are able to deal with the objection raised by Sartre. This objection is expressed as the "knowledge which is ignorant of itself". Sartre postulates the censor as self-censoring. It must be aware of the information that it represses, and furthermore

Pears's use of Freudian terminology seems at times to differ from classic Freudian usage. For instance, his use of the term "preconscious" seems to contradict the Freudian usage of the term. For instance, Pears states that the "preconscious belief would always be shut up in its sub-system and it would never be able to use the sub-system as a base for operations in the main system ... "[21] But as defined above, the contents of the preconscious are readily accessible to consciousness. Terminology aside, Pears raises an interesting problem. The relegation of the cautionary belief to an unconscious sub-system is to prevent the intervention of this belief in the main system; that is, the existence of the cautionary belief in the sub-system ensures that it does not have any control over or effect in the consciousness of an agent. Pears's point is that in fact such unconscious beliefs do still continue to influence our actions - that is, they seem to have an effect in consciousness, and very often the effect is significant. In some instances it may seem as if the repressive process is there to aid rather than hinder the irrational and repressed belief, as now it is able to operate - albeit covertly - whereas if it was recognized by consciousness and then rejected it would not be able to operate in the same way at all. An example of how an unconscious wish or belief can influence behaviour is the following: suppose a woman suspects her boyfriend of having an affair with the waitress of a certain restaurant that the woman often used to frequent. She tries to suppress her suspicions, but she no longer frequents the restaurant, and even goes out of her way to avoid it. When questioned, the woman might mutter something about the food, and say that she no longer enjoys eating there. We are assuming in this example that the woman is sincere in her replies. In this example we have the woman's behaviour being influenced by an unconscious belief. Of course, if we are to regard her as an agent, she

---

[21] Pears, D Motivated Irrationality p 79

F O U R

FREUD: THE STRUCTURAL THEORY

This section will examine the structural theory developed by Freud. The structural theory arose out of inadequacies with the topographic theory; some of these have been touched on in the previous chapter, but they shall be developed and explored here. Firstly, in the topographic theory, repression preceded anxiety; the feeling of anxiety was the result of the failure of repression. Freud came to reject this description of the workings of psychical forces, and proposed instead that anxiety precedes repression, and is in fact a precipitating factor of repression. This analysis was derived from clinical observations, and it is illustrative of a fundamental difference between the topographic and structural theories, a difference which is important for our analysis of self-deception. The topographical theory seemed to be at odds with clinical observation in the following way: Freud found that patients were very often unconscious of the fact that they were repressing undesirable thoughts or ideas. Patients resisted talking about certain experiences or memories, but they were not aware of the fact that they were offering such resistance to the analysis. In the topographic theory, the repressing agency is the PCS, a system accessible to consciousness. If this is so, then the act of repression should itself be conscious; the fact that it very often is not is problematic for the theory. There are other criticisms that Freud gave of the topographic theory, but since they are not directly related to the issue at hand, they shall not be explored here.

figures. The Superego would also be concerned with the regulation of the activities of the agent, especially in relation to its moral dictates, but it would regulate the activities of the Ego itself. Now all these agencies constitute the mental and psychological world of the subject, but it is the Ego which is the principle of unity of the subject. One is able to define the agencies in terms of both functional and genetic explanations. Thus the Ego develops as a result of certain conflicts between the demands of the Id and the outside world, but it is also characterized by the functions it performs in the life of the individual - say that of censoring undesirable Id wishes. So we see that the principle of distinguishing between agencies is different from that of the topographic theory, where the fundamental distinction between systems was along the lines of what was or was not accessible to consciousness. We need to explore the relationship between the various structural agencies and consciousness, because we need to assess the relation between repression and access to consciousness. In the topographic theory, repression of an idea meant that the idea was no longer accessible to consciousness. It will be necessary to examine what role repression plays in the structural theory, and hence the relationship between the various agencies of the structural theory and accessibility to consciousness.

In the topographic theory, the only means of defence is repression. In the structural theory there are other types of defence mechanisms available - for instance, projection or denial, where the ideas defended against are not so much barred from consciousness as altered in some other way. For instance in projection the undesirable proposition ("I hate him") is changed into its opposite ("He hates me"). Projection is of course a way of denying that the act of hating is done by oneself, but the mechanism of defence here is not purely to repress the "act of hating", as it were, and bar it from consciousness. Rather it is recognized, but attributed to another agent instead of to oneself. What will be of

particular interest to us is the relation between the three agencies and consciousness, and the relation between the mechanisms of defence or repression and consciousness. It is, I think, fairly safe to say that the Id can be closely associated with the Unconscious, and operates within the structural theory as an unconscious system. The Superego, as a repressing force, or rather as one that lays down the conditions under which the Ego may feel obliged to act in terms of repressing or otherwise censoring Id-contents, may be either conscious or unconscious. Certainly many of our moral strictures are conscious and easily accessible: one believes without conflict (usually) that one should not kill, say, and these beliefs govern our behaviour, usually in an appropriate fashion. But sometimes the laws or operations of the Superego are inaccessible to consciousness. The Superego may be a harsh and punitive agency, and may attempt to punish the Ego for crimes which it "believes" the Ego has committed. Very often these punitive attitudes are unknown to the person involved, and a subject may perform a punitive action towards himself without knowing either that he is doing so, or why he may be doing so. So the Superego seems to intersect both consciousness and unconsciousness at different times.

The nature of the Ego is most relevant and also most problematic to my project, and the emphasis in this chapter will be on the problematic of the Ego in relation to self-deception. It is, I think, the case that any solution to the problem of self-deception must be found within the realms of Ego-psychology, if it is to be found within the realms of Freudian psychology at all. The mechanisms of defence are located within the Ego, and so it is here that the forces of self-deception should enter. As the Ego is that part of the psychical personality that is responsible for the integration of the various psychical forces so that the individual is able to function as an agent or subject in the world, it is clear that it is to the Ego that we must look if we are to examine the mechanism of defence

that are set precisely for this purpose. We shall first need to examine the relationship of the Ego with the states of consciousness, to try to assess how the state of self-deception can be possible, given the nature of the Ego.

