

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.
2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.
4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.
5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms

300 North Zeeb Road
Ann Arbor, Michigan 48106

76-1491

SHAPIRO, Daniel, 1944-
SELF-DECEPTION.

The City University of New York, Ph.D., 1975
Philosophy

Xerox University Microfilms, Ann Arbor, Michigan 48106

© 1975

DANIEL SHAPIRO

ALL RIGHTS RESERVED

SELF-DECEPTION

by

DANIEL SHAPIRO

A dissertation submitted to the Graduate
Faculty in Philosophy in partial fulfillment
of the requirements for the degree of Doctor
of Philosophy, The City University of New York

1975

This manuscript has been read and accepted for the Graduate Faculty in Philosophy in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

June 23, 1975
date

John J. Cairns
Chairman of Examining Committee

July 14, 1975
date

Marshall Cohen
Executive Officer

Charles Landmann
Geoffrey Myers
Abraham Edel
Supervisory Committee

The City University of New York

Acknowledgements

I would like to thank my friends and teachers for all the help that I have received while working on Self-Deception. I am thankful to my friends, especially Dan Jacoff, for their willingness to put up with me during the ups and downs of the last year, and for supporting me in some of my deceptions while abusing me of others. To Dr. Oscar Sternbach I owe the glimmer of awareness that I now have of my own personal dealings with self-deception. And to Professor Charles Landesman, my sponsor, I owe a special debt of thanks for all the help and encouragement that he gave through all the stages of my work on self-deception.

It was clear that I must repair to our town next day, and in the first flow of my repentance it was equally clear that I must stay at Joe's. But when I had secured my box-place by to-morrow's coach, and had been down to Mr. Pocket's and back, I was not by any means convinced on the last point, and began to invent reasons and make excuses for putting up at the Blue Boar. I should be an inconvenience at Joe's; I was not expected, and my bed would not be ready; I should be too far from Miss Havisham's, and she was exacting and mightn't like it. All other swindlers upon earth are nothing to the self-swindlers, and with such pretences did I cheat myself. Surely a curious thing. That I should innocently take a bad half-crown of somebody else's manufacture is reasonable enough; but that I should knowingly reckon the spurious coin of my own make as good money! An obliging stranger, under pretence of compactly folding up my bank-notes for security's sake, abstracts the notes and gives me nutshells; but what is his sleight of hand to mine, when I fold up my own nutshells and pass them on myself as notes!

Dickens, Great Expectations,
p. 245

Contents

Chapter I: Introduction	1
Chapter II: Self-Deception and Unwarranted Belief	15
Some History	
Self-Deception and Other-Deception	
Self-Deception as Unwarranted Belief	
Self-Deception and Other Forms of Unwarranted Belief	
Conclusion	
Footnotes to Chapter II	
Chapter III: The "Paradox" of Contrary Beliefs	57
Belief and Paradox	
The "paradox" of Contrary Beliefs	
The Oddity of Contrary Beliefs	
Moore's Paradox	
Footnotes to Chapter III	
Chapter IV: The Self in Self-Deception - 1	111
Desires, Motives, Intentions and Self-Deception	
Some More History	
The Approaches to Self-Deception	
The Self as Desire	
Siegler	
Russell	
Butler	
Footnotes to Chapter IV	
Chapter V: The Self in Self-Deception - 2	169
The Self as Will	
Freud	
Marx and Mannheim	
Marx	
Mannheim	
Kierkegaard, Sartre and Fingarette	
Kierkegaard	
Sartre	
Fingarette	
Conclusion	
Footnotes to Chapter V	

Contents (Continued)

Chapter VI: The "Paradox" of Self-Induced Belief	234
Belief, Desire and Will	
Belief and Attention	
Conclusion	
Footnotes to Chapter VI	
Chapter VII: Conclusion	254
List of Works Cited	258

Chapter I: Introduction - The Problem

Like ostriches, men can hide their heads from what they do not want to see. There are times when we ignore, avoid, or simply fail to realize that what we know contradicts the way we want and believe things to be. In such cases our beliefs and lack of knowledge are of our own making: men, it seems, deceive themselves.

One need not look far to uncover cases of self-deception. Tolstoy tells of how Count Ilya Rostov returned from a business trip to discover that something had happened to his daughter Natasha. The count quickly puts things right with a little self-deception. "From the pretended illness, from his daughter's distress and the troubled faces of Sonya and Maria Dmitrievna, the count saw clearly that something had gone wrong during his absence; but it was so terrible for him to imagine anything discreditable occurring in connection with his beloved daughter, and he so prized his own cheerful tranquillity, that he avoided asking questions and did his best to persuade himself that there was nothing very much wrong or out of the way, and his only regret was that her disposition would delay their return to the country." (War and Peace, I, p. 698)

Traditionally self-deception has been of interest to philosophers because of its importance in ethics. Moralists are often frustrated by the fact that while men understand the point of their edifying homilies, they often avoid applying to

themselves the moral of their tales. Men can be very astute and perceptive in knowing the good and bad in others, and uniquely blind to seeing it in themselves. Joseph Butler, in a sermon on self-deceit, quotes the biblical story of King David's adultery with Bathsheba and the murder of Uriah, her husband, as an example of the evil done from self-deception. The very upright King David committed both adultery and murder with a clear conscience, and only became aware of his misdeeds when the prophet Nathan related the parable of the rich man with many flocks of sheep, who takes the pet lamb of a poor neighbor in order to feed a visiting stranger. In answer to David's moral outrage and harsh judgement of the rich man, Nathan replied: "Thou art the man." (The story is from The Second Book of Samuel, verse 12, lines 1-15.)

Although it is not often noticed, self-deception can also serve useful purposes. Thomas Mann remarked that if he did not deceive himself about the length of the books he wrote, and think of them as short and manageable when he began, he would never have been able to have started his greatest works.

One of the most interesting and complicated examples of self-deception is given by Sartre. A young woman is on her first date with a man whose intentions are all too clear to her. She finds his interest flattering, but the urgent desire which lies behind it is unacceptable to her. Instead of having to decide by answering his overtures, she disarms his flattery

of its sexual connotations, and interprets his words as the sincere and respectful thoughts of a charming companion. She, of course, realizes full well what he is about, and it is just because of this that she is able to deceive herself about him, able to put off his real intent by interpreting every word and deed as she does, and so avoids an unpleasant scene or see him as a desirable companion should she decide to respond. (pp. 55-56)

There are of course simpler, everyday cases of self-deception. Freud tells the story of a long overdue visit to his brother in England. Because of his brother's unyielding ways, Freud was asked to postpone a day's visit to Amsterdam, where he wanted to see Rembrandt's paintings at the Rijksmuseum. In taking the quickest route from Munich to the English sea-side, Freud had to change at Cologne for the Rotterdam express. At Cologne he could not find the train, and after a fruitless search he missed his connection. The result was that he had an extra day which allowed for a visit to the museum. On the following day, while traveling across England, Freud recalled that as he stepped off the train at Cologne he had seen a train waiting under a large white sign for the Rotterdam Express, a sign and train that he had earlier failed to notice in his rush to make the necessary connection. (Freud, 1901, pp. 227-228)

There are many forms and types of self-deception. We deceive ourselves about many different things, and we do this in various ways and for various reasons. We can deceive our-

selves about our circumstances, thoughts, feelings, and desires. And we can deceive ourselves about these things through perception, as when the ill see and feel themselves getting well, or reasoning, as when our moral failings by means of shady argument appear to us a boon to our fellow man, or memory, as when past failures take on new lustre with time and become the warm remembered joys of how things used to be. And our deceptions can serve many purposes: they can blind us to what is painful, help us feel good, or enable us to accept or reject what would otherwise be unpalatable. Sometimes our deceits are good and serve us well, and sometimes they are harmful to ourselves and others.

To try to list all the nuances and varieties of self-deception is probably an impossible task. To try to fit all the ways in which we deceive ourselves into one broad definition is also probably impossible. What I want to do is to argue for one particular form of self-deception. I want to argue that self-deception can involve ignoring what we know while inducing ourselves to believe what we want. In self-deceit we ignore or deny some warranted belief that we find unacceptable. What makes this belief unacceptable is that it conflicts with what we want. But not only do we ignore what we know, we persuade ourselves into holding the belief that justifies what we want. Not all of these conditions are involved in all cases of self-deception. But they are involved when self-deception is used in its fullest sense. These are cases when we deceive ourselves by taking beliefs that are

warranted by reason and using them to create self-made unwarranted beliefs that we find more acceptable.

While occurrences that fit this description appear strange, it is undeniable that they take place. Think of Count Rostov. He saw that something was wrong with Natasha. Instead of finding out what had happened, he persuaded himself that it was of only minor importance, and so preserved his peaceful state of mind. King David knew that adultery was wrong. He also knew that sleeping with Bathsheba was a case of adultery. Yet he committed adultery while managing to think that what he did was right and proper. Freud saw the train he wanted and yet was able to miss it, thus allowing for a visit to the Rijksmuseum. The woman in Sartre's example is able to neatly balance her companion's desire with the belief that he is only interested in quiet conversation. She holds both beliefs until the time when her own desire leads her to decide which belief is correct. All of these examples can plausibly be seen to involve inducing ourselves into a belief we want although we know better.

The problem is that while this view of self-deception appears to me to be both natural and right, it also seems to be hopelessly paradoxical. It strains our normal conception of things in a number of ways. The idea that we can persuade ourselves to hold a desired belief contradicts the view that beliefs are not of our own making but the result of evidence

and reason. All the desire in the world doesn't give any reason to hold a belief. So the view that we can induce ourselves into holding a belief appears to be an impossible undertaking. The idea that we can persuade ourselves to believe what we know isn't so is doubly odd. Even if we could induce ourselves into holding a belief, how can we do this when we already know better? If we believe one thing this seems to preclude the possibility that we can hold the opposite belief. The picture of the self-deceiver as holding contrary beliefs appears to be self-contradictory. That we can persuade ourselves to hold an unwarranted belief also makes self-deception look as if it were a rational activity. It is as if the self-deceiver used what he knew in order to develop a scheme to believe what he wanted. Self-deception seems to involve an impossible combination of impossible activities.

It is the paradoxical aspects of self-deception that have generated the recent philosophical interest in self-deception. Definitions are given and strategies are devised with an eye to ridding self-deception of paradox. The paradoxes come first and the definitions and explanations that follow attempt to make self-deception fit with our normal view of belief and reason. The key feature of these accounts is that belief while knowing better is dropped as a condition for self-deceit. There have been many accounts of self-deception along these lines, but all of them follow one or another of two basic approaches.

The first of these is simply to interpret self-deception as a case of unwarranted belief. Some additional characteristic is added to distinguish the unwarranted belief in self-deceit from the general category of cognitive errors. It is a belief held in circumstances where what is believed is obviously unwarranted, or it is a mistaken belief when we should know better. The self-deceiver is not seen as either inducing this belief in himself or as knowing that his belief is unwarranted. This is the position of Siegler (1) and (2), Canfield and Gustavson, and Penelhum. While self-deception involves unwarranted belief, this account does not distinguish self-deceit from any other form of cognitive error. Why does the self-deceiver fail to notice what is apparently so obvious? And why does he make just the error that he does? Imagine King David excusing his conduct by explaining to Nathan that he simply failed to realize that Bathsheba was married although he knew that Uriah was her husband. In cases where these accounts do seem plausible, they also seem to implicitly imply that the self-deceiver does know better. Thus, on analysis, what makes for adverse circumstances for belief is the fact that the believer sees these circumstances as showing that his belief is unwarranted. This tends to imply not just that the self-deceiver should know better, but that he does know better.

The second account of self-deception that avoids attributing contrary beliefs is more plausible. It construes self-deception

as a case of motivated belief or wishful thinking. This is the view of Gardiner (1), Siegler (4), Demos, Shea, Russell, and Butler. Self-deception is seen as a case of holding an unwarranted belief which we presumably wouldn't hold if it weren't for some interest of ours. It is not that we believe what we want although we know better, but that interest, not evidence, guides us to our belief. This account does explain why we make the errors we do. It does not explain why the belief is unwarranted for the believer, nor does it explain what happened to the knowledge which might contravene the desired belief. If King David didn't know that Bathsheba was married and that adultery was wrong, why should thinking that all was alright be deceptive? This account also does not do justice to the subtlety of our deceptions. The picture here is that of a man who is blindly driven by desire.

Both of these accounts leave out what is central and most interesting in cases of self-deceit. Unless we get ourselves to believe what we do in spite of what we know, or just because of what we know, self-deception is not involved. Denying what is paradoxical about self-deception involves denying that there is self-deception. The only way that these accounts can be defended is if it is impossible for self-deception to be anything more than a case of ignorance or wishful thinking. The justification for this must be that self-deception as persuading ourselves to believe what we know isn't so simply

can't occur. The denial that there is self-deception seems to be a case of what is being denied; we do manage to believe what we know isn't so.

The only way to make sense of self-deception - instead of implausibly denying its occurrence - is to accept what occurrences of self-deception imply. Room must be made for the seeming impossibility of self-induced unwarranted belief. Self-deception forces us to revise the picture of ourselves as consistent and integrated in what we believe, and as self-consciously coming to the beliefs that we hold. It is these assumptions that make self-deception appear so strange. If we are rational and self-conscious about our beliefs, then self-induced belief is impossible. If our beliefs are integrated, then contrary beliefs cannot be held. What makes self-deception paradoxical is that it does not mesh with our ordinary view of ourselves. This is why all attempts to fit self-deception in with our normal view of things are bound to fail. Self-deception shows that we are not as rational as we normally think.

Behind the positions that deny the reality of self-deceit lies a certain model of mind. On this model the mind is seen as transparent. It is assumed that we know what we believe and why we believe it and so holding contrary beliefs or persuading ourselves to hold the beliefs we hold appears as paradoxical. We cannot believe what we see or realize is contradictory nor can we accept those beliefs that appear to be self-induced.

This view of mind is built into the way we think of belief and into the psychological assumptions about how the mind works, and so self-deception as getting ourselves to believe what we know isn't so appears as either logically or empirically impossible. While this model reflects how we picture the would-be believer, it is neither logically required nor true in fact. It is difficult to give up this view of things for in doing so we are forced to accept the possibility of our own irrationality. Self-deception requires this. Instead of holding to our cherished rational view of ourselves, we need to admit that conflicting beliefs and irrational processes are as natural to us as rational and consistent belief. There is no way that this can fit with our ordinary experience of things and our conscious view of ourselves, and this is ultimately why self-deception appears to be paradoxical.

Those who do account for self-deception do so by accepting the irrational nature of man. Men can purposefully avoid what they know and believe what they want although they know better. Our thoughts are seen to be the natural outcome of irrational processes; unknown fears and desires can manipulate what we know and induce us to believe what we want; the mind is not transparent to itself nor unified in its beliefs. Given these assumptions self-deception appears neither strange nor anomalous. Of the accounts that develop this line of thought, it is Freud's that is the most interesting and inci-

sive. Although the picture of ourselves as naturally prone and even requiring self-deception is not a pleasant one, nor one that is easily accepted, it cannot be faulted on logical or empirical grounds. It is the view that we must in the end come to accept if the paradox of self-deception or the paradoxical denial of self-deception is to be avoided.

In the pages that follow I argue that self-deception involves inducing ourselves to believe what we want although we know better. The literature on self-deception is outlined. The accounts that hold that self-deception is either unwarranted belief or wishful thinking are criticized as inadequate. The paradoxical aspects of self-induced belief and believing something while knowing better are analyzed and are seen to result from an overly rational model of mind. Freud's account of self-deceit, and the positions of Marx and Mannheim, and those of Kierkegaard, Sartre, and Fingarette, are presented and seen to fit the general account of self-deception as self-induced unwarranted belief.

The position that self-deception is merely a case of unwarranted belief is considered first. The relationship between unwarranted belief and self-deception is reviewed, and the positions of Siegler, Canfield-Gustavson, and Penelhum are analyzed. The conclusion reached is that what makes our beliefs unwarranted in the sense required for self-deception is that

the believer in some way knows that his belief is unwarranted. This leads to the question of whether holding a belief while knowing better, with its assumption that we can simultaneously hold contrary beliefs, is possible. In the third chapter, "The 'Paradox' of Contrary Beliefs", I argue that the paradoxical nature of contrary beliefs is based on the mistaken assumption that the mind is transparent and integrated with respect to its beliefs. This view of mind is neither logically required nor empirically justified, and so the way is cleared for self-deception as involving beliefs held while knowing better.

Self-deception is not simply a matter of holding unwarranted beliefs. Its most striking feature is that the unwarranted belief serves some interest or desire. Without this addition the fact that we hold the unwarranted belief that we do is incomprehensible. In the fourth and fifth chapters we consider the self-induced character of the beliefs held in self-deceit. In the fourth chapter we deal with the view that self-deception is merely a matter of belief held from desire. This is the view that self-deception is a case of motivated belief or wishful thinking. The most interesting positions are those of Siegler (4), Russell, and Butler. They are analyzed and criticized in detail. By not holding that the self-deceiver knows that his belief is unwarranted, these positions fail to explain why the desired belief is unwarranted

for the believer, or why the wishful belief appears tailor-made to suit what is unacceptable to the believer. Self-deceit is not simply believing what we want, but doing so when what we want conflicts with what we know. In the fifth chapter the views of Freud, Marx and Mannheim, and Kierkegaard, Sartre and Fingarette are outlined. The common feature of these accounts is that self-deception is purposeful. It enables the self-deceiver to avoid the particular beliefs that he finds unacceptable. The fact that the believer knows better is built into the self-induced or motivated character of his beliefs. Here self-deception involves both self-induced and unwarranted belief, and so the full sense of self-deception is involved. In the sixth chapter, "The 'Paradox' of Self-Induced Belief", the problematic character of inducing or persuading ourselves into holding our beliefs is considered. I argue that subjective factors alone can be the source of belief, and so we can believe something when the belief involved has no support and is in fact contrary to what we know. There are no limits placed on what we can believe because of either what we know or because desire and will are not the proper basis of our beliefs. Self-deception as unwarranted beliefs held in order to realize what we want, although contrary to what we know, is possible.

In the end the view that self-deception involves self-induced unwarranted beliefs comes down to the fact that there

is no other adequate account of self-deception. Without the element of belief held while knowing better the need for self-deceit and the intricate moves of the self-deceiver remain incomprehensible. Without the element of inducing in ourselves the beliefs that we want, the reason for self-deception remains a mystery. The only reason for rejecting this view of self-deception is that it does not fit our conscious, rational view of things. Since this view is not logically required and rests on questionable assumptions about the nature of mind, it cannot force us to accept implausible interpretations of self-deception. As odd as it may seem on our ordinary view of ourselves and our beliefs, the fact is that our beliefs may be counterfeit truths of our own making that we gladly accept as real currency.

Chapter II: Self-Deception and Unwarranted Belief

One of the necessary conditions for self-deception is erroneous belief. At a minimum self-deception involves holding erroneous beliefs that we foist on ourselves. Thus Dorothy Sayers comments that Guido da Montefeltro, a "counsellor of fraud" burning in the eighth circle of Dante's Inferno, has deceived himself in thinking that Dante is one of the damned, for he has only himself to blame for his error. (p. 245)

This view of self-deception is attractive because of its broad scope and great simplicity. It is simple because it avoids the complicated issues involved when self-deceit implies either motivated belief or knowing better. It is broad since all forms of self-deception involve erroneous belief. The question is whether this view adequately distinguishes self-deceit from other cognitive errors, such as failing memory or stupidity, and whether more is lost than gained by lumping together all forms of self-deception under this minimal view. The defenders of self-deception as erroneous belief claim that more astute views are impossible, for self-deceit as motivated erroneous belief or erroneous belief while knowing better only leads to paradox.

In this chapter we consider the view that self-deception is merely a matter of erroneous belief. We shall see that self-deceit does involve erroneous belief. We shall also see

that self-deception must be more than this if it is to be distinguished from other sources of error. In particular, self-deception appears to involve the holding of erroneous belief when we know better. Believing something and knowing better is, it is true, well on the road to paradox, but whether or not holding contrary beliefs is paradoxical will have to be left to the next chapter. Our limited conclusion will be that self-deception can only be considered a matter of erroneous belief if in fact it can't be anything more.

Some History

Recent philosophical discussion of self-deception begins with the 1960 article by Raphael Demos entitled "Lying to Oneself". Demos argues that self-deception is similar to lying: it is, he says, a case of lying to oneself. According to Demos lying involves successfully inducing someone to believe as true what one knows to be false, and doing this intentionally. On this account there are three necessary conditions in lying: intention, result, and knowledge. By treating deceiving oneself as analogous to deceiving others by lying to them, Demos arrives at his account of self-deception:

Self-deception exists, I will say, when a person lies to himself, that is to say, persuades himself to believe what he knows isn't so. In short, self-deception entails that B believes both P and not-P at the same time. Thus self-deception involves an inner conflict, perhaps the existence of a contradiction.

Demos, p. 588

With this definition the stage was set for the debate that followed. After all it is strange to think that we would persuade ourselves to believe what we know isn't so. First, some may think it odd that we would want to ignore or avoid the truth. Once this is accepted the question becomes how we intentionally get ourselves to hold or give up our beliefs. The very idea of manipulating oneself into one's beliefs rather than discovering what one believes or having evidence guide one to a belief is puzzling. But say we found the answer to this puzzle, we then would be faced with the harder puzzle that in self-deception we try to persuade ourselves to believe what we know isn't so, and so edit knowledge with an eye to ignorance. It would seem that whatever sense we can make of intending to believe certainly can't apply to beliefs we already know are false. How can we even try to believe what we know isn't so? But even this is not the worst of it. Let's say we were very clever and made sense of trying to believe the contrary of what we knew: how could we succeed? The notion of successful lying to oneself seems to clearly be beyond the bounds of sense. It would, after all, entail believing as the deceived what we didn't believe as the deceiver, and this certainly has the sound of paradox about it.

Demos, of course, was aware of all this. He and those after him were struck by the fact that if self-deception was

a case of lying to oneself then it entailed believing and not believing the same thing at the same time. Demos tried to deal with this by arguing, in a manner modeled on Aristotle's account of akrasia, that there were different ways of believing. He thought he could deal with the problem of contrary beliefs by arguing that we can have contrary beliefs if we fail to notice that they conflict. (pp. 592-595) Now this is no solution if holding contrary beliefs is in fact logically odd or paradoxical, for the psychological expedient of failing to notice what is paradoxical or logically impossible doesn't, of course, make it any the less paradoxical or impossible.

The articles that soon followed Demos' were quick to make this point. (E.g. see Siegler, 1, p. 472) Writers like Siegler, Canfield and Gustavson, and Penelhum, argued that since interpreting self-deception as analogous to lying led to "attributing contrary beliefs to a man at the same time" (Siegler, 1, p. 470), "was odd sounding if not contradictory" (Canfield and Gustavson, p. 32), or seemed to "generate paradoxes" (Penelhum, (1), p. 258), then clearly self-deception was not like lying at all.

This in turn led to non-paradoxical renderings of self-deception that did not entail the holding of contrary beliefs. The more or less common conclusion was that self-deception was essentially just a cognitive misdemeanor: it involved holding

erroneous beliefs in circumstances where we should have, but didn't, know better. An example of this is poor Tom, mistakenly thinking that fidgety Sally is bored with the movie when she is really bored with her ignorant Tom.

Siegler's view was the simplest. Anyone who has an erroneous belief that it is unreasonable to have is deceiving himself. Self-deception simply involves ascription of responsibility to the believer for faulty beliefs. "It is not that the "deceived" knows that his belief is (or was) unwarranted; the point is that he ought to know (have known) better." (1, p. 474, and 2, pp. 33-34). Canfield and Gustavson's view is a bit more complicated. According to them, self-deception involves holding a belief in circumstances that are adverse to that belief. "All that happens in self-deception, we suggest, is that the person believes or forgets something in certain circumstances." (p. 35) Their addition to Siegler's account is that they explain this view of self-deception as being analogous to self-commands. Thus belief in adverse circumstances is construed as similar to commanding ourselves to do what we do not want to do, and this, unlike lying to oneself, supposedly doesn't involve holding contrary beliefs. Penelhum's view is rather sophisticated. Like Demos, he thinks that self-deception involves the appearance of contrary beliefs and inner conflict, but, unlike Demos, he doesn't think it can

involve actually having contrary beliefs. The solution is that the self-deceiver "partially satisfies the criteria for belief and also those for disbelief - in particular he will tend to declare his disbelief in that which he sees the evidence points." (pp. 258-259) The self-deceiver partially satisfies criteria for contrary beliefs and so appears to believe what he at the same time knows isn't so. But this is illusive. The conflict is only on the level of evidence for what is believed. The self-deceiver does not actually hold the contrary beliefs that his behavior tends to imply.

As we shall see, none of these non-paradoxical renderings of self-deception work. It will be shown that when these positions seem to give plausible accounts of self-deception it is because they entail holding false beliefs while knowing better and so are as paradoxical as Demos' original. And when these positions are interpreted as implying that self-deception merely involves holding erroneous beliefs, they are non-paradoxical but also inadequate.

Self-Deception and Other-Deception

Before going on to consider the account of self-deception as holding deceptive or unwarranted belief, and what is wrong with it, some words on the relationship between self-deception and other-deception are in order.

Ever since Demos' original article the point of focus for much of the literature has been whether self-deception is or isn't analogous to other-deception.¹ While this may have some value in pointing out what is odd about self-deception, it seems to be taken for granted that once this question is answered all else will fall in place. And this I think is one grand and misleading over-simplification. The assumption here seems to be that it is much easier to be clear on what is involved in other-deception, and that other-deception is philosophically similar to self-deception. This, or something like this, is implicit in the view that self-deception can be approached via other-deception. Unfortunately this just isn't so. Other-deception is just as complex as self-deception, and little is to be gained in dealing with self-deception by comparing it with other-deception.

First, other-deception is not so clear a notion as is usually supposed. There are many different senses in which it can be said that we have deceived others. We can deceive them by lying to them, where this implies intention, knowledge, and result, as held by Demos. But we can also be said to lie without knowledge, as when I intend to mislead but unknowingly tell the truth; and we can also be said to lie without result, as when we are caught lying; and we can even tell a lie without intending to, as when I unwittingly pass a lie on to you.

And there are many forms of other-deception besides lying. We can deceive others by saying something false with strong conviction as to its truth. We can deceive by misleadingly emphasizing the parts, or by being too selective in presenting the facts or because our make-believe is mistaken for belief, and so on. We can also deceive others by telling them half-truths, the whole truth, and even by being painfully honest and telling nothing but the truth with neither intention nor success in misleading them. Thus I may deceive someone by telling them the half-truth that they shouldn't call late because I plan to go to bed early, and while this is true, I also plan not to be at home. We may well imagine a clever spy who immediately tells the whole truth about his secret upon being captured, realizing that what he says will be the last thing his interrogators will believe. And when a wayward husband immediately confesses his thoughtless afternoon's escapade to his wife, she may feel that he has deceived her although he told the truth without intending or succeeding in misleading her.²

The point of all this is that 'other-deception' is not a term for any neat category of similar occurrences. Some cases of other-deception involve intention to deceive, others not; some knowledge that we are deceiving, others not; in some deception is in fact perpetrated, others not; and in all

of these cases the truth may have played a role in bringing about the deception. All of these subtle nuances and differences can also apply to the supposedly straightfoward and simple case of other-deception involved in lying. Thus not one of Demos' three necessary conditions may in fact be necessary for a case of lying, although one or more may be sufficient. There is no list of conditions that will lead to one definition sufficient for all cases. Consequently, saying that self-deception is either like lying or other-deception, as Demos does, or saying that it is not like other-deception, as Siegler, Canfield, Gustavson and Penelhum do, doesn't come to very much. For as there are many forms of lying and many more forms of other-deception, both positions can be right, and so nothing is settled by this.

There is the possibility that by close attention to words and astute analysis all the intricacies of other-deception can be made clear. And with this done the necessary similarities and differences between self-deception and other-deception can be drawn, and so simple-minded accounts avoided. But even if this was done, a comparison of self-deception and other-deception could only focus on the unique or common aspects of either, and this needn't explain anything. Thus if self-deception is equated with lying, the question still arises whether self-deception is like those cases of lying where intention and knowledge are involved. One can't simply say that self-deception can't be like this sort of lying for this

involves intending to deceive oneself and holding contrary beliefs, without explaining why this is impossible. The most that can be hoped by such a comparison is that the more familiar notion of other-deception may cast some light on what are the problems and characteristics of the less familiar self-deception. Since there is also a danger that such comparison may lead to over-simplifications, the wisest course is to deal directly with self-deception and by-pass the circuitous route through other-deception.

A basic assumption behind all this seems to be the view that self-deception is the mirror image of other-deception. Many writers tend to assume that other-deception is the basic notion, the word that wears the pants as Austin might say, or the logically prior concept as those doing conceptual analysis might say. Hence it is felt that one must first be clear on other-deception in order to be clear on self-deception, just as someone might claim that to be clear on what is unreal one first must be clear as to what is real. In fact there is little reason to think that self-deception and other-deception are paired terms; there may be senses of one that are not analogous to the other.

The assumption of analogy was, perhaps, due to thinking that both self and other-deceptions were twin off spring of a common parent, or that other-deception followed self-deception. But paternity and family resemblances are often tricky to track

down. We have already seen that assuming family resemblances between other-deception and self-deception may create many orphans if not down right genocide. But what about parentage? It does seem as if other-deception, with its reference to other people and intersubjective contexts, is logically prior to self-deception, which involves self-reference and the world of the individual. This follows from the common scenario that language is only possible in an intersubjective community where words are learned and first used in public contexts and only later, if at all, applied to one's own case. This model may thus be taken as showing that the use of 'self-deception' comes from applying to ourselves the 'deception' that was first used in relation to others. In other words, self-deception is literally either a form of lying to ourselves, or it is not like lying at all and is a self-referential use of some other public expression, say error or ignorance.

Now while this general view of the logic and use of language is very possibly correct, it seems to be misused in this particular context. For one thing, while the existence of language may entail intersubjective contexts, it doesn't follow that every word gains its use from such contexts. Thus words having essentially private reference seem to be logically possible, but what is not possible is that these are the first and basic linguistic concepts. Thus words with self or private reference, like first person psychological terms, cannot have formed the elementary vocabulary of any language user, but

could only have developed after public reference was well instituted. But it doesn't follow from this that first person use of psychological terms are logically derivative from third person use of those same terms, only that they are related to the use of some public terms. In fact it seems likely to me that in the case of some complex psychological terms it was the private use that first developed by giving a unique metaphorical twist to some public expression, and that this "private use" then gained public acceptance. In particular this seems possible in the case of terms such as 'humility', 'sincerity', 'concern', and 'self-deception'.³

If it is possible that language develops in this way, then 'self-deception' cannot have been one of the first words used on entering the linguistic community. But it of course doesn't follow from this that the public use of 'deception', 'other-deception', was the public term or concept that led to the more sophisticated concept of self-deception. And it might even be that of the two it is 'self-deception' that was used first and from which the notion of 'other-deception' develops. Thus it is possible that self-deception, not other-deception, wears the linguistic pants, although it could not have been one of the first pairs of linguistic pants. And in general, it seems one is on very tricky grounds when one argues that certain terms or concepts ante-date or are logically prior to others, for it is unclear how such claims can be

philosophically settled.

Whatever our view of how the concept of self-deception develops, it seems plausible to me that in fact self-deception may actually, not just linguistically, wear the pants. We may all be very adept at self-deception even if others never deceived us, and if there is other-deception it might consist in others giving us enough rope so that we deceive ourselves: other-deception may prove to be a form of self-deception. Thus it can be argued that the clearest cases of other-deception involve knowing what someone wants to think, or what one believes. In deceiving others we play to a person's beliefs and desires in order to get him to believe what we want him to believe. Someone who knows little about us is unlikely to deceive us. It is for good reason that con men look for "marks", men whose thoughts and desires they are familiar with. I suspect that the man who cares little for money, self-esteem, women, etc., although not particularly suspicious or bright, will not be an easy mark for the deceiver. And the man who cares about nothing at all, if there could be such a man, will be no mark at all. If this makes sense then perhaps deceiving others is really just giving them a chance to deceive themselves.

Put more generally: suggesting and influencing others may in the end come down to auto-suggestion and self-interest. The best we can do for good or ill is to manipulate others so that they manipulate themselves. This may be why it is usually

a vain and fruitless task to try to boldly change others or even ourselves; and that success, if it comes along these lines at all, comes only to those patient few who understand themselves and others, and so are able to try to satisfy what they or others are after in a way that is more acceptable to themselves.

To sum up: To treat self-deception as either similar to or different from other-deception may involve the mistaken assumption that other-deception is a clearer and simpler notion, and the mistaken view that the questions and problems of self-deception could be settled by such a comparison. There also seems to be no good reason to think that self-deception either logically or linguistically presupposes understanding or having the concept of other-deception. And there is even some reason to think that self-deception may, at least psychologically, be the more important notion. Thus it seems much wiser to attempt to explain and understand self-deception by considering the problems and occurrences of self-deception, not those of other-deception. To think otherwise, as do Demos, Siegler, Canfield-Gustavson and Penelhum, is very likely to put the cart before the horse, and be stampeded into confusion and error rather than clarity.

Self-Deception as Unwarranted Belief

Leaving aside the question of whether self-deception is or isn't like lying, we return to self-deception as unwarranted belief. This is at the heart of the claim that self-deception isn't like lying to others. Thus Siegler's view that self-deception involves believing what we shouldn't, Canfield and Gustavson's view that self-deception involves holding a belief in adverse circumstances, and Penelhum's view that self-deception involves holding an unwarranted belief in circumstances that justify another belief, all treat self-deception as first and foremost a case of mistaken or unwarranted belief.

The view that self-deception is basically a case of cognitive error allows for a very straightforward account of self-deceit. Here ascriptions of self-deception need not refer to other beliefs or motives of the deceived believer. This is why Sayers can comment that Guido had deceived himself, although she knew nothing of his thoughts and motives. On this view, all we need to know is the epistemic conditions surrounding the belief in order to know that someone has deceived himself. The problems of contrary belief and motivated belief are thus neatly avoided. Since most positions on self-deception hold that we need to know about the "self" in order to know if it is deceived, they may be termed "self directed" accounts of self-deception. They require their proponents to take a stand on the tricky questions of contrary

beliefs and motivated beliefs. Since treating self-deception as unwarranted belief need only refer to the external circumstances of belief and not the thoughts and the motives of the believer, it may be considered an "other-person" or outsider's epistemic account of belief. The question is: What is involved in such a view, and is it a plausible account of self-deception?