It seems obvious that, of all the agencies, it should be the Ego that has the most access to consciousness. Certainly many of the activities associated with the Ego require awareness - decisions to act, for instance, or the very act of thinking or speaking, are conscious acts. But the Ego, like the Superego, has portions of it that are unconscious as well. Freud wrote in his New Introductory Lectures that "large portions of the ego and superego can remain unconscious and are normally unconscious ... it is a fact that ego and conscious, repressed and unconscious, do not coincide."[22] This suggests several things to us. Firstly, we see that the terms "conscious" and "unconscious" are no longer used to delineate topoi or regions in the mind, as was done in the topographic theory. So the field of reference of these terms has now been narrowed. The second is the point stated above - namely that parts of the Ego itself are also unconscious. Thirdly, in the above quotation, we also see a distinction being made between the repressed and the unconscious. This suggests that we have discussed above - namely that we have in the structural theory other forms of defence that do not require keeping ideas out of consciousness. However, it may also suggest that what is unconscious is not so necessarily because it is repressed. It shall be necessary to examine the relationship between the unconscious aspects of the Ego itself and the repressed, because it is the Ego that is involved in the act of con-

[22] Freud, S "Dissection of the Psychic Personality" in New Introductory Lectures p 102

to consciousness; unlike the UCS, which actively strives to break through into consciousness in order to attain fulfilment of its desires, the censor would have no such strivings. It would be counter-productive were the censor to desire such access to consciousness, and hence there may be no need for repression of the censor at all. Also, the censor could be depicted as a kind of screen or filter; on this description, we would not have the creation of a new agency, and thus there would be no distortion of the topographic model. So if there is a censor that operates between the PCS and UCS, we shall postulate that it must operate in a covert manner without necessarily being repressed.

The operations of the censor that lies between the PCS and CS would work in a similar fashion. This censor would also act as a kind of filter, and it would operate between the agencies without necessarily belonging to either one. We could, perhaps, postulate that this censor operates within the confines of the PCS; it is a kind of check on the contents of the PCS, and such contents must pass through the filtering process before becoming conscious. Again, the censor would also work covertly, but possibly in such a way so as not to require repression. The nature of this censor would probably be less repressive than the one that lies between the CS and PCS.

Another problem that arises has been raised by David Pears.[21] According to him, repression is made possible by a "cautionary belief". The notion of a cautionary belief assumes the divisions of the mind into systems. It works as follows. On this account, a person consists of two separate systems -- a main and a sub-system. The main system is the dominant one, and it would include the irrational belief as well as the information and

[21]. See his Motivated Irrationality, esp. Chap V

the evidence that would make the belief irrational. The sub-system would house the "cautionary belief" that would mitigate against the formation of the irrational belief. So, for instance, in the case of a deceived husband who continues, despite contrary evidence, to believe in his wife's fidelity, the main system would house both the belief that the wife is faithful, as well as the evidence to the contrary. So the belief in the fidelity of the wife would clearly be irrational. The cautionary belief would stand in opposition to the favoured but irrational belief; in this case the cautionary belief would be one regarding the wife's infidelity. If the cautionary belief were part of the main system, it would prevent the formation of the irrational belief. So its existence in the sub-system explains and ensures - rather than prevents - the continuing existence of the irrational belief. In this way the division into systems acts as a "permissive cause" of the irrational state, as it permits the irrational belief to exist, and hence makes the irrational state possible. On the other hand, the "productive cause" of the irrational state would be the (irrational) wish or belief itself that was the cause of the whole operation - in this case the desire of the man to continue to love and trust his wife . Pears's theory - as too Freud's topographical theory - operates within the conscious/unconscious division of the mind; the sub-system to which the irrational belief is relegated is unconscious, and it is this unconsciousness that keeps it out of the main system and allows it to operate covertly. It is the unconscious nature of the cautionary belief that allows it to act as a permissive cause of the irrationality. The converse must also be true; "[i]f a person's cautionary belief, that the belief that he was forming was irrational, were conscious, it would prevent him from forming the irrational belief." So the conscious / unconscious distinction operates as a permissive cause of the irrationality.

soring ideas, and so the notion of a part of the Ego that is unconscious to itself must be examined.

There is one more point to make about the nature of the relationship between the Ego and the Id that will indicate where the solution to the problem of self-deception does not lie. Freud characterized the Id as that region which stands outside the Ego. It is described as a place of chaos and contradiction, where conflicting wishes and desires can exist side by side. Now this nature of the Id suggests that we should not look for the solution to the problem of self-deception here; we cannot. I think, claim that self-deception could be possible if the belief that p resides in the Ego, and the belief that -p resides in the Id, and hence the subject holds both beliefs simultaneously. The one problem - and there are others - is that it seems impossible to talk of the Id as having a belief system at all. The Id seems to be characterized by irrationality, and so it is absurd to attribute beliefs to it. So it seems that the system Id should be left out of the discussion altogether, and analyses of its nature not undertaken here. The only contribution that the Id would make to the problem - and it is not an insignificant one - would be that the impulses or wishes of the Id are acknowledged or taken up by the Ego itself, and in this way "converted" into beliefs. This would occur were a desire of the Id to be recognized and acted upon by the Ego. So the Ego would acknowledge the desire as its own, and would thus be able to form beliefs concerning it. So the focus of concentration should be on the Ego itself - and perhaps on the Superego - in an analysis of the problem.

It may be a good time now to raise a worry in regard to the structural theory. Can the problem of self-deception be solved by positing that the two contradictory beliefs are held be separate agencies, whether it be Id and Ego, Ego and Superego, or Id and Superego? Here we have the problem

of the divided mind presenting itself. Now this type of division may assist us in understanding other aspects of human psychology, but I am doubtful as to whether it can help us solve the problem of self-deception as it presents itself to us. The problem is that if the two beliefs are held by separate agencies - especially agencies that are unaware of one another's existence - there does not seem to be a single agent holding these two beliefs simultaneously, and somehow deceiving himself with respect to them. Multiplying agencies does not seem to work for the simple reason that the notion of self-deception requires that the contradictory beliefs be noted by the same subject - Sartre's point once again makes its stand. So positing the split in the psyche between Ego and Superego may not be very useful for our argument. The place to look, then, if we are to attempt to solve the problem from within the Freudian model, is at the Ego itself, and it is to a discussion of Ego psychology that we shall now turn.