Most writers tend to agree that self-deception involves cognitive processes, such as belief, and some sort of failure in such truth directed processes, and so mistaken or unwarranted belief. Thus, as we shall see, Butler, Mannheim, and even Sartre, hold, with current philosophical accounts, that self-deception involves problems in belief. But this doesn't seem to be the case with all accounts of self-deceit. Freud can be interpreted as considering self-deceit the outcome of strong unconscious motives or intentions where this doesn't involve any cognitive processes, except if we happen to frame beliefs about ourselves or our actions when such non-cognitive forces are at work. The neat and natty dresser may be deceived about his deep down interest in muck and mess, although he never gave it a moments thought and isn't about to entertain beliefs about it. Similarly, Kierkegaard considers a form of self-deception where a man is "untrue" to himself, where this doesn't involve mistaken or unwarranted beliefs, but rather a failure to be ethically consistent or a lack of faith in God. With an eye to views such as these, Fingarette

argues that self-deception is not essentially a cognitive occurrence at all, but rather a matter of will, a case of volition or action, not thought. In self-deceit we will or act wrongly, we fail to be what we are, and so lead false, inauthentic, or meaningless lives, rather than have false beliefs.

Whether or not such views lead to coherent accounts of self-deceit, I do not think they actually remove self-deception from the context of cognitive error. While the stress here is on the motive behind the action or on a certain view of self, the notion of mistaken or unwarranted belief is waiting in the wings. Without being able to cash in talk of unconscious motives or true selves into descriptions of ourselves that we can believe, the notion of being false to ourselves or leading untrue lives really doesn't make sense. The natty dresser is deceived because he doesn't understand his "true interest", and the man who is untrue to his moral or religious self is so, at least in part, because he has mistaken beliefs about what is of value. In general, it doesn't even begin to make sense to say someone's life is false or his existence is meaningless, unless we believe that life really is about something and that there is a true source of value. General talk about the meaninglessness of life is either emotional discontent, empty rhetoric, or bombastic cant. Underlying any sensible claim about the meaning of life or personal existence there are always cognitive claims about the nature of self and experience.

If it makes sense to say that our lives are false and our existence empty it also makes sense to say that we believe wrongly or fail to understand "real life" and "true self". Talk of the emptiness and lack of meaning in an individual's life makes sense, and unfortunately is often true. Thus even in cases where self-deceit is construed as basically due to unknown or mistaken motives, it still involves ignorance or unwarranted belief, on the part of the individual.

So self-deception involves mistaken or unwarranted belief. As it stands this is a rather broad and innocuous view that characterizes all cognitive errors. We need to narrow down the sort of contexts and occurrences that involve self-deception, and thus clarify the particular sort of cognitive error that is symptomatic of self-deceit. The key here is that self-deception doesn't just involve getting something wrong or holding an unwarranted belief, but that we get things wrong when we could and normally would get things right. At the least self-deception requires contexts where we end up in error when belief is appropriate and truth is expected. I think some light can be shed on self-deceit as unwarranted belief by considering the general contexts in which it is possible.

The person who deceives himself must in some way be interested in the truth. If I refuse to read a botany book, and do not erroneously claim to know about botany, I protect my ignorance without deceiving myself. Similarly, the woman

who does not want to know of her boyfriend's escapades, and the mother who refuses to hear of her son's mischievous acts, do not deceive themselves as long as they knowingly choose ignorance and stay within bounds of their unwarranted beliefs. They simply do not want to know, perhaps because they are afraid of the hurt or useless confrontation that knowledge would bring. A person who does not want to know what he can and should know is a fool. A person who avoids holding beliefs and doesn't make any claims as a result of his ignorance, may be someone who is reprehensible, dangerous, or dumb, but he isn't a self-deceiver. In general, where the truth knowingly goes unsought or lies unclaimed, deception is not possible, because erroneous belief needn't be involved. But if a person is wrong in thinking he doesn't want to know, or is mistaken about the interest or importance of knowledge, or holds unwarranted beliefs because of knowledge avoided, then we are back to the context of erroneous belief, and so back to the possibility of self-deception.⁴

In a similar way, phantasy, dreams, jokes, pretense, and make-believe, can all be distinguished from self-deception. All these occurrences involve contexts where truth is out of place, and so belief in its cognitive sense is inappropriate. Thus an essential feature of make-believe is that the usual connection between belief and truth is either severed or suspended. In make-believe or pretense what is "believed" is

✓

either not considered in truth related terms, or is considered "as-if-true" in a realm where our sense of reality is suspended, or is thought of as removed from our other thoughts and beliefs. It is because of this that I can make-believe that I am a dog, a chair, or even myself. In these and similar contexts, since relation to the truth is either attenuated, suspended, or removed, mistaken or unwarranted belief cannot get hold, and so self-deception is out of place. But if I make believe about something with the intention of getting someone else to believe it, or if I am taken in by my own pretense, then we can again speak of deception. This is because we are no longer in a state of make-believe or pretend, but have returned to the realm of truth-related belief, and so can be wrong.⁵

Self-deception as a form of unwarranted or mistaken belief is not only distinguishable from contexts and occurrences where the truth is unsought or out of place, it is also distinguishable from contexts and occurrences where there is an insurmountable obstacle to acquiring the truth. Thus self-deception is unlike cases of delusion, utter confusion, imbecility, or total irrationality. In the case of delusion, for instance, a man cannot tell what is or isn't so and when the ability to make this distinction goes, the concept of deception cannot be applied. The man who really thinks that he is Christ or Napoleon, depending on his position of and on arms, would be thought to have lost his wits rather than have deceived himself. Perhaps the often keen ability of psychotics

to see through the world's madness and the deceits of others can be explained by the inability to apply the truth to their own case: madness allows for great insight because the truth never hits home. This may explain the common paradox, that so impresses R. D. Laing, that the insane often appear saner and more insightful than the rest of us. And it may also help explain the tragic metaphor of insight that leads to blindness, as with Oedipus, or the madness that allows for insight, as with King Lear's garland crowned vision into how this world really goes. It may be more than poetic justice that enables fools and madmen to speak the truths of others; they may be the only ones able to see it and the only ones we are willing to hear it from. And it may be more than just a dramatic point that insight into ourselves often leads to blindness, madness, and other tragedies. But again, wherever delusion, confusion, etc., are compatible with belief, then again the possibility of mistakes arises and so too the possibility of deception. Thus the man who comes to believe he is Christ because it allows him to avoid what he knows is a cruel world, and enables him to garner sympathy and other medicine, may be either a self-deceived or clever man, and so mental illness may be a myth as Szasz points out.⁶

There are countless subtleties and difficulties involved in giving even partial accounts of such complex occurrences as make-believe and delusion. Fortunately our thumbnail

analyses of these concepts needn't be astute in order to clarify our sense of self-deception. The important point is that self-deception can't be said to occur when the would-be believer either removes himself from a truth related context because he is not interested in knowing as do auditors in the school of ignorance is bliss; or because he suspends the relationship between belief and truth, as in make-believe; or when he loses all contact with reality so that he no longer understands what is true and what isn't, as in delusion and madness. In all cases where the normal conditions for belief are clearly absent, it would be at best misleading to speak of mistaken belief and so deception.

The "normal conditions of belief", whatever they may be, are hard to delineate, and because of this there will be many gray areas where it will not be clear whether 'self-deception' is applicable. What is clear is that if 'self-deception' is a form of mistaken belief, then it can only be appropriately applied to those who satisfy the normal conditions of belief. Among the normal conditions of belief are: being reasonably well informed, ability to reason and observe, having the appropriate linguistic skills, not being overly fatigued, or drugged, and so on and on. There are obviously enormous difficulties in knowing if these conditions are fulfilled, even when dealing with those usually considered normal in normal situations.

So far we have it that self-deception involves mistaken or unwarranted belief by normal believers in contexts where belief is appropriate and true belief is expected. What is essential is that self-deception involve deceptive belief. But deceptive belief is ambiguous. It can mean a belief that is in fact false, or it can mean a belief that is not justified by the available evidence for that belief,⁷ in which case the belief needn't be false. The ambiguity stems from the fact that 'deception' can apply either to the falsity of the resulting belief, or it can apply to unjustified procedures leading to belief. Since process and result are related, the difference between the two gets blurred. The distinction can be kept clear if we speak of 'mistaken belief' when a belief is in fact false, and 'unwarranted belief' when the belief is not appropriate given the available evidence and correct reasoning.

Now if deception is taken to involve mistaken belief, then no matter how ludicrous one's belief, no matter how contrary it is to what should be believed, if the belief turns out to be true, we cannot have a case of self-deception. The emphasis here is on the resulting belief; if it be true, there is no self-deception. A man who disbelieves his doctor's diagnosis that he has cancer, although all the tests, symptoms, and consultations indicate cancer, certainly has an unwarranted belief. But if it should turn out after all that he doesn't

have cancer, his unwarranted belief wasn't self-deceptive. Penelhum seems to hold such a view of deception and falsity when he says that if an unwarranted belief turns out to be true it is a matter of faith, not deception. ((1), p. 259)

If, on the other hand, we hold that deception involves unwarranted belief, then we can talk of self-deception even if the belief ultimately proves true, as long as it was unjustified for a rational believer given the available evidence. This approach emphasizes what we do with the evidence we have, rather than the truth or falsity of the belief we end up with.

In general, I think this second interpretation of deceptive belief is preferable. For one thing it allows self-deception to have a wider scope than that of false belief, and this seems intuitively correct. Our would-be cancer patient is, after all, very possibly deceiving himself in holding his unwarranted but correct belief. This interpretation also stresses an active sense for 'self-deception', implying as it does that the source of deception is the believer, and not simply what is believed. This too seems intuitively plausible for self-deception. If all mistaken beliefs were also unwarranted beliefs, then we could simplify things by reducing self-deception to cases of unwarranted belief.

Unfortunately not all mistaken beliefs are unwarranted. We can speak of mistaken and warranted belief whenever right reason, available evidence, and the appropriate criteria for

belief lead to mistakes in normal circumstances or when proper caution was taken. It can happen. It happened to Guido da Montefeltro. Such cases may be termed passive occurrences of self-deception. But there are much more important exceptions to self-deception as unwarranted belief. Not all beliefs are based on evidence or involve reason. In such cases we cannot, of course, hold an unwarranted belief, for there is no evidence to mis-evaluate or reason to wrong. Examples of this are errors in belief about our thoughts and feelings when we do not come to such beliefs by observation or inference.

There is a final point about deceptive belief. To hold that self-deception involves either mistaken or unwarranted belief entails that self-deception always applies to some belief one has and shouldn't. But self-deception can also apply to the omission of beliefs we should have. The cuckold who fails to believe that his wife is unfaithful, when the neighborhood is loud with rumors and his head is full of shady late night excuses, is very apt to be self-deceived. It may be argued that omissions of belief always turn out to be lack of belief because the contrary belief is held, and so turn out to be cases of mistaken or unwarranted belief. But there is a difference. We can omit a belief without having a contrary belief. This clearly is involved in self-deception through forgetfulness, when we fail to have a belief we should have and don't hold a contrary deceptive belief. This can also

be the case of our cuckold if he fails to believe that his wife is unfaithful without believing that she is faithful. This could well be the case as he might prevent himself from thinking about such an indiscreet matter.

What is essential to self-deception is failure to know or believe what the available evidence or conditions warrants. This covers all the relevant cases; it includes mistaken belief, unwarranted belief, and omitted belief when we normally would be expected to have the right belief. If we add to this the proper context for self-deception we have the following full account of self-deception as deceptive belief: self-deception involves failure by a normal, interested, and able believer to have a correct or warranted belief where available evidence or conditions would normally lead such a would-be believer to the right or justified belief. Simply put: self-deception involves an unusual or abnormal case of error or ignorance. For the sake of expedience, in the place of this long-winded formulae, self-deception will be simply said to involve unwarranted belief. All the views we are now considering fit a part of this general formulae. Thus Siegler's "mistaken belief we are responsible for", Canfield-Gustavson's "belief in adverse circumstances", and Penelhum's "disbelief in that which he sees the evidence points", all, in their way, involve failing to hold a correct or warranted belief where such beliefs are expected.

The question now is: is this a plausible account of

self-deception? The answer is: no. The reasons follow.

Self-Deception and Other Forms of Unwarranted Belief

There are many so-called normal sources of error and ignorance. Most mistakes are due to: faulty reason, difficult material, lack of knowledge or familiarity with the subject, inattention, bad memory, naivete, fatigue, confusion, stupidity, and on and on. The list of things that can lead us into error is long and well traveled. Such a list includes the sort of things that cause normal believers, interested and able to know, to hold unwarranted beliefs, although in circumstances that warrant and usually lead to correct beliefs. As we saw, self-deception is a similar source of error. And so we begin to run into difficulty. All the sources of error seem to be covered by the list of usual mistakes. But if this is so then either self-deception is no different than naivete, confusion, stupidity, etc., or, if it is different, an additional condition is needed to distinguish it from these more common sources of error and ignorance. And how is self-deception as an "abnormal" or "unusual" source or case of unwarranted belief to be distinguished from the run-of-the-mill varieties of error and ignorance?

This problem is particularly acute for Siegler and Canfield-Gustavson. Siegler characterized self-deception as simply any unwarranted belief we shouldn't have. Canfield

and Gustavson's view is a little more complex: self-deception is a belief in adverse circumstances. Now both "unwarranted belief" and "belief in adverse circumstances" applies to error due to simple ignorance, confusion, a naivete, as well as to self-deceit. There are two possible approaches to this problem. One is simply to deny that self-deception is distinguishable from other forms of unwarranted belief. The other way is to specify an additional characteristic that applies to self-deception and not to other forms of unwarranted belief.

Siegler takes the first approach. For Siegler, our cuckold is self-deceived simply because he is ignorant of what he should have known. His position amounts to the heroic denial that there is anything special about self-deceit. The result is an elegant and simple, if hopelessly inadequate, account of self-deception.

I deceive myself...is often accompanied by reference to a time I did something which I now regard as silly, wrong or unbearable. But it is not a claim that I intended to mislead myself into believing what I know to be false, or that I actually did believe what I know to be false. It is more like a reference to an attitude or action about which I now feel that I should have known better, but did not.

Siegler (1), p. 474

Our claim that Jones "deceived" himself would be a roundabout exhortation to Jones to "face the facts".

Siegler (2), p. 42

This account is heroic because cases of self-deception can be special. Thus take Siegler's view that ascriptions of self-deceit are really reprimands to face facts. Now what happens when such reprimands are given and understood, and our advice accepted? Perhaps the usual result is that the unwarranted belief is given up, and the facts faced. But it can also happen that while agreeing with our reprimand, and even while seeing that his belief is unwarranted, a person continues to hold his unwarranted belief. And this can happen with self-directed as well as other-directed exhortations. Sometimes good advice well taken needs to be repeated. In such cases clearly a second "reprimand" of self-deception is not adequately characterized as merely imputing unwarranted belief. We now need to add that the deceiver seems in some way to know better. But this of course takes us back to the problem of contrary beliefs that Siegler wanted to avoid. In the end, the only plausibility to Siegler's view is if self-deception logically cannot be more than normal errors. And so its final reckoning must wait on whether or not contrary beliefs are paradoxical.

The second approach is that self-deception is different from the common sorts of error. Here the attempt is made to specify additional characteristics that, while differentiating self-deceit from other forms of unwarranted belief, do not re-institute the paradox of contrary belief that the account of

self-deception as unwarranted belief was meant to avoid. Penelhum and also Canfield-Gustavson take this line of approach. Canfield and Gustavson can be interpreted along these lines when they attempt to analyze "belief in adverse circumstances" as similar to self-commands. For Penelhum the additional characteristic is that the self-deceiver sees and understands the evidence against his belief, and holds his belief in spite of this. This is what Penelhum means when he says that self-deception is a conflict-state where the believer partially satisfies the criteria for belief and disbelief.

The problem with these accounts is whether they actually avoid raising the question of motivated belief and the problem of contrary belief. On the surface, Canfield-Gustavson's view that self-deception is like self-command comes perilously close to involving both motivated and contrary beliefs. On the other hand, Penelhum's view that self-deception involves belief when we recognize the evidence against our belief seems very close to requiring the holding of contrary beliefs. We shall now see that these two positions are either no more plausible than Siegler's in distinguishing self-deception from other forms of unwarranted belief, or, if successful, they involve the very problems that they were meant to avoid.

While Canfield and Gustavson, like Siegler, hold that self-deception merely involves unwarranted belief, they attempt an analysis of this. According to them, belief in adverse circumstances can be equated with making oneself believe something

or forget something, and they treat this as similar to self-command. The father who forgets his son's past mischief, and who willfully believes that his son is innocent, although the police are convinced otherwise, is an example of this form of self-deception.

We shall show that it is possible to interpret self-deception as a special case of self-command, i.e. as making oneself believe something or forget something. Under this interpretation, 'self-deception' inherits the non-paradoxical features of 'self-command'.

Canfield and Gustavson, p. 33

This view makes explicit what is implicit in Siegler's account of self-deception as reprimand. The implication is that somehow the source of our deceits is up to us, for, we are, after all, held responsible for them as they are mistakes we "shouldn't" make. On both accounts, the claim to personal responsibility is made while attempting to avoid reference to what goes on within the mind of the deceiver. The interpretation of self-deceit as similar to self-command is the explicit attempt to make this claim good. It fails. For one thing, there is a certain implausibility involved in treating self-deception as similar to something seemingly so different as self-command. Another thing is that the notion of self-command is itself in need of analysis. What sense does it make to say we can command ourselves? What happens if our commands aren't obeyed - was it said too low, or without

conviction? And if they are obeyed - who gave the command and who followed it? The problems here are similar to Siegler's difficulties with accepted reprimands that need repeating.

But more important, the sense in which self-command can be construed as similar to self-deception involves the same problems that this treatment of self-deception was meant to avoid. Thus, the basis of this account is the supposed similarity between "doing something under adverse conditions" and "believing something under adverse conditions", the first being a case of self-command, the second a case of self-deception. The point of the comparison is that just as someone may do something although disinclined, so a person can believe something although disinclined. Now there are two senses to "disinclined to believe something". On the one hand we may be disinclined to believe what is against our will or interest. In this sense the analogy between making ourselves do something and making ourselves believe something holds: we force ourselves to believe what is disagreeable or against our wills. While it is true that there is nothing paradoxical in this, it is also true that this is the opposite of what takes place in self-deception. This description applies very well to the cancer patient who believes he is going to die, refusing to falsely believe, as is his wont, that there is hope. We can non-paradoxically command ourselves to face the facts, but

this surely isn't self-deception. On the other hand, we may command ourselves to believe something when the evidence is against it, as when we try to accept an implausible newspaper story. In this case we try to believe what our evidence, rather than our will, disinclines us to believe. If successful, our belief will be contrary to our evidence, and we will believe what we have reason to disbelieve. Now while this case does fit self-deception as holding unwarranted beliefs, it also implies that our belief is unwarranted because we have evidence that it is an erroneous belief. And so it seems to implicate us in the paradox of contrary belief or willing to believe what is false. This of course was what Canfield-Gustavson's account was attempting to avoid.

Perhaps Canfield-Gustavson only meant "make ourselves believe" to apply to circumstances where we hold an unwarranted belief, but not a belief we have reason to think is unwarranted. But if this is so, then we cannot speak of the believer being in unfavorable circumstances in a way that leads to a comparison with self-command or making ourselves believe something. For a person can only be said to make himself believe or do something if he thinks he is in circumstances that are unfavorable for his belief or action. And in the case of belief this seems to involve having evidence or reason against holding that belief.⁸

In the end, the plausibility of Canfield-Gustavson's account of self-deception comes from smuggling in, under the notion of self-command, the idea of forcing ourselves to hold a belief when we know better. Their account thus implies, rather than avoids, the problem of contrary beliefs. As with Siegler's account, this view of self-deception as unwarranted belief can be plausibly resurrected only if it should prove impossible for self-deceit to be anything more than a case of error or ignorance.

This leaves us with Penelhum's account. Penelhum is aware of the pitfalls in the positions of Siegler and Canfield-Gustavson. ((1), p. 258) He is also aware that the needed addition to self-deception as unwarranted belief is that the believer in some way knows that his belief is unwarranted, and that this addition seems to involve reinstatement of the paradox that was to be avoided. ((1), p. 258) Penelhum's way of dealing with this is both simple and subtle. He accepts the fact that the self-deceiver seems to hold unwarranted beliefs while knowing better, thus distinguishing self-deception from other forms of unwarranted belief. And by interpreting this conflict state of belief and disbelief as the partial satisfaction of the criteria for conflicting belief, he avoids the view that the self-deceiver actually believes contrary things at the same time. He avoids this because the conflict is on the level of criteria or evidence

for belief, not on the level of belief itself. The self-deceiver appears to believe the contrary of what he knows. If he really knew better and didn't seem to believe otherwise we would have hypocrisy, not self-deception; and if he didn't seem to know better but only believed wrongly, we would have only simple ignorance. Penelhum gives errors made about the objects of our pleasure as examples of self-deceit. The teacher who believes his pleasure is due to the fact that his students are about to begin a well earned vacation, when it is his coming vacation that is the source of his pleasure, is deceiving himself. The teacher would seem to know better, and yet does seem to non-hypocritically think that it is his students' coming pleasure that pleases him. Penelhum sums this up as follows:

So I would suggest as severally necessary and jointly sufficient criteria for self-deception over matters other than one's inner states⁹ (i) belief in the face of strong evidence, (ii) the subject's knowledge of the evidence, (iii) the subject's recognition of the import of the evidence. These add up to a conflict state in which there is partial satisfaction of the opposed criteria for belief and for disbelief, with the subjects declarations likely to be against the evidence.

Penelhum (1), p. 259

The first problem is whether an unwarranted belief held while cognizant of the evidence against that belief sufficiently distinguishes self-deception from other forms of unwarranted belief. While this does distinguish self-deception from simple

ignorance, naivete, confusion, etc., there are occurrences other than self-deception where a person holds an unwarranted belief although fully apprized of the evidence that should lead to correct belief. Cases where a man believes a hunch, or holds a belief on faith, or where a scientist believes in his hypothesis although unsubstantiated, and where a man refuses to hold a belief out of stubbornness or habit, all involve unwarranted or omitted beliefs although apprized of the evidence for or against these beliefs.¹⁰

But more important, Penelhum, like Canfield-Gustavson, has not avoided the problem of contrary beliefs. According to Penelhum, in self-deception we partially satisfy criteria for conflicting beliefs. By holding that the conflict is on the level of criteria for belief, the actual status of the beliefs involved supposedly does not arise. This is simply a clever sleight of hand. Penelhum's gambit is to try to keep the question of contrary beliefs to an epistemic level, thus avoiding the question of whether the beliefs are actually held. But what is satisfying criteria for belief if the beliefs are not, at least sometimes, held?

Thus imagine someone who satisfies Penelhum's definition of self-deception. This will entail that the person acts and speaks as if he believed one thing and at the same time knew or believed the opposite. It is possible that such a one holds contrary beliefs, and it is also possible that he is confused

or crazy, a very skillful hypocrite, or only half-believes¹¹ what he does or says. In all these later cases not only are contrary beliefs not involved, neither is self-deceit. Crazy and confused men needn't, and perhaps couldn't, be self-deceived; hypocrites are rarely taken in by their own deceptions; and half-believers, while inconsistent and irrational, need not be deceiving themselves. This leaves us with actually holding contrary beliefs as the only possible occurrence that satisfies Penelhum's conditions and possibly involves self-deception. So satisfying Penelhum's definition either does not involve self-deception, or, if it does seem to be a case of self-deception, it also involves holding contrary beliefs. Penelhum has done no better than Siegler or Canfield-Gustavson.

Conclusion

There is a common pattern to all this. Self-deception can't be merely unwarranted belief when we should know better. Either self-deception is merely a matter of unwarranted belief, and so is indistinguishable from more common forms of unwarranted belief - this is the problem with Siegler's account; or, if an additional characteristic is added to make this distinction, as in adding that self-deception involves "making ourselves believe" or "belief when we appear cognizant of the evidence against our belief", then it seems that we really know better, and so are back with the problem of contrary

belief - this is the problem with the accounts of Canfield-Gustavson and Penelhum.

The problem here is that self-deception as an abnormal unwarranted belief or a belief when we should know better collapses into either simply unwarranted belief or unwarranted belief when we do know better. There is no middle path. This is because there isn't anything between merely unwarranted belief and unusual warranted belief other than contrary belief or warranted belief held because of desire, motives, or intentions. If self-deception is simply unwarranted belief, we have Siegler's position. If self-deception is unwarranted belief when we are aware of the evidence against belief, this seems to imply we hold the correct belief, or at least believe that the unwarranted belief is unwarranted. If self-deception involves unwarranted beliefs we command ourselves to hold, self-deception involves motivated beliefs as well as contrary beliefs. And so if contrary beliefs are impossible, we can no longer hold our unwarranted belief and so can't be deceived. If we continue to hold the original unwarranted belief, and so remain self-deceived, then we hold contrary beliefs, since we at the same time believe it to be unwarranted. This is the see-saw position of Canfield-Gustavson and Penelhum.

This leaves us in the following position. We can accept the non-paradoxical but seemingly inadequate view that self-deception simply involves holding unwarranted belief. Or we

can see whether holding contrary beliefs is in fact paradoxical. If there is no paradox here, then we can accept the view that self-deception involves unwarranted beliefs when we do know better. But even this may not prove sufficient for capturing the full sense of self-deceit. In many cases, self-deceit seems to involve not only unwarranted beliefs when we know better, but also willingly holding these beliefs. This brings in the added problem of whether beliefs can be held because of desire, or whether we can believe something because of motives or intentions. But this is jumping ahead. First we need to see if holding contrary beliefs is paradoxical.

Footnotes to Chapter II

¹ Articles other than Demos, Siegler (1) and (2), Canfield-Gustavson, and Penelhum (1), that take the comparison of self-deception with other deception as basic are: Paluch, Pugmire, Siegler (4), Shea, Canfield-McNally, King-Farlow (1), Hamlyn (1), Mounce, Pears, Gardiner (1).

² For more detailed accounts of some of these distinctions, see: Siegler (3) (On Lying); Mannison; Lindley; Isenberg; Linsky; King-Farlow (1); Eck; Hamlyn.

³ Perhaps examples of contemporary expansions of language that follow along these lines are colloquial expressions like "off the wall", "buggy", or "wacked out". Such terms may have just been used self-referentially to apply to uniquely private experiences, say experiences had through drugs, and involved a metaphorical use of otherwise public terms. Perhaps more and more language users heard or themselves developed these expressions, finding them fitting as applied to their own experience. The result is that we are now in the process of coming to refer in a publicly accepted way to personal private experience through what we originally a private referring expression, albeit a private expression that was itself based on intersubjective notions.

⁴ The reason why the conscious rejection of knowledge through lack of interest, desire, or whatever, can be a case of self-deception is that such a rejection is itself based on belief, and if this belief is mistaken it may be due to self-deception. Members of the ignorance is bliss school of thought, while not necessarily deceived about their resulting ignorance, may have deceived themselves in enrolling at the school in the first place. Perhaps they joined because of a mistaken belief about the value of knowledge, or because they thought knowledge must bring misery, or because they couldn't accept their own ignorance unless they worked at it. Since these mistaken beliefs could be due to self-deception, their ignorance or bliss might be deceptive. And if we choose to run cause together with effect, we can also say that their specific ignorance or mistakes are due to self-deception, for they are the result of their general deceit about knowledge.

On the latter point, it seems better to hold that while a man may deceive himself that ignorance is bliss, the mistaken beliefs that follow from this needn't be considered self-deceits. For although he may be mistaken that ignorance is bliss, if he does believe this then he really doesn't care for knowledge, and so is not deceived in the specific case. It is the self-removal

Footnotes to Chapter II - Continued

from the community of truth seekers that is so unfortunate for those who hold ignorance to be bliss; for unless they are very lucky, their immunity from the possible help and correction of those who know will likely lead to much greater misery than the bliss their ignorance can bring.

We can of course assume that all members of this school are self-deceived, but this is so only if we assume that all men really desire to know, are hereditary members of a rational community, and so on. And this, I suspect, may be as much an erroneous and deceptive belief as the one we are criticizing.

I think that a possible justification for thinking that those who hold ignorance to be bliss are deceiving themselves, is that it is very unlikely that anyone ever graduated from this school. No doubt there never has been a man who successfully kept the purity of his ignorance and was untainted by the desire or at least temptations for knowledge. Since no one ever succeeds in truly believing in ignorance, it can be claimed that expertise, even in particular areas, is really a matter of knowing better. Thus anyone who says that ignorance is bliss really doesn't fully mean it and is deceiving himself about the virtues of ignorance, and so is brought back to the rational fold, where ignorance isn't bliss, and knowledge may help.

5 For more on make-believe see Stern, esp. pp. 28-29, Demos, p. 590; on pretense see Austin (1) esp. p. 215, and Anscombe, esp. pp. 308-310, and p. 298; and for dreaming see Linsky, pp. 165-167.

6 For more on delusion see Demos, p. 590, Freud, 1927, pp. 48-50, Austin, (2), pp. 23-24; for an interesting attempt to work out the metaphor of madness and insight in King Lear see Cavell; and for the extremes of psychiatry on madness and truth see Laing and Szasz.

7 By available evidence I mean the stock of information actually available to the believer, and not the evidence that would or should be present if we are to come to a reasonable or justifiable belief sans phrase. For more on this see the notion of "information basis" in Hempel's "Rational Action", pp. 282-283.

8 This argument is a revised version of one given by Gardiner, 1, pp. 229-231, in his excellent article reviewing the literature on self-deception.

Footnotes to Chapter II - Continued

9 Penelhum's reason for saying that this definition of self-deception doesn't apply to one's own inner states is that he thinks beliefs about such states are not based on evidence or observation. Since such beliefs are not based on evidence they cannot satisfy the conditions for self-deception that he is setting forth. Penelhum later makes the point that errors about our inner states cannot be simple errors for there is nothing to get wrong. But he holds that while someone cannot make a simple error here, since he can satisfy contrary criteria for belief and disbelief, he can nevertheless be said to deceive himself. In such cases, the mistake is explained in terms of some motive, something which is logically out of place in the case of simple errors. He concludes: "Perhaps one cannot be mistaken about one's inner states, but one can have reason to wish them otherwise and to dissemble them." "If I really enjoy the prima donna's dress (sic) but think it is the arias, the continuance of the enjoyment might be conditional on my not admitting its real object to myself." ((1), p. 263) While I think Penelhum's account must in the end be rejected as inadequate as well as timid, it is full of the subtlety and perception that is needed in this area and which is too often missing.

10 There is the possibility of shoring up this inadequacy by adding that hunches, faith, stubbornness, habit, unsubstantiated hypothesis, etc., involve beliefs held with the hope of having future confirming evidence, and that this is not so in self-deception. Also, See Gardiner, 1, pp. 233-236, for some interesting examples of belief held in the face of contrary evidence that are not cases of self-deception.

11 In the case of half-belief a person satisfies some of the conditions for belief but not all of them. He appears to be a "part-time" believer, like someone who truly believes that exercise is good but always avoids exercise, or someone who believes in his coming mortality but doesn't draw up a will, or someone who is superstitious and believes he shouldn't cross a black cat's path, but does when he is in a hurry. A half-believer sincerely believes what he says, but his belief comes and goes with different contexts, where these contexts normally give no reason for changing our belief. See Price (3) especially pp. 156-162, and Braithwaite.

Chapter III: The "Paradox" of Contrary Beliefs

Can a man believe contrary things at the same time? If the answer to this is no, then we can't believe something while knowing better. This view of belief puts a limit on human irrationality for it excludes all but unwitting belief in what is false and makes unintelligible holding beliefs we realize are inconsistent. But self-deception seems to involve these forms of irrationality. Thus, the denial of contrary beliefs leads to difficulty in distinguishing self-deception from other sources of erroneous belief, and the claim that self-deception does involve contrary beliefs leads to "paradoxical" views of self-deception.

In this chapter we consider whether and in what way holding contrary beliefs is paradoxical. We shall see that having contrary beliefs is not logically odd, and the "paradox" of contrary belief only appears if the mind is viewed as inherently rational. This rational view of mind is not required by logic and there are good reasons to think it false. Thus our conclusion will be that men can hold contrary beliefs, and that self-deception can plausibly be construed as involving this sort of irrationality.

Belief and Paradox

There can be little doubt that there is something odd about holding contrary beliefs. Thus Sartre's young woman, who

believes her date's only interest is seduction and who at the same time believes him a caring and concerned companion, truly interested in her self, makes a strange picture indeed. But strange or not, it seems that people do hold contrary beliefs. An indecisive man may believe conflicting things and be unable to choose between them. Often, because of failure to draw the logical consequences from what we believe, or because of failure to realize factual incompatibility between our beliefs, we hold conflicting beliefs. The man who believes that marriage is bliss although all his experiences and other beliefs contradict this, hasn't logically thought things out. The man who believes he will drive to work because of the cold, and who also believes that his car doesn't start when it's cold, has factually incompatible beliefs. Sometimes we even have a logically inconsistent belief. A distraught woman might search for other married bachelors in order to show the world that her John was not simply a married Tom, a bachelor Dick, or a confused Harry. There are other cases too. All men, so the syllogism goes, are mortal, and all men no doubt know this, but few personally accept and believe it. And many are the Sunday worshippers who believe in love and charity to all, and also believe, when they pass the alms box, that God, and men, can only help those who help themselves. Although, it is strange, people do hold contrary beliefs simultaneously.

Contrary belief lies in the shade of paradox. In order to get out of the shade we first need to be clear about the

supposed paradoxes of contrary belief, and then try to shed some light on them. In this section we take the first step of clarifying what is meant by 'paradox', and seeing in what ways contrary beliefs can be paradoxical.

Paradoxes arise whenever two truth related claims conflict and when both claims seem equally proper. What is paradoxical is that while something must be wrong if truth is to be preserved, it also seems that nothing is in fact wrong. Paradoxes are odd or strange because they involve conflicts between commonly accepted, and seemingly acceptable, truth claims. Having your cake and eating it too, when it comes to the truth, is paradox.

Paradox only exists when the preservation of truth is in doubt. This is why paradox is so unpleasant for those with a rational bent of mind. There are a number of different types of paradox.

One type of paradox involves the strange but true. Here the two truth claims only seem to be inconsistent, in fact both claims turn out to be true and consistent. The claim that someone had breakfast in both London and New York at nine the same morning need only be an apparent contradiction; and so too the claim of being forty on one's tenth birthday. In the first case, if one were hungry and liked breakfast and took the 9:15 from London to New York, and forgot about time changes, one can be said to have eaten at both places at nine. And if one was born on February 29, he would be older than his birthdays

show. In these cases the usual connection between time and place, birth and age, do not hold, but this can be accounted for without logical absurdity. Such cases can be called falsely paradoxical.