It is the Ego that is the principle of synthesis of the psychic personality. The Ego mediates between the Id, Superego and the external world, and has connections with all three loci. The Ego strives for synthesis and organization, and attempts to perform the most rational and realistic course of action under the circumstances. But Freud also talks of the unconscious part of the Ego, and in a later - and unfinished - metapsychological paper ("Splitting of the Ego in the Process of Defence"). Freud talks of the Ego's ability to "split" into two (or more?) parts as a method of defending itself against undesirable beliefs. Both of these aspects need to be examined in turn, as I am not at all sure if they coincide. It may be that the way that the Ego splits is not along the lines of a conscious/unconscious distinction at all.

In a paper in Wollheim and Hopkins (Eds),[24] Herbert Fingarette raises the possibility of a split in the Ego itself to deal with the problem of self-deception. Fingarette states that there is an Ego structure that becomes split off from the main part of the Ego system because the latter is unable to integrate this "Ego-nucleus" into itself, and defends itself against it. The Ego-nucleus is defined as a complex organization of "motive, purpose, feeling, perception, and drive towards action." The way that it is characterized suggests that this nucleus does have access to, or is influenced by, reality, but is undeveloped and infantile in its operations, being unable to "learn" from its past behaviour. What Fingarette needs to do is to characterize this Ego-nucleus in such a way so that it is sufficiently like the Ego to be a part of it, but sufficiently unlike it to have become estranged from the central system. Fingarette claims that this defensive action on the part of the Ego - this splitting - on the part of the Ego is motivated - it is something the Ego "does". Fingarette does not sketch in any detail how this Ego-nucleus is either formed or repressed. He postulates that the repressed system is Ego-like rather than Id-like for the following reason: he argues that it is not the case that many symptoms that would be classified as belonging to the Id seem to have many characteristics of the Ego, say. Fingarette argues that the postulation of separate non-intersecting agencies does not sufficiently explain psychological behaviour. His suggestion is to locate the fundamental split between Ego and some other system within within the Ego itself, so that this "other system" becomes the Ego-nucleus. I would agree with Fingarette's conclusion that a split within the Ego itself seems to be a necessary move to make. The problem is not so much re-describing the agencies of the structural theory so as to accommodate this new entity, but understanding how this new de-

---

[24] Fingarette, H "Self-Deception" p223

scription of the mind actually works, especially in the area of repression, which is central to the problem at hand.

One word of caution is needed at this point. We seem to be dividing up the mind into more and more entities. Not only do we have now the three agencies of the structural theory, but we have a division within the agency of synthesis itself - namely the Ego. This move alters to some extent the nature of the Ego, which had hitherto been described as the centre of organization - the unifying system that made agency possible. We shall need to investigate whether this move does in fact destroy the notion of agency altogether. But first the mechanism of the splitting of the Ego itself needs to be examined.

The first question to pose here is: what is the psychical quality of that part of the Ego that is repressed? Is it conscious or unconscious? We noted that in the structural theory, making something unconscious was not the only mechanism of defence. In this case, however, we do not have the undesirable idea relegated to another system, but we have instead the creation of a split within one particular system. It is possible to conceptualise defending against a belief by a process other than that of repression. The idea of projecting an unsavoury desire held by oneself onto someone else seems to make sense - we do it all the time. But it is difficult to see how one could defend against the existence of an entire system other than by repressing it. Another alternative would be to posit the systems as alternatively having access to consciousness, as in the case of so-called multiple personalities. But this alternative does not accord with most people's experiences, and is hence unsatisfactory, to say the least. Another attempt might be to argue that both systems are conscious? Could we coherently postulate the existence of more than one consciousness within the subject? This seems a strange suggestion at best. The idea of two consciousnesses existing in ignorance of the other

contradicts the very notion of consciousness. As Freud himself wrote in his essay "The Unconscious": " ... a consciousness of which its own possessor knows nothing is something very different from a consciousness belonging to another person ... Those who have resisted the assumption of an unconscious psychical are not likely to be ready to exchange it for an unconscious consciousness." The work done in the area of commissurotomy might suggest that this notion cannot be dismissed out of hand; nevertheless, the postulating of two consciousnesses within a single subject raises all sorts of problems concerning personal identity and questions of subjecthood too numerous to enter into here. I trust that the bizarre nature of this area of enquiry will make it apparent why I would not like to raise the problem here; I also trust that this problem is not central to that of self-deception. Within Freud's own theory, it is safe to say that this Ego-nucleus does not form a separate consciousness. Now since it belongs to the Ego, which is the locus of subjecthood, as it were, if it is to retain its Ego-like character and yet be separate from the Ego, it must subsist in the unconscious part of the Ego.

It seems that there is no alternative but to suppose that this Ego-nucleus is defended against by being made unconscious. So now we have a repressed, unconscious part of the Ego existing alongside the conscious Ego. It may be that the two qualities do not overlap entirely, that is, it may be the case that there are parts of the unconscious Ego that are not repressed, but merely latent. The problem is not so much postulating that aspects of the Ego are unconscious - this had been postulated by Freud fairly early in the structural theory - but it is problematic to propose an unconscious repressed within the Ego itself. And yet this move is in some sense necessary. It is necessary because without this postulation we could not have the formation of symptoms. If the only threatening wishes belonged to the Id, then the Ego could simply repress them completely and not allow them to influence the Ego's behaviour in any way.

But it was discovered by Freud that the formation of symptoms arises as, and in the form of, a compromise between two or more agencies. The Ego tries to satisfy the conflicting demands of Id and Superego and the external world, and it may do this by forming some sort of a compromise between the various agencies and their wishes. What this means is that the ideas or wishes of the other agencies must be acknowledged and made sense of by the Ego. It is as if the Ego has to translate the language of Id and Superego into a language that it can understand. This is especially true of Id impulses, the organization and origin of which may be irrational in nature. It is difficult to see how the Ego could assimilate such Id wishes into itself without changing them in some way and making them its own. But, as noted before, many if not most of these wishes are not known to the subject - that is, they are unconscious. So the existence of an unconscious within the Ego seems almost necessary for the coherence of the structural theory.

The problem arises with respect to the existence of a wish that the Ego does not wish to acknowledge, either because it is irrational or for some other reason, and yet this wish affects the behaviour of the individual. In other words, how is the repression that is necessary for self-deception possible within the Ego? What is the agency that does the repressing? How is it possible for the censor to exist at all within the Ego? In the case of self-deception we have the problem raised by Sartre that throughout the deceptive procedure both beliefs - the desired and the undesired - must be held by the subject. In the formation of the irrational belief, it is the Ego that has access to the evidence at hand, evidence which is recognized and at the same time ignored in order to allow the irrational wish to be formed. In the case of self-deception, the evidence for the rational but undesired belief that p is recognized by the agent. This is a restatement of the paradox, but now it is shifted to the level of the Ego, and from here it seems that there is nowhere else

to go. We cannot keep on multiplying entities endlessly, and if a solution is to be found to the problem it should be at this level of the analysis. I shall attempt to explore the nature of repression within the Ego.

The first thing to note is that repression of an idea from the main Ego system is not sufficient to prevent it from having an effect on the behaviour of the subject. The woman who suspected that her lover might be having an affair managed to keep this belief out of consciousness, but it still affected her behaviour. From this we should conclude that repression does not deprive an idea of its force or effect on the individual. All that repression requires is that the irrational or unwelcome belief does not enter into the main belief system. Since the unconscious belief is able to affect behaviour, it is possible that the subject may be able to infer it from observing her own actions. This could result in at least two things: either the subject continually attributes the actions to some <u>other</u> more acceptable reason, or there is an ongoing monitoring of the belief that ensures that it is kept from the consciousness of the agent. Both of these results could work simultaneously. Of course, it may happen that the person does not infer the real reason for his behaviour - it may be the case that the person just doesn't know why he is acting in such a way - that is, the person may avoid using the evidence of his own behaviour in a way akin to the avoidance of other evidence. All this emphasises the central problem, which enters at the level of the censor. How and where does the censor operate? There are three possibilities. Either the censor operates (i) in the sub-system (the Ego-nucleus), or (ii) it operates between the main system and the sub-system, or (iii) it is located within the main system itself. Suppose we analyse each of these possibilities in turn.

The last possibility - that it is located within the main system - does not seem feasible for the reason that this would imply that the operations of the censor are accessible to consciousness, and as noted, this does not happen. The other two options will be examined in more depth.

The first option - that the sub-system is self-censoring - does not involve the same problems that exist if we supposed that the censor were located within the Id. The latter case would be problematic given the nature of the Id - Id impulses strive for satisfaction and are not bound by practical or other such considerations. But the Ego-nucleus is part of the Ego, and hence would be governed by many of the principles that would govern the Ego. So it would, I would argue, be bound (to some extent) by the reality principle. This would imply that it would not strive for immediate gratification, but would be able to delay the fulfilment of its desires. Furthermore, we may assume that it has access to the contents of the main system, even if the converse is not true. The sub-system would thus be able to monitor both its own beliefs or wishes, and the beliefs of the main system, and keep out of the main system - and hence out of consciousness - any of its beliefs that would conflict with the main system. On this description the sub-system is in a sense more in control in the sense that it has access both to its own beliefs and to the beliefs of the main system. This inter-connection is only one-way. Let us take the case of the cheated girlfriend, and see how this type of analysis could accommodate the factors. We have two central beliefs held by Suzy: Sam is faithful (p) and Sam is having an affair with the waitress (-p). Each belief is accompanied by its set of related beliefs, and by its various causes. The belief that -p is the undesired belief, and it is held unconsciously by the sub-system; the belief that p is held by the main system. The sub-system with the undesired belief that -p is still able to affect Suzy's behaviour - she avoids going to the restaurant where Sam and his lover are likely to meet. Now the censoring activities here

would be confined to the sub-system itself - in this sense the system is covert, and it hides its operations from the main Ego system. The censor hides both its existence and its operations - and hence the existence and operations of the ego-nucleus - from the main Ego system.

The second option is to position the censor between the main system and the sub-system. The censor would have access to both systems, without necessarily sharing the beliefs of either. It would be able to know what beliefs would be compatible or incompatible with the belief/desire system of each system, and would censor beliefs accordingly. The censor would work either by not allowing beliefs from the sub-system into the main system, or vice versa. However, as the main system is the conscious one, the censoring procedure would work only one way - from sub- to main. It would also be responsible for repressing beliefs that were once in the main system into the sub-system. The censor would not hold any of these beliefs in itself, and thus Sartre's paradox - that the censor would have to hold both the desired and the undesired beliefs simultaneously - would not arise. The censor acts as a kind of screen, filtering beliefs but not accepting any imprints. With regard to the activities of the censor, the question arises as to how its activities remain hidden from the main system. We can postulate that the censor is covert - that it hides its activities from the main system of its own accord, to facilitate its activities. Another possible option is that the workings of the censor do not themselves require censoring. It could be that the activity of censoring is simply not the kind of thing that can be known to the subject. The censor is not an agency, but simply a structural mechanism, and as such cannot become conscious. Certainly the holding of a belief can be known or not, but it may be that the actual activity of censoring is just by nature inaccessible to consciousness. Whether this is a feasible suggestion or not I am not sure.

Now the question to ask at this stage is whether there is a major difference between these two options - namely that of the censor lying within the sub-system or between the sub- and the main systems. Do the operations of the censor differ that much in either case? In each case the censor is covert, and in either case it has access to the beliefs of both systems. The central difference is that if we posit the censor as being internal to the sub-system then it is in principle possible that contents of the main system would enter the sub-system - that is, the workings of the censor would be one way only - from sub- to main. This is not a problem, however - in fact this is what normally happens. It is the main system that must be guarded from beliefs that it cannot accept; the unconscious system would not be threatened by beliefs of the main system. So in the Suzy example, the main system could not share those beliefs that led Suzy to believe that -p. The question that arises here is - does the censor create the split, or does it simply guard it, and ensure that it is maintained? If only the latter - as seems more likely if the censor is internal to the sub-system - then we need to ask how the split is formed - what are the principles that govern the creation of the sub-system of the Ego? As this sub-system is not a separate agency, are there rules of genesis and functioning that govern it in the same way as the agencies are governed? These questions are difficult to answer, because Freud himself never developed the notion of the Ego-nucleus. I think that it would be fair to say that the nucleus does not share a special place in the development of the individual as the other agencies, and that its role is defined in terms of providing a space for the existence of beliefs formed by the Ego that would be threatening were they to become conscious.