Unfortunately, there are also cases of true or real paradox. A "true" paradox involves two seemingly true claims that are truly inconsistent, so that one of the claims must be abandoned on pain of logical absurdity. The story of the village barber who shaved all and only those men who did not shave themselves involves a true paradox. It seems that the barber must shave himself and also that he cannot shave himself. If he does shave himself then he doesn't, for if he shaves himself then he doesn't get shaved by himself as the barber; if he doesn't shave himself then he does, for if he doesn't shave himself then he must be shaved by himself as the barber. The absurdity can be avoided by denying that there can be such a barber; this seems strange because our intuition leads us to think that the barber exists. The common form a true paradox takes is that of showing that a certain claim leads to an absurd conclusion, so that something must be wrong with the original claim. Formally put this is the principle of reductio ad absurdum.

One final sense of paradox is possible: "real" paradox. Real paradoxes point to inconsistencies in accepted ways of thinking, and so lead to revision in our theories or conceptual

schemes, rather than rejection of presumed truths as in true paradox. Russell's conundrum of whether the class of all classes that were not members of themselves was a member of itself is a real paradox. Like the barber, if it isn't a member of itself then it is, and if it is a member of itself it isn't. While the two cases are very similar, there is a key difference that distinguishes Russell's real paradox from the true paradox of the barber. In the case of the barber, we are no more than surprised by having to deny the barber's existence on purely logical grounds: his existence one way or the other isn't crucial, and so his late demise for logic's sake is quite acceptable. But in Russell's paradox the presumed basis of general set theory was at stake, for the paradox queried the then fundamental principle of class existence: for any class condition there is a class of things whose members meet these conditions. The choice here was to get rid of general set theory or find a new basic principle for class membership, for as things stood, either set theory was inconsistent or the principle of class existence had not as yet been adequately formulated. The difference then between a real and a true paradox is that in the latter we give up a fact and save our reasoning, in the former we feel forced to excise a part of our reasoning, and this involves relocating, reinterpreting, or abandoning, a whole bevy of facts.¹

We now need to see how contrary beliefs can be paradoxical. As with all paradox, two things are involved: conflict and truth claims. So we must see how beliefs can be true or false, and how contrary beliefs are related, before we can come to belief's paradoxes.

Belief has a dual relation to truth. On the one hand, 'belief' refers to a mental state or characteristic of a person. Thus we speak of person's thoughts and ideas as his beliefs, and we explain a person's actions as the expression or outgrowth of his beliefs. Belief as a state of mind is two-faced. One face is pointed out towards the world and makes up part of a man's mental character. It is in this sense that belief is linked to action, and so to know a man's beliefs is to have a sense of what he will be disposed to do or inclined to say. A man who believes it gauche to be on time is likely to be characteristically late. The other face of belief is more private. It involves the thoughts or ideas that are mulled over, entertained, or just flicker across the screen of consciousness. This is the face of belief that need only show itself to the actual believer, the occurrences in the mind that may or may not flicker across one's lips or face and lighten one's step. When we think of belief's public face, we think in terms of behavior and give a dispositional account of belief; when we think of belief's private face we think in terms of inner goings-on and give an occurrence account of belief. But either way, belief when it refers to a state of

mind is true or false depending on whether the person is actually in that state of mind, whether he has the inclination or awareness that goes with a certain belief. In this sense, to say that someone believes that it is raining, implies that he is either inclined to act in certain ways, say put on his galoshes and take his umbrella, or that it has occurred to him that it might rain, has entertained or is aware of the possibility of rain.

Not only is belief two-faced, it is also two-sided. 'Belief' can also refer to the content of one's state of mind when one believes something. This is the cognitive side of belief. In this sense belief is the appropriate attitude taken towards the truth: we believe what we think true. When belief is seen in this cognitive light it is a mental state that is truth related although it is so in a very broad, uncritical way. Unlike knowledge, what is believed can be held because of social pressures, habit, without supporting evidence or examination, and even when we are not clear as to what it is we believe. What is believed can range from the cold clear truth held with deep and justified conviction after much intellectual labor, to undeclared, indeterminate, questionable notions that were imbibed while we were still in the crib. While there are many explanations for how we come to our beliefs and many reasons for and reasons to believe what we do, whatever we believe must be presumed by us to be true. This is why belief is held to take propositions as its object. It

is in this way that someone's belief (state of mind) is identified with what is believed, say the proposition 'it will rain'. When taken in this sense, the belief, 'it will rain', will be true only when it does in fact rain and not simply when someone takes his galoshes or entertains this notion.

So belief can be true in two senses, as well as having the two-faced aspects of mental states of being either dispositions or occurrences. First, a belief statement is true when the person actually has the belief a statement says or implies he has. This is the truth of belief as a state of mind. Second, a belief statement is true when what is believed, the object of belief, is true. This is the cognitive side of belief. And it is this side that connects belief as a mental state to truth. Talk of beliefs being true can mean that someone is in a belief state about a certain thing or proposition, or it can mean that some proposition believed is true, or it can be a combination of these, that someone's state of mind has captured truth or involves a true proposition.

It is because of these different sides and aspects that belief is such a hard concept to analyze. And even when these different aspects of belief are analyzed there is the problem of fitting everything back together to get the picture of a coherent phenomenon. For instance, how does belief as an occurrence square with belief as a disposition? And how is it that when we refer to a belief as a state of mind, as in 'I believe...', we seem to point past the description of our minds

to the truth of the fact believed? Many of the problems and aspects of self-deception can only be approached and not answered because of a lack of clarity about belief. But I think enough has been said about belief's relation to truth to see how paradox can arise.²

Following 'belief's' dual relation with truth, we come to two ways in which beliefs can be in a truth related conflict. First we can speak of conflict between belief as a state of mind. This conflict leaves aside the question of the truth of what is believed and focuses on the mental state of the believer. Here the conflicting truth claims would be that a person had a belief or was in a state of belief about something and at the same time didn't have the belief and so wasn't in a state of belief about that thing. This does have the air of paradox about it. We need to see if this is a solvable or false paradox or if it is made of sterner stuff.

Belief as a mental attitude or state can be described in a number of different ways. It is not just a yea or nea, a mental pro or con, to some mental content. Since belief in the Judeo-Christian god is often presumed to be something about which we all must take a position, it can be used as illustrating the possible belief states for a given content. A theist is someone who believes that god exists. (BsBsp where p = 'god exists', and Bs = 's believes that...')³ An atheist is someone who believes that god does not exist (BsBs-p). The agnostic does not commit himself to either god's

existence (-Bsp) nor to god's non-existence (-Bs-p), he sits on the fence truly believing in his own absence of committed belief ((Bs(-Bsp · -Bs-p))). There is also the modern day non-believer. Unlike the atheist and agnostic of old, he is neutral or unconcerned, god is not a "live issue" for him. Such a person not only lacks belief in god (Bsp), he also lacks a denied belief (Bs-p), and a withheld belief (-Bsp·-Bs-p). Our modern man is neither theist, atheist, nor agnostic, he doesn't think about god, and so doesn't take a stand with respect to belief about god (-Bs(Bsp v Bs-p v (-Bsp·-Bs-p))).

There are two other forms of non-belief. One is the non-belief of a man who never heard of the Judeo-Christian deity. Such a man, having no concept of this god, has no primary belief about him. Unlike the modern non-believer who fails to exercise his concept of god, the primitive pagan - if there be any left these days - has no belief in god to exercise. The primitive pagan never hearing the tale of the missionary has no belief whatever (-Bsp). Because of his total absence of Judeo-Christian belief the naive pagan was presumptuously thought a religious virgin, and so was held as fertile ground for the missionaries' zeal. There is one final sense to non-belief. This is failure to believe because belief is impossible. A '71 Buick, although a possible object of worship, has no belief in god (-Bsp). But then, unlike the modern or naive pagan, being the crass material thing that it is, it could never have a belief in god.⁴

With all these fine distinctions, the point of conflict between belief as a state of mind should become clear, although perhaps hard to find. In the case of the religious believer, contrary belief states exist between the theist, atheist, and agnostic on the one side, and the pagan on the other. All of the former are in some state of belief about god. The content of their beliefs varies according to their primary belief about god: the theist believes in god, the atheist doesn't, and the agnostic hasn't made up his mind. But all hold beliefs about god. $(Bs(Bsp \vee Bs-p \vee (-Bsp \cdot -Bs-p)))$. On the other side of the religious fence is the pagan. Unlike his neighbors across the way, it is false that he is in a religious frame of mind: he just doesn't believe any of these things. $(-Bs(Bsp \vee Bs-p \vee (-Bsp \vee Bs-p)))$. All of the believers' minds are in some way filled with the belief in god, while the mind of the pagan is a religious blank. One can find theist, atheist, and agnostic, in their various states of belief in animated conversation about god, a conversation which the pagan can only look on from across the gulf of his disbelief. If we try to roll believer and non-believer into one person, we can't. In the case of belief as a state of mind, it is logically impossible for one person to be in contrary belief states. Either you have a belief or you don't $(- (Bsp) \cdot (-Bsp))$.

Now if self-deception involves contrary belief states then self-deception would be truly paradoxical. Only Demos appears willing to say that self-deception involves this logi-

ical contradiction.

Now I come to the far too long postponed problem of self-deception. Believing and disbelieving are pro and con attitudes; they are contraries and therefore it is logically impossible for them to exist at the same time in the same person in the same respect. When B lies to himself he comes to believe what he knows is false, to accept this as the description of a fact is to admit a violation of the law of contradiction.

Demos, p. 591

Demos' view is strange for two reasons. First, it seems clear that self-deception does not involve being in contrary belief states about some object of belief. We don't think of Tom's self-deception about Sally as involving the belief that Sally loves him (Bsp) and also his lacking this belief (-Bsp). But rather that he believes that she loves him (Bsp), while at the same time having the contrary belief that she doesn't love him (Bs-p). The question troubling Tom is, does she or doesn't she (p v -p), not does he believe she does or does he fail to have this belief. (Bsp v -Bsp). While it possibly makes sense to answer the first question by believing both, it makes no logical sense to do this with the second question. The conflict of belief involved in self-deception is on the level of content of belief, not in the state of mind of the believer.

That Demos implies that self-deception involves the logical contradiction of belief and non-belief (Bsp • -Bsp) is doubly strange. The point of Demos' article was to show that self-deception even when interpreted as lying to oneself made sense and did in fact occur. But if Demos really meant what he said about self-deception violating the law of (non) contra-

diction, then this position makes no sense. For if self-deception is logically impossible then it certainly doesn't in fact take place, and even the idea of it would be logically non-sensical. Demos, it seems, has placed himself in the very ludicrous position of saying that self-deception, like square circles, is truly paradoxical, and then squarely attempts to argue for both its sense and the reality of its occurrence. This is quite a strange circle to go in.

Demos was no doubt misled into all this by the confusing ambiguities of 'disbelief'. He confuses lack of belief (-Bsp) with denial of belief (Bs-p). Thus the second line of the quoted passage implying belief as a state of mind becomes believing what one knows is false in the next line. In ordinary use 'disbelief' can mean either believing the contrary of something (Bs-p), the atheist is a disbeliever in this sense; or 'disbelief' can mean that it is false that someone has a certain belief (-Bsp), the position of the pagan non-believer. By confusing the contrary beliefs of theist and atheist (Bsp · Bs-p) with the conflict of belief states between believers (theist, atheist, and agnostic) and pagans (Bsp · -Bsp), Demos mistakenly sees self-deception as involving the logical absurdity of having a belief and not having a belief at the same time, something truly paradoxical.

Seeing the pitfall of Demos' confusion kills two birds with one stone. First, being in conflicting belief states at one time involves a true paradox. Second, self-deception

doesn't appear to involve this paradox. This leaves as our quarry queries over holding beliefs whose contents conflict. This is the second way in which beliefs can involve conflicting truth claims. What sort of paradox, if any, is involved with contrary beliefs? And does self-deception involve holding contrary beliefs?

The "Paradox" of Contrary Beliefs

The question whether self-deception involves contrary beliefs depends on whether contrary beliefs are found to be irremediably paradoxical. If yes, then if self-deception involves contrary beliefs it too will be paradoxical. And this, of course, will not do, for wherever paradox exists supposedly true claims contradict each other. So we must first be clear whether contrary beliefs are paradoxical.

It is interesting that writers on self-deception tend to hedge on whether there is paradox to contrary belief; or, if they hold that there is paradox, they are muddled as to what sort of paradox it is. And, as with Demos, there is often confusion as to what is being considered paradoxical in the first place. This is odd since the question of contrary beliefs is perhaps the single most important and discussed issue in the literature on self-deception. Here are some examples of the confusion.

The concept of self-deception has seemed to some to generate paradox, to make us claim, for example, that a man both does and does not believe the same proposition.

Penelhum, 1, p. 258

It is awkward if not paradoxical to suppose that one could both know something was not the case and yet persist in believing that it was, at least where knowledge is taken to imply awareness of belief.

Pugmire, p. 339

There is something disturbing about attributing contradictory beliefs to a man at the same time.

Siegler, 1, p. 470

The paradox is, then, that although such cases of self-deception exist, i.e. belief in cases where we somehow know better our description of them seems to be self-contradictory... we shall try to show that there remain unsolved two paradoxes of self-deception.

Canfield and McNally, p. 141

From these and other passages it appears unclear whether self-deception involves paradox or merely awkward descriptions or disturbing attributions of belief. It is also sometimes unclear whether self-deception, as paradoxical, is taken to involve conflicting belief states or contrary beliefs. While a large number of writers take self-deception to involve contrary beliefs, it is not clear whether they think this involves logical problems, factual impossibilities, or merely a case of the strange but true. Are contrary beliefs ruled out by the laws of logic, as were conflicting belief states; or are they empirically impossible because of the way our minds work, as would be the case if thinking was always one-tracked; or do contrary beliefs just seem to be odd, like being forty

on one's tenth birthday? There is also a question whether 'contrary belief' applies only to logically inconsistent beliefs (p and $\neg p$) or whether it also applies to empirically inconsistent beliefs, as in holding two logically independent beliefs that can't be true together, given the world as it is.

Since self-deception does not appear to involve conflicting belief states, the important questions for self-deception are: Is holding contrary beliefs at the same time paradoxical? What is the source of paradox? The key here is whether logically or empirically inconsistent beliefs can be held, or whether logically or whether holding such inconsistent beliefs itself involves some form of logical absurdity or is empirically impossible. The literature has spawned a number of reasons for holding that contrary beliefs do involve some form of paradox, and so can't occur.

First there is the problem of how the existence of contrary beliefs in the same person at the same time could ever be established. The problem is that supposed evidence for contrary beliefs can be interpreted as conflicting evidence for what is believed. Instead of evidence that a man has contrary beliefs, it might be thought that what we really have is obscure and confusing evidence, and so cannot be sure about what is believed. (Siegler, 2, p. 38) The ascription of contrary beliefs then is really a euphemism for a believer's confusion. Similarly,

the man who says that he believed something although he knew better can be seen as someone who now realizes that his past belief was faulty and now feels that he should have known this all along. (Siegler, 2, p. 31) Instead of accepting the memory claim of having had contrary beliefs, we can interpret this claim as the present understanding of past mistakes about what should have been believed. And aren't these non-problematic interpretations to be preferred since they avoid the treacherous road of paradox?

This argument has some force in showing that situations which seem to involve contrary beliefs can plausibly be interpreted to avoid this problematic ascription. But this is not an argument against contrary beliefs. Such an argument requires that contrary beliefs are impossibly paradoxical, for only this can force us to make the non-paradoxical interpretation. To hold that we should always settle the question of interpretation in favor of confusing evidence or present realization about belief is to presume that having contrary beliefs is impossible. And this is to beg the question at issue.

Perhaps an argument against contrary beliefs can be found merely on the fact that alternative interpretations for would-be occurrences of contrary beliefs are always possible. Thus let us assume that we can have evidence for contrary beliefs rather than conflicting or confusing evidence for what is believed. Now cases where there is such evidence are not necessarily cases where contrary beliefs are actually held, or, if held,

are held simultaneously. If we catch a person lying it is because we have evidence that he holds contrary beliefs when we know him to have only belief contrary to the one he most wants to show. And we have evidence for contrary beliefs in cases of confusion, vacillation, or indecision over belief, which, while involving contrary beliefs, do not entail that they are held simultaneously. If these interpretations are always possible, there will never be a situation where the only possible account will involve holding contrary beliefs at the same time. But then even with the best of evidence we cannot verify the existence of contrary beliefs. And if we cannot verify or know that a person simultaneously holds contrary beliefs in the best of circumstances for knowing this, then what right do we have to think it does in fact occur?

A possible answer here is that while we could never directly verify the simultaneous holding of contrary beliefs, we could verify this indirectly by having evidence that excludes all alternative interpretations. Cases where a man appears to have contrary beliefs and is not lying, confused, indecisive, etc., are cases where contrary beliefs are held simultaneously.

And now the argument against contrary beliefs continues: it is impossible to exclude all the alternative interpretations, for to exclude one is always to support some other alternative to contrary beliefs. If a man seems to hold contrary beliefs and isn't lying then he is confused, and if he isn't confused then it is likely he just changed his mind, and so on.

This really boils down to assuming that situations where we have evidence for contrary beliefs can always be interpreted as one of a set of exhaustive alternatives, where this set does not include simultaneously holding contrary beliefs. And this, like the first argument, is to beg the question at issue. (Cf. Shea, pp. 11-21)

In general, it seems that epistemic arguments such as the above are inherently inadequate. One needs to do more than point to difficulties in knowing whether something does in fact occur to be able to decisively conclude that it does not occur. Arguments with negative existential import need logical or metaphysical support, and are not plausible when based on only epistemic grounds. Perhaps the argument here is really that evidence that someone believes p is at the same time evidence against his belief in $\text{not-}p$. While this argument is logical in nature it merely begs the question of whether contrary beliefs are possible, only now the question begging is on the level of evidence. The above argument cannot exclude the possibility of holding contrary beliefs simultaneously, but it does point out very real difficulties in knowing that this description is correctly applied in any particular situation.

There is a second argument against contrary beliefs that is based on belief's relation to evidence. The arguments here also come from the fact that there could not be adequate evidence to support contrary beliefs. But this time the problem

arises for the would-be holder of contrary beliefs, rather than someone trying to decide if another holds contrary beliefs. If a man holds contrary beliefs then the evidence will warrant at most one of the beliefs held. At most either Tom's belief that Sally loves him or his belief that she doesn't is justified by the evidence. And if Tom holds his belief without any evidence, or if the evidence is equally divided between the beliefs, then both of Tom's beliefs will be unwarranted given the available evidence. Now the question of why Tom holds this belief arises. It seems that the answer cannot be that it is based on evidence. But then what?

The problem here is to explain why someone holds a belief when short on evidence. There are a number of possible answers. The belief may be held because desire adversely influences judgment and leads us to think that our evidence is better than it really is. This is the view of Russell, Siegler (4), and Butler, among others. The problem now is to explain how belief can be based on desire rather than evidence. (Siegler (2), pp. 34-35) Another answer is that the belief is in some way motivated. We persuade or induce ourselves into our beliefs. This seems to be the position of Freud, Mannheim, and perhaps Butler too. Again the problem here is to explain how belief can be the result of motives and not evidence, and to justify the a priori psychological move that contrary beliefs are always motivated. (Penelhum (1), p. 259) There is a third possibility. Tom persuades or intentionally induces himself to believe that Sally loves him although he knows it isn't so. This is

the most problematic view of all. It is held by Fingarette, Sartre, and perhaps Freud. It is possible to argue that this is really involved in all cases of motivated belief, for if evidence sufficed motives would be unnecessary, thus motives only apply to cases where we should know better. (Canfield-Gustavson, p. 32, and Siegler (2), p. 34) Since the intention itself involves having contrary belief, we now have the problem of motivated belief in addition to the old problem of contrary beliefs.

The attempt to explain how beliefs can be held without evidence leads to problems of how desire and motive are related to belief. Suffice it to say that belief from desire and motivated belief themselves raise large problems, that in the case of some forms of motivated belief one of these problems is whether holding contrary beliefs makes sense. Thus the question of how we can hold contrary beliefs, since at least one of these beliefs isn't based on evidence, involves two separate questions. One, our original question, are contrary beliefs possible? Two, how can beliefs be due to motives or desires? The question of how we can hold beliefs not based on evidence doesn't itself argue for or against the possibility of contrary beliefs, rather it explains how the possibility can be actualized. Thus in some ways it is a totally separate problem, in others it arises only if the problem of contrary beliefs

has already been settled. Either way it neither helps nor hinders us in finding out whether contrary beliefs held simultaneously are possible.

That simultaneously holding contrary beliefs and belief from desire or motive raise two distinct sets of problems often goes unnoticed, and is a source of confusion. Thus Demos speaks as if the only problem with persuading ourselves to believe what we know isn't so is the problem of contrary beliefs and forgets the problems of motivated belief. (p. 588) Siegler argues that since intending to believe doesn't make sense neither does holding contrary beliefs. ((2), pp. 34-35) Because the question of belief from desire and motivated belief is not involved in deciding whether contrary beliefs are possible, it will be left for later when we consider whether mistaken belief from desire or motive is possible and whether it gives a plausible account of self-deception.

I think that behind the mistaken epistemic argument and the misplaced argument against belief held without evidence is the presumption that simultaneously holding contrary beliefs is incoherent. It seems clear that the intuition behind these arguments must be that contrary beliefs are empirically or conceptually impossible, for this is the only plausible way to generate conviction for the negative existential claim that there are no such beliefs. The intuition here is so primitive, or the argument for this view so obvious, that intuition or

argument often go unstated. The presumption is simply that one cannot believe what one at the same time believes isn't so. This insight seems to be so deep that it serves as a basic unexplained assumption about the nature of belief for those who have it. Here are some examples of what I mean:

We could not establish that Jones believed what he knew to be false, or that he knew what he believed was false. For, as these words are normally understood, 'A believes (knows) what he knows (believes) to be false' does not make sense.
Siegler, 2, pp. 31-32

To believe that p is to believe that p is true, so the discovery that two of my beliefs cannot both be true is itself a step on the way to my not holding at least one of them.
Williams, p. 107

One can mistrust one's own senses, but not one's own belief. If there were a verb meaning 'to believe falsely', it would not have any significant first person present indicative.
Wittgenstein, p. 190

A man who becomes aware of a conflict in his beliefs (logically) cannot retain them intact.
Cohen, p. 29

I do not share this insight into the nature of belief. It seems to me that we all know that we hold false beliefs. All but the most arrogant must accept the fact that they hold false beliefs since only an omniscient God is presumed immune from fallible beliefs. Some may even realize that they have particular false beliefs tucked away among their cognitive repertoire, but because they are lazy, indulgent, or too much creatures of habit, they maintain these false beliefs. It is

an unfortunate part of my intellectual baggage that I believe that I can help others avoid their all too human follies by my pointing out to them how I think things really stand. I make my rational gestures countless times each day. They rarely help. And it doesn't much matter if I'm right and my listener knows it. And by now even I know it. But I still gesture, apparently unable to give up a die-hard belief in others' rationality even when I know better. Luckily others forgive me my all too human follies.

Since my intuition tells me that contrary beliefs are possible, I want to show that the intuition into the impossibility of contrary beliefs is mistaken. This is not an easy thing to do. When it comes to basic intuitions perhaps the best or worst one is able to do is share his own and disparage others. I have already given a first attempt at the former. I shall now try to muddy the supposed insight of my opponents by flushing their views out into the open and by criticizing the arguments that are informed by this insight. I can only hope that my view that simultaneous contrary beliefs are possible will become more plausible, and that the certainty that surrounds the view that contrary beliefs are impossible will become weaker.

Although many seem to feel that holding contrary beliefs simultaneously is logically or empirically impossible, few ever come out and say it. We already saw this sort of hedging

on the question of whether self-deception involved contrary beliefs. Some of the above quotes certainly imply that holding a belief we know is false is logically impossible. This is clearly the case with Cohen, and seems to be implied in Wittgenstein's claim that 'to believe falsely' cannot have a first person indicative use and Siegler's claim that knowingly believing what is false "does not make sense". On the other hand, the claim by Williams that the discovery of contrary beliefs tends to force us to give up at least one of the beliefs, seems to be an empirical claim, for he implies that we can hold contrary beliefs even if only for a short while. I shall first try to undercut the idea that holding contrary beliefs is logically impossible, and then try to disparage the more plausible claim that they are empirically impossible or non-existent.

Holding contrary beliefs or believing contrary things is logically possible (i.e. (Bsp . Bs-p)). Demos, as we saw, disagrees. But then he confused contrary beliefs with conflicting belief states. But perhaps Demos' formal contradiction of belief and disbelief ($\neg(\text{Bsp} \cdot \neg\text{Bsp})$) does arise in the case of contrary beliefs. This would be so if we could infer disbelief in the sense of no belief ($\neg\text{Bsp}$) from disbelief in the sense of believing the contrary of something (Bs-p). The ordinary use of 'belief' does seem to allow this move. Thus we tend to go from the disbelief of the atheist, from the fact that he believes that there is no god (Bs-p), to his supposed

disbelief, in the sense of lack of belief, about god (-Bsp). (We cannot logically move the other way, from no belief (-Bsp) to believing the contrary (Bs-p), for then "disbelievers" like agnostics would turn out to be atheists.) But at the heart of this attempted reduction of contrary belief to the formal contradiction of conflicting belief states is the insidious ambiguity of disbelief. For clearly the disbelief of the atheist is not a lack of all belief about god (-Bsp), but rather the lack of a specific object of belief, the belief that god exists (Bs-p). The disbelief involved in holding contrary beliefs does not involve lack of belief as a state of mind (-Bsp), and so the contradiction of conflicting belief states does not arise. (For more on the senses of 'disbelief' see above p. 69 and below footnote 4.)

The point here can also be put as follows. The content of one's belief is not a truth functional component of having a belief, since we can believe false things. Thus we can believe what is either true or false. Conflict between what we believe needn't involve conflict on the level of having beliefs. We can simultaneously believe contrary things (Bsp . Bs-p) without incurring a logical debt to truth with respect to holding those beliefs, although we do fall short of truth in the content of those beliefs. (See Canfield and McNally, p. 142)

The "paradox" of contrary beliefs at last comes down to a question of empirical possibility. Is it possible for the mind to entertain contrary beliefs at the same time? This is

a very difficult question. The usual and very plausible answer is no. I think the answer is yes. And I think that what makes the denial of contrary beliefs so seemingly obvious, so intuitively clear to many, is the unthinking acceptance of a certain rational model of the mind. For want of a better name we can call this the transparency model of mind. The model takes many forms, but its key assumption is that the mind is transparent to itself. We assume that we can "see" what we think about things: our beliefs are pictured as propositional mosaics that are viewed by us on the screen of consciousness. This is a very attractive model. It fits our view of ourselves when all is working well, when we are aware of our beliefs and are rational about them. The rational believer who self-consciously evaluates his beliefs in terms of their corresponding evidence and their coherence with his other beliefs is the paradigm for this model of mind. It's a nice model, with a pretty picture and a self-satisfying paradigm, and it is even sometimes true.

The transparency model is the source of most metaphors for talk about minds, and so is deeply entrenched in our thinking about thinking. It is hard enough to argue against basic assumptions and intuitions, and it is doubly hard when these assumptions are basic to our view of ourselves. Nevertheless, the assumptions and view of mind that the model generates are false. Our minds are not transparent; we do hold contrary

beliefs. The transparency model is not logically forced on us nor factually true of us. But there are difficulties in showing this. Besides pointing out that some of our beliefs must be false, that people, like Sartre's women, appear able to believe in contrary things, that most of us now and then indulge in a belief we know isn't so, and that all minds are sometimes opaque to what seemingly must be obvious to all, little can be done to impugn the transparency model. The reason for this is that we naturally assume that we would notice contrary beliefs and so feel compelled to reject at least one of these beliefs.

The empirical assumption that the mind is transparent and rational leads to a mistaken claim that it is impossible to hold contrary beliefs. Arthur Pap, in Semantics and Necessary Truth, argues that we cannot believe explicitly contradictory sentences because he assumes that the mind must be rational in belief. The argument against the possibility of holding contrary beliefs comes up during a general review of the linguistic theory of necessary truth: the view that analytic statements merely express linguistic rules and so cannot properly be characterized as either true or false. Pap argues for the impossibility of holding contrary beliefs in order to show that contradictions and truisms cannot be believed, and if whatever is true or false can be believed, then analytic statements are neither true nor false. Thus the linguistic theory of truth is at least partially correct. (See pp. 173-181) The general point of

Pap's argument is not important for us. What is important is that Pap's view that belief in explicitly contradictory sentences is impossible is based on empirical considerations about how the mind works and is not, as Pap thinks, a logical claim.

Explicit self-contradictory sentences do not express anything that could possibly be believed: that there are round squares, for example, is not something that could possibly be believed, and the possibility in question is not just psychological. That somebody should believe both (and at the same time) p and not- p is itself a self-contradictory supposition. The frequent claim that people, alas, are capable of holding self-contradictory beliefs notwithstanding, the statement "X believes at t that p and not- p " is itself contradictory. It entails "X believes at t that p and X believes at t that not- p ", and hence "X believes at t that not- p ". But if someone reports that he believes that not- p , we all deduce analytically from the report that he does not believe that p ; no psychological assumption about the working of his mind is needed to justify the inference.

Pap, p. 173

This argument does not show that contrary beliefs are logically impossible. Pap's argument is based on the implicit assumption that the would-be believer is aware that his beliefs conflict and is rational in his beliefs. As we shall see, these assumptions are necessary in order to think that holding contrary beliefs ($B_{sp} \cdot B_{s-p}$) is equivalent to believing and not believing the same thing ($B_{sp} \cdot -B_{sp}$), which is logically impossible. Pap mistakenly builds these assumptions about how the mind works into his view of belief and so sees contrary beliefs as "logically impossible". The argument presumes that contrary beliefs are empirically impossible and then construes this as a logical fact about belief. It is the questionable assumption that the mind is rational in belief that leads to the denial that contrary beliefs are possible.

The original "self-contradictory supposition" is "X believes at t that p and not-p". This involves holding contrary beliefs ($B_{sp} \cdot B_{s-p}$). Pap moves from this to belief in "explicitly self-contradictory sentences". The assumption here is that the two are logically equivalent; they're not. Unlike contrary beliefs, explicit contradiction involves not just believing one thing and at the same time its opposite, but also being aware of all this ($B_s (B_{sp} \cdot B_{s-p})$). The conflict in explicit self-contradiction is in the realization or awareness that what we believe conflicts. Awareness of our beliefs is a psychological fact, and here it is a "fact" based on the assumption of the mind's transparency. Pap assumes that we all "see" what we believe, and so "see" that our beliefs conflict. The mistaken move from contrary beliefs to explicit contradiction is due to the transparency model of mind; it is neither a logical nor psychological fact.

Pap's next move is from awareness of contrary beliefs, and especially belief that not-p ($B_s (B_{s-p})$), to no belief that p ($\neg B_{sp}$). The mistake here involves construing disbelief in the sense of denial of belief or "belief-not" as disbelief in the sense of negation of belief or "no-belief" ($B_{s-p} \equiv \neg B_{sp}$). This is like Demos' confusion of the atheist with the pagan non-believer. (See above p. 68, and below footnote 4.) In Pap's case it seems to be more than a logical confusion. I think Pap is saying that anyone who is aware that he holds contrary beliefs must automatically jettison one of those be-

liefs. As in logic, where it is analytic that contrary beliefs cannot both be true, so in our mind's eye we must "see" that contrary belief is incoherent and reject it. Thus Pap's conclusion that we know analytically that a man can't hold contrary beliefs. All of this is the work of the transparency model of mind; this time it is being used as a metaphor for the mind's rationality. The view that a man cannot hold contrary beliefs is the result of a psychological view of mind that is disguised as a logical fact.

Pap's argument involves the following. First we are asked to imagine a man with contrary beliefs. (Bsp · Bs-p) This is taken as equivalent to a man realizing that his beliefs conflict. (Bs (Bsp · Bs-p) The man "seeing" that both beliefs can't be true immediately give up his beliefs (-Bs (Bsp · Bs-p)). All this is construed as logical, not psychological; our imagined man is held to have a belief he at the same time does not have (Bsp · -Bsp). This is properly seen as logically impossible (-(Bsp · -Bsp)), and so our imagined man is also a logically impossible man, a figment of more than human irrationality who is sent to join the barber who does and doesn't shave himself.

Unfortunately, the man with contrary beliefs cannot be gotten rid of with a logical razor. All the talk of contrary beliefs being logically impossible, far from being analytic, is based on the psychological assumption that the mind is transparent to itself and rational, and this assumption in turn

leads to "logical" facts. This presumption of the mind's pellucid and rational workings clearly is false, for people, alas, can believe what others find obviously self-contradictory: Hobbes did not think that his attempt to square the circle was an impossible task.

Pap has swallowed the transparency model hook, line, and sinker. Because this model of mind is almost second nature to us, it is hard to see, and so is often swallowed and used without notice. But many writers on self-deception are a bit more circumspect. All those who hold that self-deception involves contrary belief take some exception to this model. Thus Demos and Paluch hold that we can have contrary beliefs because we can fail to realize that our beliefs conflict. (P. 593, and p. 344 respectively.) Rorty holds that the self-deceiver thinks that there is some strategy that will enable him to reconcile his contrary beliefs ((1), p. 344). While Hamlyn and Freud hold that beliefs can be unconscious, and Fingarette, King-Farlow, and also Freud, think that we can have unintegrated beliefs. In all these cases, exceptions are made to the view of mind as clear and rational; and strategies are devised to explain how contrary beliefs are possible. In the next two chapters we shall go into some of these views in detail.

The view that holding contrary beliefs simultaneously is impossible finally comes to rest with this enticing but mistaken psychological model of mind.

In the end the "paradox" of contrary beliefs proves to be only a phantom or false paradox. Contrary beliefs are epistemically, empirically, and logically possible. The shadow of paradox was cast by the transparency model of mind. And yet, I suspect, one is apt to feel that contrary beliefs still lie in the shade of paradox.

The Oddity of Contrary Beliefs

Having been unable to detect any formal paradox in the case of contrary beliefs, we are in the awkward position of being unable to explain the feeling of paradox that surrounds them. The woman who believes that her platonic friend is taking more than a philosophical interest in her does, after all, still make a strange picture. The fact that this picture isn't logically, empirically, or epistemically absurd, only makes matters worse. We are left with the residue of a felt or intuited oddity with all the redeeming explanations gone. We need to track down the source of our misleading intuition, the phantom that leaves the impression of false paradox.