Now - to get back to the original problem - namely whether positing the censor within the sub-system is a workable suggestion. If the censor is confined to the sub-system, it can do an excellent job of preventing beliefs and desires that are already within the sub-system from entering

the main system, but it is not clear how it is able to censor the contents of the main system, and relegate undesirable beliefs into the sub-system. That is - the censor here would not act as a principle of division, and it would not be able to limit the entry of mental items from main to sub-systems, assuming this should ever be necessary. We would need to postulate another principle or agency of division. However, if we postulate the censor as lying between the sub- and the main systems, then the censor will have access to both systems, and it will be able to both censor the contents of the sub-system from entering the main system, and it will be able to seize upon undesirable beliefs in the main system and place them in the sub-system. I do not think that it is necessary to posit that all of the contents of the main system are conscious all the time - they are, however, easily accessible to consciousness, and we must assume that the censor is able to relegate these undesirable beliefs to the sub-system before they have a chance of becoming conscious. I think that the situation where the censor exists between the sub- and the main systems is the more feasible one, and it is also a simpler solution. It does not force us to define the nature of any particular system as involving essentially the function of censorship; we can simply posit the function of censorship as existing between systems, and we need not be concerned with attributing beliefs to the censor itself. The censor is not an agent, nor need it possess its own belief states. It must, however, be sensitive to the workings of the three agencies, and its operations would certainly be directed by the demands of the various agencies - in particular by the Superego and the Ego. The workings of the censor may be mechanistic in operation, attempting (and not always succeeding) to prevent beliefs or desires from a lower, more primitive system from entering into a higher system. The censor itself need not decide what elements should be assigned up what agency in order for secondary repression to occur; this would be done by the agencies themselves. So Id desires press for gratification, and, depending on the receptivity of the

Ego under the rule of the Superego, the demands of reality and considerations of optimal functioning for itself, censoring activities will be accordingly more or less powerful. The censor may be "agent-sensitive" - it may acknowledge the desires of the more developed system that is doing the repressing - but it will not have desires or motivations for repression of its own.

Another problem that arises is how the belief in the sub-system is able to influence behaviour - how is it able to produce the belief in the main system that is used to rationalize this behaviour? It seems that we can posit a solution along the following lines, taking the Suzy example as the case in point. Suzy desires to believe that p (Sam is faithful). The belief that he is having an affair with the waitress is unacceptable to Suzy, and is relegated to the sub-system. However, if this unconscious belief were not able to influence behaviour in any way, then Suzy might attend the restaurant and be confronted with Sam's infidelity - an undesirable consequence. So Suzy's avoidance of the restaurant ensures that the unconscious belief will not become conscious. It is not simply the case that the belief that -p keeps Suzy away from the restaurant, but rather that confrontation with sufficiently strong evidence would be sufficient to make the unconscious belief conscious, and hence is avoided. In this sense the sub-system is able to regulate behaviour so as to ensure that one of its beliefs does not enter the main system. There are various ways one can attempt a solution. One way would be that, whenever the thought of going by the restaurant enters Suzy's mind, she feels anxious, and a non-threatening reason is created by way of explanation. The feeling of anxiety arises out of the repressed but active belief -p that threatens to become conscious. Another suggestion would be that the sub-system itself motivates for the production of a modified rationalization within the main system. The censor notes the belief in the sub-system, and allows for or causes a rationalization to appear in the main

system so that the agent will act in accordance with her desires but still not become aware of her real reasons for acting. It can be noted here that the more elaborate structure of the Freudian thesis has an advantage over the functional account: Freud's complex picture of the mind is able to account for just such cases, where unconscious contradictory beliefs are still able to affect the individual's behaviour. In the functional account, contradictory beliefs are kept strictly apart and do not interact, and so the functional account would not be able to account for Suzy's situation. So the very complexity of the Freudian account may provide motivation for its acceptance over a more simple account of the mind, given that this complexity is able to provide for a fuller explanation of human psychology.

Another query relates to the relation between the evidence at hand and the formation of the beliefs within the various systems. Is it possible for beliefs to be formed in the sub-system directly, or must they first be formed in the main system and then repressed? That is - can beliefs in the sub-system be directly influenced by the evidence at hand? One argument against this is that if we allow too much independence to the sub-system, we are in danger of creating a very definite rift between the sub- and main systems, and thereby creating another agency. We also run the risk of creating two consciousnesses within the agent. It seems that the main system must have priority with respect to the formation of a belief in direct response to the evidence. What we must simply emphasise is that not all beliefs that are formed in the main system are conscious; they may be unconscious but not repressed. Before they have a chance to become conscious, and thereby threaten the agent with anxiety, they are repressed by the censor and sent to the sub-system. Some of the dynamics of the topographic theory seem to be re-emerging in this description - there seems to be a variation on the structures of UCS, PCS, and CS

emerging on a smaller scale within the Ego - but in a way, I think, that is not damaging.

A question raised earlier needs to be addressed here: what is the principle that underlies the formation of the creation of the sub-systems? How many sub-systems are there? With respect to the latter question, it could be suggested that there could be an endless number, if we postulate that sub-systems arise whenever incompatibility of one kind or another arises. However, I do not think that this is necessary. As a sub-system of the Ego, we need not regard its operations as concerned with the total functioning of the agent. Also, it is a primitive part of the Ego, and as such would not be that affected by the constraints of reality as the main system would be. Most importantly, though, confinement within the sub-system is the result of repression; that is, confinement within the sub-system ensures that the undesirable beliefs will not reach the main system. It is therefore the main system that has to be protected from undesirable beliefs, by having their access to consciousness barred, and not the sub-system. There is thus no good reason for us to postulate an infinite amount of sub-systems. There certainly may arise differences of belief within the sub-system, but this can be handled simply by postulating the existence of multiple belief systems existing alongside one another - in much the same way as belief systems exist in the main system.

We need, however, to be able to provide an account of the creation of the Ego-nucleus. The principle of the creation of this structure cannot simply be one of logical incompatibility, as then we would run into the problems noted in chapter two. In keeping with the Freudian tradition, we should be able to give a genetic and not merely a functional account of the Ego-nucleus, and as noted above, a detailed account is sadly not available in Freud's writings. We can say some things about the Ego-nucleus. Firstly, it is an undeveloped part of the Ego that has become

split off from the main Ego as a result of the Ego's defensive action against the contents of the Ego-nucleus. Such contents may be Id desires that the primitive Ego had attempted to act upon at an earlier stage in the Ego's development, but whose drive towards fulfilment are now becoming too threatening, or have become inappropriate, for the more developed rational Ego. So we see that the Ego-nucleus may, at some stage in the subject's development, have constituted the identity of a primitive Ego or self; as the subject develops, this primitive Ego fails to develop alongside other aspects of the Ego, and eventually becomes rejected and cut off from the main Ego. This position in turn makes the Ego-nucleus a suitable receptacle for beliefs and desires that the Ego now cannot itself accept. So the Ego defends itself against some part of itself; it treats the Ego-nucleus as "other". This rejection of or defence against a part of itself lies at the root of self-deception, and makes it possible. We can also say that the sub-system, or undeveloped Ego-nucleus, is repressed by a censor operating between the sub- and main systems; that we need not postulate the existence of more than one sub-system; that the principle of division should not be one of logical incompatibility that would involve us in the dilemmas of chapter two. What we need is an exploration of Ego psychology within the Freudian tradition, and such a comprehensive account is sorely lacking. Such an account would explain how the split in the Ego comes about, and how the mechanisms of defence operate within the Ego itself.

Another problem that emerges is that of agency: if there is a split within the system of unity - the Ego - itself, can we speak of the notion of agency at all? If not, then are we able to hold onto the notion of self-deception? Self-deception seems to involve a single unified subject involved somehow in the process of deceit - hence the paradox. One possible way out is to answer by saying that the notion of agency that we have - as that of a single unified entity - is misguided. In reality, a

subject - the Ego - is divided, and there are levels of censorship that ensure these divisions. That is, we can reply that the Freudian thesis requires us to loosen our definition of agency. As regards self-deception, perhaps explaining the problem from within a Freudian paradigm does indeed dissolve the paradox by turning it into one of deception by one part of the Ego over the other. So perhaps the solution is reached by changing our notion of the constitution of the self. Even so, the Freudian conception of the self is not simply that of divided agencies warring against one another. Censorship, repression, deception do not occur without cost. Symptoms may arise, feelings of anxiety - these would indicate that some principle of unity is at work, and that the self is threatened or disturbed by radical splits within it. But it must be admitted that the attempted solution to the problem may involve changing the face of it somewhat.

In concluding this chapter, it may be helpful to sketch the workings of self-deception within the Freudian paradigm used above. How does self-deception occur? Let us use Othello's predicament as an illustration. Othello loves Emilia but he cannot accept this desire as suitable for a man of honour, married as he is to the fair Desdemona. Othello desires that he should love his wife, and that this love should exclude any love for Emilia. We may, for the sake of the argument, assume that Othello has loved Emilia in the past - his desire for Emilia was not confined to the Id, but has been taken up by the Ego - and that he continues to love her. What Othello's Ego must do is repress the desire - and the corresponding belief about the desire - for Emilia, and confine the belief/desire to the F -nucleus. Here, unconscious and out of bounds of the main system, the belief concerning Othello's love for Emilia remains hidden. It may affect Othello's actions, but in ways that would not be recognized by Othello as being indicative of the fact that he loves Emilia. It would not become conscious. For instance, its presence may

result in Othello's "protesting too much" that he loves Desdemona and not Emilia - here the Ego would mount a strong defensive campaign against the acknowledgement of the belief/desire. Or, on the other hand, Othello may especially love and praise Desdemona's eyes <u>because</u> they remind him of Emilia's, this association would necessarily remain unconscious. Othello is self-deceived - he denies that he possesses certain beliefs and desires, and relegates them to the sub-system in the Ego. The censor, in accordance with the demands of both the Superego (which contain Othello's moral standards), and the Ego (which assesses the desire for Emilia in terms of both moral standards and real situations), and prevents the belief/desire from entering the main system, from becoming conscious and being acknowledged by the subject, although the repressed belief may still continue to affect the agent's behaviour. So repression is able to favour both systems to some extent - it keeps harmful ideas from upsetting the equilibrium of the Ego, while it allows partial fulfilment to the contents of the nucleus. The self-deceptive nature of these operations should be apparent: the Ego rejects certain beliefs and desires, and denies that they belong to itself. It deceives itself with respect to its own beliefs, with respect to what it acknowledges as belonging to itself, and consequently with respect to the reasons it provides for the subject's behaviour in the world.

CONCLUSION

The paradox of self-deception presented us with various problems at various levels, and there exist in response various attempted solutions to the problem. There are two broad lines of attack, both of which, in their own way, attempt to dissolve the paradox. The one attempt tries to argue that self-deception may not involve the holding of two contradictory beliefs simultaneously by an agent, and that the agent is involved in some type of irrational behaviour other than deception. The problems associated with this line of approach were examined, and the approach was seen to be unsatisfactory. The second line of attack took the deception seriously, and examined the nature of the self that could allow for the possibility of self-deception. Here a Freudian approach to the problem was examined in some detail. I would like to examine some problems that may arise in connection with this approach.

A central problem raised in connection with the approach taken in this paper is that of agency. The Freudian model of the subject presents problems for a picture of a unified, coherent agent acting rationally and for the best possible reasons. Part of the reason for this is that the Freudian thesis attempted to explain, not only rational behaviour, but irrational behaviour as well. It attempted to understand the reasons for the behaviour of individuals who were most often unable to supply these reasons themselves. Freud's thesis presents problems for the notion of agency for two reasons. Firstly, the notion of unconscious beliefs that are able to influence both behaviour, and the acquisition of conscious beliefs, must do some damage to the notion of agency. If we behave for

reasons that are unconscious, then we will be unable to articulate these reasons, and so will be unable either to give the true reasons for our behaviour, or we may act without knowing why we do so. Secondly, the description that Freud gives of the self is essentially that of a divided subject; not only is the mental life of the subject divided into what is conscious and unconscious, but the subject consists of three agencies which are for the most part unaware of the existence or operations of one another, and which are often in conflict. One agency may find itself performing - often unconsciously - a certain action for the sake of, or in protection against, the beliefs or desires of another agency. Furthermore, as was seen, there exists fragmentation within the very agency that was to serve as the principle of unity - the Ego itself. This presents a problem for a solution to the problem of self-deception : the paradox of self-deception arises precisely because we appear to have a single unified agent who holds contradiatory beliefs. By splitting the self into multiple agencies, we solve the problem, but at the risk of losing the original problem altogether. For now we do not have a single self that is deceived; rather we have one agency being deceived by another. We could reply that the notion we have of agency is misguided - as nice as it may sound, the notion of a single unified rational agent acting in full knowledge of all the reasons for her actions is but an idealization of human psychology; in reality, human subjects fall far short of any respectable notion of agency. This solution would be to acknowledge the divisions postulated by Freud as real, and it would thus challenge the notion of agency altogether. Whether this is satisfactory is far from clear; the problems and puzzles associated with this area, however, would require further and independent research, and go beyond the scope of this paper.