After much unsuccessful sleuthing, I believe I may have hit upon a more than ghostly answer. The answer involves the relationship between belief as a state of mind and the content of one's belief. It is the combination of these aspects of belief that leads to the oddity in holding contrary beliefs. With regard to the content of belief, a person can believe what is false as well as what is true. There is no contra-

diction in holding contrary beliefs, although, of course, both beliefs cannot be true. With regard to belief as a state of mind, belief is the attitude appropriate to the truth, and one is in this state whenever he thinks, or is disposed to act as if he thinks, some proposition true. Contradiction involving belief as a state of mind requires being in a belief state and it being false that one is in such a state - the position of the theist, atheist, and agnostic on the one side and the modern or naive pagan on the other. So far so good.

It is when we combine these two aspects of belief that we get into trouble. When belief as a state of mind is combined with belief as referring to the content of what is believed, we have the following in the case of contrary beliefs: believing that contrary beliefs are both true together. The oddity here is that the proper content for belief is what can be considered true, but in the case of contrary belief, the believer has logically sufficient evidence to know that at least one of his beliefs is false, for this is what is meant by contrary beliefs. But then belief as a state of mind is logically inappropriate in the case of contrary beliefs. The oddity or informal paradox is clear: how can a man believe contrary things, for the content of one of his beliefs must be false, but then belief as a state of mind is inappropriate to it.

While the clash or conflict is clear, the offense is not so clear. Is it logical or psychological? If it is the former,

then we have a logical contradiction. If it is the latter, then we have the interesting case that under some conditions a man's beliefs are absurd and incomprehensible, and we need to explain why and how this is so.

I think that the clash between contrary beliefs, while having logical underpinnings, is basically psychological. The logical element is that in holding contrary beliefs simultaneously a person believes what cannot be true, while belief is logically appropriate only with what can be presumed true. Self-deception and holding contrary beliefs involves, so to speak, believing the unbelievable, and so transgressing what seems to be the logical limits of belief. But the conflict in "believing the unbelievable" is really psychological in nature. We can believe what is "logically unbelievable", or what cannot be true, if we do not realize that we hold contrary beliefs. In fact, we only offend against the "logic" of belief when we believe what we are aware cannot be true. It is the realization that one is holding contrary beliefs that is odd, for then one is in the strange position of believing what one has sufficient reason to disbelieve. This explains why it makes sense to say of another that he holds contradictory beliefs, while it sounds strange to say this about oneself. The conflict only arises in the case of explicit awareness of contrary beliefs, or with reiterated beliefs, or with first person, self-referential belief statements.

The oddity of simultaneously holding contrary beliefs finally rests with certain psychological assumptions about belief. It is only when belief is seen as an occurrence where we "see" or "picture" what is believed that contrary beliefs become impossible. When the would-be believer is assumed to picture or be aware of the state of affairs or proposition believed, then holding contrary beliefs is like "seeing" square circles or objects that are green and red all over, and involves assenting to an impossible mental picture. This is indeed odd. But if beliefs are taken as dispositions to behave in certain ways, then holding contrary beliefs is not so odd. On a dispositional account, contrary beliefs will involve confusing and self-frustrating behavior, like the man who does something one moment and undoes it the next, or who is prone to self-created failures. (I think a close observer of human behavior can distinguish when action is due to confusion and when it is due to ambivalence in belief.)

The occurrent view of belief itself seems traceable to the view of man as a rational believer. The model of belief as a mental occurrence involves the notion of rational or justified belief, where belief is presumed to follow upon the thoughtful, conscious evaluation of evidence. This view is part and parcel of a transparency model of mind. On such an account, the rational or occurrent believer "sees" his beliefs for what they are, and since this, in the case of contrary beliefs, involves seeing that what is believed is false, hold-

ing contrary beliefs becomes impossible as it contradicts the "logic" of belief, for belief requires presuming truth.⁵

I believe that the oddity of contrary beliefs only arises when we assume that the would-be believer is explicitly aware of his beliefs and aware that his beliefs are impossibly inconsistent. A man can have contrary beliefs. A man can be aware of two beliefs that happen to be contrary. A man may even be aware that his beliefs are contrary and conflict and continue to hold both if he can somehow manage to think that all will turn out true in the end. It is only when a man is aware that he has contrary beliefs and aware that they are irredeemably inconsistent that the oddity of contrary belief arises. Thus a man can believe what is contradictory; what is problematic is whether someone can believe what he at the same time realizes is contradictory. And I think a case can even be made for this rather outlandish form of irrationality.

The view that all cases of contrary belief involve believing what we realize is impossibly contradictory only arises if we assume a transparency model of mind and construe believing as a form of mental picturing. If one "sees" that his beliefs conflict, then one "sees" that one belief cannot be true and so cannot properly believe both. If believing is seeing then one cannot believe what cannot be pictured, and so one cannot believe contradictions.⁶ Thus it is really the assumption that belief is a transparent mental occurrence which, when combined with the conceptual fact that belief is

only possible where truth can be presumed, leads to the oddity of contrary beliefs. Both the logical and psychological combine to yield the "paradox" of contrary beliefs.

What practical limit does all this place on self-deception? When must we realize that we have contrary beliefs and so find it impossible to hold both beliefs? It might be thought that there is a clear answer to this: we cannot believe paradigm or overt cases of contradiction, cases like "the grass is green and yet not green", or disbelieve truisms like "all AB are A". (This position and these examples are in Cooper, p. 56, and Pap, p. 178, respectively.) But even in the case of explicit or obvious contradictions it is possible to fail to realize that the beliefs involved conflict and so hold both beliefs. Realizing that our beliefs conflict is a psychological fact, and there is no guarantee that we will see even the most obvious conflicts between our beliefs and so find them unbelievable.

The psychologically obvious and the logically impossible tend to converge in the case of overt contradictions, and this is probably the basis of talk about logical absurdity in the psychological idiom of the unbelievable or inconceivable. But while the logical and psychological are closely intertwined, they can still be distinguished. What is seen as an obvious contradiction to some may go unnoticed by others. To hold that we all notice obvious or overt contradictions, that explicit contradictions are the same for all, is either unhelpfully analytic or empirical and suspect.

I suspect that we all agree by now that the view that contrary beliefs are odd, that we cannot believe what is overtly contradictory, is in fact false and based on a mistaken assumption about the mind as transparent and rational. Contradictions are not obvious and explicit to all. Some men believe what others see as hopelessly unbelievable. Our minds are not always keen, clear, or transparent. Even the clearest minds are sometimes opaque; we do not always know what we think or why we think it. And even when we are thinking clearly and so "see" what we believe, we can believe contradictory things. This is because even transparent minds can fail to be rational; we can dissociate our beliefs or fail to integrate what we believe. Minds needn't be clear and men needn't be rational.

Dostoevsky's underground man says that no matter how obvious a belief is, if we do not want to "see" it we can always avoid doing so by going mad.⁷ There is real truth to this. Instead of going mad on a grand scale, we might merely forget, fail to notice, leave our beliefs unsynthesized, become momentarily confused, refuse to avow our beliefs, keep some beliefs confused and indistinct, and so fail to realize either what we believe or that our beliefs conflict. And even when we are aware that our beliefs are contradictory, truth and reason need not prevail, for we can "rationalize" our contradictory beliefs to ourselves, and so allow ourselves our impossible beliefs by warping reason to make the impossible appear true. I suspect that some of the most creative and

ingenious things the human mind has ever conceived involved contorting reason to justify the irrational.

At any rate, any answer to what must be seen as an explicit contradiction, and so be in fact unbelievable as well as logically "unbelievable", involves some psychological assumption about the nature of the mind, and this will always be a questionable assumption about the mind's transparency and rationality. There is no limit to what we can be deceived about, no contradiction that must be seen as explicit and so impossible to believe. Psychology leaves room for contrary beliefs.

The desire to link the psychological with the logical is understandable. There is, after all is said, a close connection. If men as a matter of fact were always irrational then we would have no notion of logic nor any truth preserving discourse. But be this as it may, it does not give us the right to draw any a priori link between the psychologically unbelievable and the logically unbelievable. The only conclusion that appears allowable is the very broad one that if we were completely mad then we wouldn't have science, logic, or any other form of truthful talk. But this doesn't, of course, preclude that we are all sometimes mad. And I'm not even sure that what we suppose to be truthful talk irrevocably proves that we logically cannot all be madmen all the time.

To return to our beginnings: in the end the "paradox" of contrary beliefs turns out to be a false paradox. That men have contrary beliefs appears odd at first sight, but when looked into the oddity vanishes, like the oddity of being sixteen on one's fourth birthday. The source of the strangeness was ultimately found to inhere in the assumption that the mind is transparent and rational. This, taken together with the correct view that we can only believe what we think true, does make holding contrary beliefs a case of explicit contradiction and so impossible to believe. It is also clear what sort of impossibility this new paradox involves. It is a case of true paradox. The two claims, the assumption of the mind's transparency and rationality, along with the definition of belief, leads to the denial of contrary beliefs, and is a case of reductio ad absurdum. The solution to the new paradox requires the rejection of the transparency model of mind and the attendant claim of the mind's rationality. If we replace this view of mind with a more complex account that allows for "unpicturable" and unintegrated beliefs, then paradox is avoided and the existence of contrary beliefs is explained. Thus the "paradoxes" of contrary belief comes down to a mistaken conception of mind, an inadequate theory parading as an obvious fact. The answer to "Why does self-deception as involving contrary beliefs appear odd?" is that we wrongly presume a too rational account of mind.

Moore's Paradox

The combination of both the logical and psychological that is the source of paradox with contrary beliefs is also the source of another so-called paradox, Moore's paradox. Moore's paradox involves asserting something while at the same time denying belief in what we say. "It is raining, but I don't believe it", is the usual example, but any expression of the form "p, but I believe not-p", will do. As with all paradox, the problem is to find the source of the oddity: is it merely strange sounding but true, does it involve conflict between accepted truth claims, or does it show that our conceptual scheme needs to be revised? As with the paradox of contrary beliefs the source of oddity is elusive. But I think the approach we took with contrary beliefs also proves fruitful when applied to Moore's paradox. Moore's paradox can also be seen to be a case of the strange but true, where the strangeness is due to an overly rational view of mind.

Moore thought that the oddity in not believing what we say came from the fact that in general people do believe what their words imply. It was, he thought, an empirical fact that we believe what we assert and so asserting p had come to imply belief that p. Since the conflict is between what is asserted and what is implied but not itself asserted, no contradiction is involved, and the unusual utterance may, in its fashion, be true. (See Moore, "Russell's Theory of Description", pp. 203-204, and "Reply to My Critics", pp. 541-543.) Others,

refining Moore's position, said that assertion without belief under-cut the practical use of language, and so was pragmatically odd or contradictory; the statement was without use but not without sense. Here too no formal contradiction is involved. (See MacIver, "Some Questions About 'Know' and 'Think'", especially pp. 49-50; Bar-Hillel, "Analysis of 'Correct' Language", pp. 333-337; and Langford, "Notions of Analysis in Moore's Philosophy", pp. 332-333.) But it is easy to feel that the oddity in Moore's paradox goes deeper than either an unusual or practically useless form of words, and so a stranger, logical or conceptual, source of oddity is sought. This led to seeing the paradox as violating the rules for "significant" use of language, or as being "necessarily unbelievable by the speaker". (The first position is that of Black, in "Saying and Believing", especially pp. 31-33; the second that of Hintikka, in Knowledge and Belief, pp. 64-78.) On both these views there is no formal contradiction, but there is also no assertion, for a necessary presupposition for coherent use of language or belief statements is missing when we don't believe what we say. (See Landesman, Discourse and Its Presuppositions, pp. 66-68.)

It does seem, as the latter writers point out, that at its heart Moore's paradox involves a conceptual or logical oddity, which although not amounting to a formal contradiction is something more than an unusual or useless use of language. It is the nature of this "quasi"-logical or conceptual oddity

that is so hard to capture. How can something be logically odd and not involve a formal contradiction?

The answer may be that, as with the paradox of contrary beliefs, both logical and psychological elements are combined to bring about a "half-logical" oddity. This would explain both the logical or conceptual "feel" of the paradox and the fact that it doesn't reduce to a formal contradiction. I believe this is what is involved. The strange status of Moore's paradox can be explained once the psychological assumption that generates it is uncovered. The role of such an assumption easily goes unnoticed in the case of Moore's paradox because the paradox is usually dealt with as involving a conflict in the conditions or rules for linguistic use, and not, as in the case of contrary beliefs, a conflict in the conditions of belief. But I think it is clear that at least part of the oddity of "p, but I don't believe it" involves the fact that the person who asserts p is normally presumed to believe that p, and so contrary beliefs are involved.

The role of suppressed psychological assumptions in the case of Moore's paradox can be seen in Hintikka's treatment of the problem. For Hintikka the oddity in "p, but I do not believe that p" comes from the fact that it violates "the general presumption that the speaker believes or at least can believe what he says". (p. 67) Thus "p, but I don't believe it" is held to be "unbelievable" or "doxastically indefensible" for the speaker, where the "unbelievability" is taken to be

of a logical sort without any psychological underpinnings.

This unbelievability of theirs [i.e. statements of the form 'p · Bs-p'] is of a logical character; it can be seen from the very form of words the speaker is using (provided we know how he is referring to himself). Doxastically indefensible statements are therefore self-defeating. By making such a statement the speaker gives his hearers all they need to overthrow his statement (to the extent of showing that he cannot consistently believe what he says). They are for the same reason absurd to make if the speaker is aware of their self-defeating character.

Hintikka, p. 72

So far all is as it should be. The absurdity remarked on is that of someone who consciously holds beliefs he realizes are in conflict. As Hintikka says, in such a case one has logically sufficient evidence to know that his belief is false and so unbelievable for him. This was also the logical part of the oddity of contrary beliefs. The logical oddity is conjoined with the fact that statements must be believable for speaker and hearer, or else they are self-defeating and not coherent statements at all. The psychological element in this is that the speaker and hearer will notice and be aware of the logical absurdity or self-defeating character of what is said.

The psychological element of awareness is suppressed by Hintikka. But it is clear that if the speaker was not aware that his words were self-defeating they could in fact be believed, and then the logical absurdity, although present, would not be noticed, and so logically self-defeating statements may be believed.

Hintikka does take note of the psychological element in his solution to Moore's paradox when he remarks that it is "a (so to speak) logical necessity" that one hears what one says and so the speaker must notice the self-defeating character of his words. What is interesting here is that while apparently aware that he is taking a psychological position on the nature of mind to explain part of the absurdity of first-person belief statements, he treats this assumption as a "simple logical fact". I quote:

This peculiarity of first-person statements (i.e. that one necessarily hears what one says) may not be entirely unrelated to the idea that one's mind is transparent to one-self. We have seen that there is a sense in which one cannot help hearing - rather, knowing without listening - what one is saying. If we are also right in construing thinking as a sort of interior dialogue, then it is tempting to conclude that one must always be able to know what one is thinking, including what one knows and what one believes....How close we have come to the actual doctrine of the mind's automatic self-knowledge is shown by the fact that we have ended up endorsing, with heavy qualifications, one of the metaphors by means of which the doctrine is often illustrated. When Gilbert Ryle explains the doctrine as follows, "To use another simile, mental processes are overheard by the mind whose processes they are, somewhat as a speaker overhears the words he himself is uttering", he is in fact giving us more than a simile. He is referring to one of the simple logical facts of which it seems to me, the whole doctrine is an extension and dramatization.

Hintikka, pp. 94-95

It is clear that the "unbelievability" of "p, but I don't believe it" rests on the assumption that the speaker is aware that he believes p and also not-p and that he is logical or rational, and so finds it impossible to maintain both beliefs. Hintikka sees this. But, like Pap, he swallows the transparency model of mind and the mind's rationality hook, line, and sinker, and so treats the psychological element in Moore's paradox as a "logical fact". Unfortunately men are not as logical as Hintikka likes to believe. Men may assert p and believe in not-p, although this is "logically absurd" and men should know better. The absurdity of Moore's paradox is thus not simply logical, men can believe their own "unbelievable", self-defeating utterances.

Like the paradox of contrary beliefs, Moore's paradox involves both a psychological and a logical element. Hintikka fails to see the importance of the psychological element because he is only interested in the logically consistent use of belief statements, and so defines into 'belief' and 'believer' the conditions of belief that automatically allow us to exclude inconsistent or irrational beliefs. To do this Hintikka has regimented into the conception of belief the condition of "automatic self-knowledge" or a transparency model of mind, making a questionable psychological assumption into a "logic of belief". Hintikka assumes that we are always aware of what we believe, and so are aware that we have inconsistent beliefs, and that we as rational will always find such

beliefs "logically absurd". It is this assumption that makes believing the opposite of what we say seem so strange. It is also this assumption that allows Hintikka to develop a "logic" out of our beliefs.

As with contrary beliefs, Moore's paradox is not so odd when the psychological assumptions underlying the supposed "logical facts" are exposed. Men can believe the contrary of what they say or believe if they are not always rational; men are not always rational nor are their minds always transparent; and so men do believe what their words belie. Moore's paradox also turns out to be a case of the strange but true, a false paradox. And the true source of oddity, as in the paradox of contrary beliefs, lies with a mistaken assumption about the mind as clear and rational.