In my paper I proposed the Freudian thesis as a better alternative to the Davidsonian approach, for problems noted in chapter two. From chapter

four, where the notion of the Ego-nucleus was proposed and examined, it appears that the Freudian approach - certainly as used here - has similarities with that of Davidson. We have the picture of an Ego that splits in order to accommodate irrationality, just as in Davidson's analysis we have a split in the subject occurring whenever irrationality occurs, the purpose of the divisions within the self is to allow for the co-existence of contradictory (and hence irrational) beliefs. The question is: how does the Freudian line differ from the Davidsonian, or functional, approach? Does it present similar problems? The first difference to note between the theories is that the Freudian thesis is able to provide us with a genesis that is lacking in the functional account. Part of the problem with the functional account was that it did not provide us with an explanation of either the genesis or the mechanism of division; we were not told how the divisions arise, nor were we told much about the connections that may or may not exist between divisions. Also, since the divisions appeared whenever irrationality appeared, their function was to avoid creating any tension within the agent; the irrational beliefs were safely accommodated within separate partitions. Thus the element of tension that arises in the case of self-deception at the level of the entire agent is not sufficiently explained. If the non-interaction of the irrational beliefs were guaranteed by the existence of the schisms - and this is, after all, their function - we would not require a motivational element that seems to exist and that seems to steer the agent away from becoming conscious of an undesirable belief. A second problem with the functional account was that it defined its mental divisions simply in terms of the job it wanted them to do, and hence no independent corroboration was given for their existence.

The Freudian account differs from the functional accounts in both respects mentioned above. Firstly, in the Freudian approach, we are provided with both a genesis of the mental divisions, as well as some indication of

their interaction; indeed, the problem of repression is a central concern in Freudian theory. The history of the mental systems provided by the Freudian account provides a fuller analysis of the mental compartments with which we are dealing. Secondly, there appears to be independent corroboration for Freud's postulations. Since the Freudian account aims to provide a full picture of the nature of the mental, his theory provides explanations of various other types of symptoms, and work done in these areas provides us with alternative evidence for assessing his theory. The existence of an unconscious, say, seems to be a theoretical notion whose existence or functioning seems able to be independently corroborated. The Freudian account is able to provide a fuller picture of the mind - the nature of the divisions, a genetic as well as a functional account of their existence and operations, an examination of the notion of repression - and in this way differs from, and escapes some of the problems of, the functional account. In the functional account, we were given an indication of the product - irrationality is accommodated by divisions within the mind - but we were not given much of an account of the process that led to this state. The Freudian account, on the other hand, is able to provide an account of the processes that underlie irrational states, and is, for this reason, a fuller account. This ability to provide a fuller description of the workings of the mind would favour the Freudian thesis over the functional one, and provide more support for its acceptance.

The area of similarity, however, may arise within the discussion of ego-psychology. The existence of the Ego-nucleus ensures that any tension between beliefs, any irrationality that cannot be accommodated by the main system, can be resolved by accommodating the disruptive or unwelcome belief in the sub-system. So, whereas in Davidson we have the existence of various partitions that arise in order to accommodate irrationality, in the Freudian thesis we have the Ego-nucleus that plays this role - here

is the similarity between the two accounts. Also, since there has not been much development in the field of Ego-psychology, and since the notion of the Ego-nucleus is a fairly recent one, we cannot offer much of a genetic account of these ideas, apart from that sketched above in chapter four. In order to highlight the distinction between the Davidsonian and the Freudian accounts, we would need to sketch the process or mechanism of division at the level of the Ego. In both accounts, irrationality is accommodated by the creation of one - or, in the case of Davidson, more - partitions that house the irrational belief, and keep it apart from the main system. How are the pictures different for Davidson and Freud? Firstly, in terms of the censorship between partitions, Freud, I believe, presents a fuller picture of the workings of the censor. The censor operates under the dictates of both the moral (Superego) and rational and pragmatic (Ego) centres of the mind; the system is designed around keeping unwelcome or threatening ideas out of the main system and out of consciousness. It is true that we are not presented with a full and detailed picture of the censor itself - Freud's discussions are centred more on the nature and desires of the various agencies. It may be that such a discussion is not important for Freudians - as long as we understand the motivation for censorship in terms of the agencies, we may not require an analysis of the censor itself - it may operate on fairly simple mechanistic grounds, the nature of which are not central to Freudian theory. It is true that we have much discussion within Freudian theory of the various defence mechanisms - repression, projection, denial, reaction formation. These tell us what the censor does - projection is, for instance, the denial that an undesirable belief is had by oneself and a belief that it is had by some other person - but we do not know what the nature of the censor is in terms of our demands. The censor does not have a personality - if this allies the theory with the functional account, then the similarity between the two accounts is perhaps unavoidable.

Another area of examination is the creation of partitions within the main system of the agent. In Davidson we are given the principle of division but not the mechanism of division. Is there a similar lack in the Freudian account? We know the reason for the creation of the division - it is the result of the Ego's repudiation of the contents of the Ego-nucleus. We know some of the reasons for this rejection - firstly the Ego-nucleus contains beliefs and desires that would threaten the Ego were it to acknowledge them, and secondly the Ego-nucleus is a more primitive, less rational, aspect of the Ego whose level of functioning differs markedly from the later more developed Ego, and hence is not easily accommodated by the Ego. This failure of simultaneous development would indicate the division between the Ego and the Ego-nucleus. Although this picture does not present us with a fully mechanistic account of the division, it appears to be a more complex one than the Davidsonian account. Also, the genetic as well as the functional nature of the account avoids the alleged circularity of the functional account. Indeed, I am not entirely certain whether the Freudian account is a more fully developed version of the Davidsonian account, or whether it is able to provide an alternative - and better - account. I hope that the areas of similarity and dissimilarity have been adequately sketched out in order for the problem, at least, if not the solution, to be clear. This may be as useful.

Another question that needs to be asked is whether the process of self-deception as described within the Freudian schema is mechanistic, and if so, whether this mechanism presents problems for our conception of self-deception as both motivated and teleological. I do not think that the two are necessarily mutually exclusive. Of course, mechanistic systems are for the most part not intentional at all, as well-thought out and executed plans cannot rely solely on simple mechanistic procedures if they are to succeed. However, it may be the case that a certain type of

behaviour is motivated or teleological, in the sense that the actions and decisions of the agent are directed towards a certain goal, and yet many of the means used by the agent to affect these consequences may be automatic. It is not mechanism as such that is problematic to any analysis of self-deception as long as the mechanistic procedure is sustained within a theory that takes into account the intentional nature of the project. In fact, one of my objections against the Davidsonian thesis was that it did not provide us with a satisfactory picture for doing precisely this. So I do not think that a mechanistic account of, say, repression, is necessarily problematic for the Freudian analysis of the problem, because I think that the Freudian account examines the process of repression (necessarily) from the perspective of the overall wishes and desires of the agent.

In assessing the Freudian approach, the central question that needs to be asked is this: does it solve the problem of self-deception? Have we come any closer to resolving the paradox? The answer to this question turns on whether the concept of the <u>self</u> employed here is not problematic. In the analysis, the method of deception was taken seriously - the paradox was not solved by showing that, when we speak of self-deception, we are speaking of a self engaging in some other kind of activity mislabelled as "deception". Instead, the resolution took the form of creating divisions within the self, and arguing that, when we speak of self-deception, what we are really saying is that one part within the self is deceiving another. Now - apart from the problem of agency mentioned above - the question is whether we are left with a self sufficiently coherent for us to posit as <u>self-deceiving</u>. The paradox of self-deception depends on our having a unified self in a state of self-inflicted deception; the resolution adopted here has been to explain self-deception as deception of one part of the Ego over another. The question now is whether we have a real solution, or just an alternative picture. The question is difficult to

answer: if we are content to accept an alternative notion of the self, then the resolution to the problem may be acceptable; if, however, we review Sartre's problem - namely that the deceiving self must know the truth that it hides from itself - then we may insist that the self described here cannot be divided. If we are content to alter our conception of the self and see it as consisting of various agencies - or of the Ego as split in the way postulated - then we may see the Freudian approach as a way to resolving the paradox; but, if we insist on the latter description, then the analysis defended here will be unsatisfactory.

B I B L I O G R A P H Y

Arlow, J. A and Brenner, C — Psychoanalytic Concepts and the Structural Theory International Universities Press: New York; 1964

Audi, R — "Self-Deception, Action, and Will" Erkenntnis 18 1982 pp 138-158

"Self-Deception and Rationality" in Martin (Ed) op. cit. pp 169-194

Canfield, V and Gustavson, D. F — "Self-Deception" Analysis 23 1962/3 pp 32-36

Champlin, T. S — "Self-Deception: A Reflexive Dilemma" Philosophy 52 1977 pp 281-299

Davidson, D — "Actions, Reasons, and Causes" in Davidson, D Essays on Actions and Events Clarendon Press: Oxford; 1980

De Sousa, R. B — "Self-Deception" Inquiry 13 1970 pp 308-321

Demos, R — "Lying to Oneself" Journal of Philosophy 57 1960 pp 588-595

Elster, J (Ed) — The Multiple Self Cambridge University Press: Cambridge; 1986

Fingarette, H — Self-Deception RKP: London; 1969

"Alcoholism and Self-Deception" in Martin (Ed) op. cit. pp 53-67

Foss, J

"Rethinking Self-Deception" American Philosophical Quarterly 17 1980 pp 237-243

Freud, S

"A Note on the Unconscious in Psychoanalysis" (1912) in On Metapsychology: The Theory of Psychoanalysis (Pelican Freud Library Volume 11) Penguin: Middlesex; 1984 pp 45-57

"The Unconscious" (1915) Ibid pp 159-222

"Repression" (1915) Ibid pp 139-158

"Negation" (1925) Ibid pp 435-442

"The Ego and the Id" (1923) Ibid pp 339-407

"Splitting of the Ego in the Process of Defence" (1940 [1938]) Ibid pp 457-464

"The Dissection of the Psychical Personality" in New Introductory Lectures on Psychoanalysis (Pelican Freud Library Volume 2) Penguin: Middlesex; 1983 pp 88-112

Gardiner, P

"Error, Faith and Self-Deception" Proceedings of the Aristotelian Society (PAS) 70 1970 pp 221-243

Haight, M

"Tales from a Black Box" in Martin (Ed) op. cit. pp 244-260

Heil, J

"Minds Divided" Mind 47 (392) 1989 pp 571-583

Martin, M. W (Ed)

Self-Deception and Self-Understanding University Press of Kansas: Kansas; 1985

Mele, A. R

"Self-Deception" The Philosophical Quarterly 33 pp 365-377

"Pears on Akrasia, and Defeated Intentions" Philosophia 14 1984 pp 145-151

Irrationality: An Essay on Akrasia, Self-Deception and Self-Control OUP: New York; 1987

Mizi, M

"Self-Deception" Philosophy and Phenomenological Research 34 pp 576-585

Mounce, H. O

"Self-Deception" PAS 45 1971 pp 61-72 ak2

Palmer, A

"Characterising Self-Deception" Mind 88 1979 pp 45-58

Paluch, S

"Self-Deception" Inquiry 10 1967 pp 268-278

Pears, D

Motivated Irrationality OUP: New York; 1984

Pugmire, D

"'Strong' Self-Deception" Inquiry 12 1969 pp 339-346

Rorty, A

"Adaptivity and Self-Knowledge" Inquiry 13

"Belief and Self-Deception" Inquiry 15 pp 387-410

Sieglar, E. A

"Self-Deception and Other-Deception" Nous 2 1968 pp 759-764

Sorensen, R. A

"Self-Deception and Scattered Events" Mind 94 1985 pp 64-69

Szabados, B

"Wishful Thinking and Self-Deception" Analysis 33 pp 201-205

"Self-Deception" Canadian Journal of Philosophy 4 1974 pp 51-68

Wilson, T. D

"Self-Deception Without Repression" in Martin (Ed) op. cit. pp 95-116

Wollheim, R and Hopkins, J (Eds)

Philosophical Essays on Freud Cambridge University Press: Cambridge; 1982