Conclusion

Moore's paradox and the paradox of contrary beliefs do not show that men cannot hold inconsistent beliefs. So self-deception as involving contrary beliefs is not impossible. What these "paradoxes" do show is that contrary beliefs are limited to beings who are less than perfectly rational. This is, no doubt, an obvious truth, and one that sets little limit on the scope and possibility for self-deception. The supposed paradox that led to the view that self-deception could only involve unwarranted belief, for anything more was logically

impossible, has proved to be a false paradox. The truth behind the paradox is that a truly rational believer would not deceive himself, and this led us to reject the assumption that the mind is transparent and that believers are always rational. The true source of paradox was a mistaken claim about the nature of mind. For good or bad, logic and psychology only now and then cross each other's paths and run together. For each of us the place where they cross may differ, and for some they may coincide longer than for others. We now need to turn to the accounts of self-deceit that attempt to explain the self's irrationality in self-deception.

Footnotes to Chapter III

¹ Philosophers have been interested in all three forms of paradox. What Ryle terms dilemmas are instances of false paradox. Consider the supposed dilemma or false paradox of perception. The conflict here is between the ordinary man's claim that he sees things like chairs, and the physiologist's or psychologist's claim that seeing involves nerves, reaction-times, sodium and chlorine ions, and so on, where one never hears talk of chairs. Ryle holds that these claims and counter-claims about "seeing something" are merely parts of two different views or theories, views and theories that serve different purposes, or involve different categories and concepts. Since they do not share any common ground they do not in fact conflict. (See Ryle, Dilemmas, p. 11 and chapter viii.) Ryle's notion of "category mistake" involves undiscovered, true paradoxes. Thus Ryle uses reductio ad absurdum arguments to try and show that purely mental minds are categorically absurd. If the mind was an immaterial, disembodied, spiritual thing, then, so the argument goes, we could not know anything at all about it; and so rather than absurdly making the mind into a spiritual ghost that mysteriously moved a machine-like body, we should drop this whole view on the nature of mind. It seems to me that some of the things that Ryle thinks are false or true paradoxes may be considered as real paradoxes, and so involve deeper philosophical problems that Ryle supposed. At any rate there are clear philosophical uses of real paradox. These come up in Kuhn's speculation on scientific revolutions, in Wisdom's work on metaphysical speculations, and in Nietzsche's speculations on science's and metaphysics' work. See Kuhn's The Structure of Scientific Revolutions, Wisdom's "Paradox and Discovery", and Nietzsche's Beyond Good and Evil, part 1, "On the Prejudices of Philosophers".

(Throughout the discussion of paradox I have taken both distinctions and examples from Quine, "The Ways of Paradox": what he calls veridical paradox, faritical paradox, and antinomy, I have been calling true, false, and real paradoxes, and I have stressed the relation of paradox to contrary belief more than does Quine. I have also used Quine as my authority on Russell's paradox and its relationship to general set theory.

² For a general account of belief as a state of mind see Price, 1, Series I, Lectures 8 and 9, Series II, Lectures 1-3, and for the treatment of belief as an "uncritical attitude appropriate to the truth", see, Griffiths, 1, especially p. 140, and Rorty, 1, especially pp. 389-391. Also see Ginsberg, Mind and Belief, and O'Connor, "Belief, Disposition and Actions", as attempts to put things together.

Footnotes to Chapter III - Continued

3 I have written the various forms of religious belief as involving self-conscious or reiterated belief for the following reasons. One, by holding that the belief involved is a belief about what one believes the focus is thrown on the fact that a belief is held rather than on the content of the belief. Two, I think this stress is justified. The various statuses of "religious believers" are more references to self-conscious positions that are taken about belief that god (Bsp) than references to the primary belief in god (p). The theist as one who believes he believes in god (BsBsp) is different from the naive native believer with his unreflective belief in god (Bsp). Arguments about religious belief tend to involve belief about god (Bsp) and are different from arguments about god (p). The former, unlike the latter, often leaves god in the lurch and focuses on the prideful reasonings of the self-conscious believer. Third, by using reiterated belief it is possible to distinguish by logical form the differences between the disbeliever who never had a given belief (-Bsp), and the disbeliever who considered and rejected a given belief (Bs Bs-p).

4 It is interesting to note that it is not clear how the '71 Buick and the naive pagan can be adequately distinguished vis modern logical notation. While both of them lack the primary belief in god (-Bsp), it seems that the further mental primitiveness of the Buick cannot be marked. A reason for this is that denial (...-p), failure of predication, and negation (-...p), failure in a statement's truth, are not clearly distinguished. While belief in god is predicable of pagans, it is not predicable of Buicks. Thus pagans can be said to be non-believers in the sense of failing to have a certain belief (i.e. (-Bsp) read as denial that a belief is held rather than negation of a belief statement), where similar denial of belief is not applicable in the case of Buicks. For both Buicks and pagans, all statements about belief in god will be equally false (i.e. -(Bsp) read as negation of all belief statements about (p). If denial of belief (-Bsp) is taken as equivalent to the negation of a belief statement (~(Bsp)), then this distinction becomes blurred. And so the disbelief of the pagan in the sense of failing to hold a certain belief (-Bsp) will become indistinguishable from the fact that no assertion of a belief statement is true for either the pagan or the Buick.

Problems in distinguishing these two senses of non-belief (negation - no belief, and denial - belief no) also blurs the differences between the non-believing pagan and the atheist or agnostic. The atheist and agnostic both have beliefs about god, but this belief includes the atheists' denial of god (BsBs-p) and the agnostics' denial of both belief in god (Bs-Bsp) and his denial of god denied (Bs(-Bs-p)). The modern

Footnotes to Chapter III - Continued

and naive pagan on the other hand, fails to believe in god in the sense that all belief statements attributed to them are false $(\neg(Bsp))$, and $(\neg Bs (Bsp \vee Bs-p \vee (\neg Bsp \vee \neg Bs-p)))$. This is why god is not a live issue to him. If denial of belief is taken as equivalent to negation of belief statements (i.e. $(\neg Bsp)$ s does not believe p $\equiv \neg(Bsp)$ it is not the case that s believes p), then atheist, agnostic, as well as the pagan all become non-believers in the same sense $(\neg Bs \dots)$. But this is an over-simplification. The disbelief of the atheist and agnostic is not due to their negation of all belief statements about god $(\neg Bs \dots)$, but rather to the denial of specific beliefs, the denial that god exists in the case of the atheist $(Bs-p)$, and the denial of belief in god's existence or non-existence by the agnostic $(\neg Bsp \cdot \neg Bs-p)$. Thus it is sometimes remarked that the atheist and agnostic are not so dissimilar from the theist, for both have beliefs about god and both can be just as proselytizing about those beliefs. It is the pagan who is the true disbeliever. This is so either because the pagan doesn't play the religious game since he has not yet been initiated - the naive pagan; or, worse, because he thinks the religious game uninteresting and not worth playing, and so doesn't exercise his belief - the modern pagan.

The difficulty involved in all this is perhaps due to the fact that modern man and modern logic try to avoid existential commitments wherever possible. Traditionally, affirming or denying a predicate of something involved the assumption that an appropriate subject existed. This is the case of belief denied $(Bs-p)$ or in the denial that a person held a belief $(\neg Bsp)$, both entailed some existential commitment, either that a belief was held or that belief could be held. In these cases a "thinking subject" is assumed as real or possible. In negating statements no ontological commitment is made. By creating a logic without existential commitment we avoid the temptation to over-populate the universe because of the logical and grammatical structures of language. But we pay a price for this in failure to distinguish denial and negation. And sometimes the price is costly, as when existential import is an aspect of a word's meaning or is an aspect in or description of things as with mental states that commit us to things.

5 Thus Price, in defending an occurrent account of belief, says that unreasonable beliefs came about in a "non-occurrent" manner. "Unreasonable belief comes into being in a behind the scenes manner. The believer does not notice acquiring them. He acquires them without knowing what is happening to him." (Price, (1), p. 298)

Footnotes to Chapter III - Continued

The occurrent account has problems in explaining how unreasonable or unjustified beliefs are acquired. This is because the model for occurrent belief is that of the rational, self-conscious believer, and we cannot rationally and self-consciously hold unreasonable beliefs. The analog to this is the problem of contrary belief for the transparency model of mind. On both accounts, the acquiring of unreasonable and contrary beliefs must be due to unconscious mental occurrences, or in a "behind the scenes manner". Or, if this too be thought paradoxical, then a dispositional account is required in the case of all unreasonable beliefs.

The dispositional account of belief while able to account for unreasonable beliefs, has problems with reasonable or justified belief. This is because it is unable to explain the conscious assent or rational choice that is often an element in acquiring justified beliefs. In a sense, all beliefs on a dispositional account are unconscious, and so there are difficulties in accounting for the conscious sifting and evaluation of evidence. The problem of explaining why we fail to act on our beliefs in the dispositional account is the mirror image of the occurrent accounts problem of contrary beliefs. The problem here is that there is no mental occurrence such as thought or choice to explain the slack between inclination and action.

I think adequacies and inadequacies in both accounts show that either approach to belief, when taken alone, is likely to prove wanting.

6 The picture theory of mind has had many proponents. It is enshrined in our ordinary talk about thinking by terms like "seeing the point", and in philosophical talk by "clear and distinct ideas". If thinking or understanding involves picturing something, then we could never think or understand a contradiction, for this would involve an impossible picture. Thus Hume said: "'Tis in vain to search for a contradiction in anything that is distinctly conceived by the mind. Did it imply any contradiction 'tis impossible it could ever be conceived." (See Hume, (1), p. 43, and (2), p. 18, p. 27, and p. 157-158.) But people do conceive contradictions, and people can have a clear awareness of their ideas and not be aware that they involve a contradiction, etc.

If we regard all thinking as the production of clearly articulated mental diagrams and blueprints, it follows inevitably that one cannot understand a contradictory sentence. A picture which purported to be a picture

Footnotes to Chapter III- Continued

of the self-contradictory would have itself to be self-contradictory, that is, a self-contradictory description would have to be truly applicable to it. This is logically impossible and thus, if thoughts are pictures, there can be no self-contradictory thoughts. But not all clear thinking is picture-thinking. We have a clear conception of even numbers, but very few people have mental pictures of them. Nor is all thinking necessarily clear thinking. If it were, nobody would ever be muddled.

Cooper, p. 55

7 The underground man says:

Even if man really were nothing but a piano key (i.e. if man's thoughts and actions were no more than soundings or consequences from being struck by determinants), even if this were proved to him by natural science and mathematics, even then he would not become reasonable, but would purposefully do something perverse, out of sheer ingratitude, simply to have his own way.....If you say all this, too, can be calculated and tabulated, chaos, and darkness and curses, so that the mere possibility of calculating it all beforehand would stop it all, and reason would reassert itself - then man would purposefully go mad in order to be rid of reason and have his own way.

Dostoevsky, p. 28

Chapter IV: The Self in Self-Deception - 1

So far we have considered self-deception as unwarranted belief in circumstances where we could and normally would know better. In order to explain such untoward error, we were led to the problem of contrary beliefs. But in all this we have left out what is the most salient feature of self-deception. The picture of the self-deceiver is not simply someone who makes an unexpected mistake in belief, but someone who foists his mistakes on himself. In self-deception our mistakes are of our own making: our selves are the source of our deceits. In this chapter and the next we consider the role of self in self-deceit.

There is perhaps nothing with which we are better acquainted and know so little about as the self. The result is the countless views on the nature of our elusive selves. The multiplicity of positions on self-deception reflects this. What follows is an attempt to set out some views of the self, and then set out some of the positions on self-deception. In doing this we will be fleshing out the bare epistemic bones of self-deception as unwarranted belief. The result is that self-deception begins to take on some of its real life appearances.

Desire, Motive, Intention and Self-Deception

It is strange to think of ourselves as being willfully ignorant or mistaken, or as holding distorted beliefs for the sake of our selves. And yet the self-deceiver seems to do just this. The self-deceiver appears to be someone who believes what he wants, whose thinking follows the twists and turns of desire and interest rather than reason and evidence. In self-deceit our beliefs, at least in part, are reflections of our selves, and not a clear reflection of the way of the world. In fact, if men were only interested in knowledge of the way things are, self-deception would be very strange indeed.

That we are often the source of our own ignorance, that our beliefs appear tailor-made to suit ourselves and not the facts, cannot be denied. We all know people who are deaf to what they do not want to hear. We have all seen how others are able to turn their self-made failures into the unforeseeable whims of an unkind fate, of how growing illness can be seen as the sign of returning health, and how the hard times of yesterday become the warm remembrances of the good old days. The social graces, such as tact, are often conventionalized ways of helping others with their deceits, as when we tell someone that they look good because they look terrible. And how often have we seen others lead themselves by way of the subtlest, most earnest use of sweet reason, to the most outlandish conclusions, as when a moral reprobate by way of wonderful reason is able to fit his indiscretions comfortably

among his moral scruples.

It is a bit more difficult to see ourselves in a similar light, to see ourselves as masters of our own deceits. But there can be little doubt that we all manage our deceits quite well. We all find it easy to see the truth in the unbelievable compliments and outlandish flatteries that others only too rarely bestow on us. And we have all felt how difficult it is to see and then admit that we were wrong, although our error be clear to all the world, and would be clear to ourselves if it were only made by others.

Examples of such frailty and folly can be multiplied without limit. The one characteristic common to so wide a variety of cases is that the errors come not from the material but from the maker: the errors are, so to speak, self-made. This shows itself in the unusual character of our self-deceits, where intellectually obvious and easy to avoid errors are commonplace occurrences made with great relish. And even when such mistakes are flushed out into the light of day they often remain uncorrected.

An explanation is needed for the unusual and uncanny errors of self-deceit. It is that such errors serve some desire, inner purpose, motive, or intention of the deceiver. Errors that appear unaccountable when seen as purely cognitive failings, begin to make sense when seen as the result of the influence of

our selves on our beliefs. In some cases the deceiver even appears to know his true circumstances and to know the undermining objections to his beliefs, and simply warps these to explain and justify his desired or intended ends. Thus the truly extravagant person is able to avoid realizing that his bank accounts are low, his bills high, and his creditors near, and so prolongs a most enjoyable spree until poverty and creditors overtake him by surprise.

There are many ways in which the self can be worked into accounts of self-deceit. The easiest way is to identify the self with a person's desires, and hold that we deceive ourselves when desire disrupts thinking leading us into error. On this view self-deception is cognitive error due to an emotionally obstreperous self. A somewhat more involved view arises when the self is seen as an expression of the inner forces or drives that move us. Here the self is the source of action, not just the cause of intellectual flutterings or impulsive movements. On this view self-deception is the result of the deepest aspects of our personalities; our deceptions are part of our character or one more form of personalized action. Our selves can also be thought of as the seat of intelligent action, and so linked with more than a moment's desires or a lifetime's motives. This is the thoughtful, intentional self: intellect, desire, and motive all rolled into one. On this view self-deceit takes on the air of having been purposefully planned, willful and wilily, not just a moment's or lifetime's emotional outburst.

The distinction between desire, motive, and intention is notoriously difficult to make. Similarly, beginning with Plato's tri-partite soul, it is very difficult to capture and explain our intuitive view of the self. But I think there is a common role to the variety of selves seen in self-deceit. The basic view is that the "self" is needed to account for the individual nature of our deceptions. The mind, as tradition has it, is the same in all where knowledge is concerned; error, on the other hand, has its source in our subjective side. Given the view of a transparent, objective mind, it is only fitting that we have an opaque, error bound self. And so it was only natural that after a while contemporary accounts of self-deception should turn toward the self and return to more traditional views of self-deception.

Some More History

With the possible exception of Sartre and those who maintained that self-deceit was merely unwarranted belief,¹ all writers on the subject turn to some aspect of our subjective selves in order to explain the self-induced character of our deceptions. The view that self-deceit has its source in our selves is perhaps the only agreed upon "fact" in the literature on self-deception. Thus Gardiner writes: "It is, for example, not clear what could be meant by, or what justification there could be for, speaking of somebody as deceiving himself if it

were at the same time contended that what he was said to deceive himself about was a total matter of indifference to him, in no way related to his wants, fears, hopes, and so forth: could we intelligibly talk about 'disinterested' or 'gratuitous' self-deception?" (p.242)

Shortly after the initial critical reaction to Demos, articles criticizing the non-paradoxical positions of Siegler, Canfield-Gustavson, and Penelhum began to appear. The main thrust of these articles was that self-deception involved motive or desire as well as unwarranted belief. Thus self-deception began to be treated as more than just an epistemic anomaly; desire and motive were fitted to belief in order to account for particular cases of self-deceit. The first result was a number of philosophical analyses of 'self-deception' that included desire in a long list of conditions necessary for occurrences of self-deceit. Among those taking this position were Siegler (4), Shea, Paluch, and, more recently, Rorty (1). The basic tenet to these views is that our desires can distort our thinking and lead us down the path of unwarranted belief.

Next self-deception was construed as akin to motivated belief, and with this turn of events the literature began to expand in an historical direction. Patrick Gardiner offered the position that self-deception might be explained and analyzed as "being mistaken with a motive" (p. 242), and he called attention to the views of the 18th Century religious moralist, Bishop

Butler, to bolster this perspective. At about the same time, D. W. Hamlyn took the position that self-deception can be viewed as putting a troublesome piece of knowledge out of consciousness, and he referred to Freud's notion of "defense mechanisms" as a similar approach to self-deception. (p. 53) The key to these positions is that self-deception is linked to what motivates us, our beliefs may express an unknown debt to our subjective selves.

But the most interest and the most hackles were raised by Herbert Fingarette's little book, Self-Deception. Fingarette took strong exception to the articles of Demos, Siegler, Canfield-Gustavson, and Penelhum. He argued that all these positions mistakenly conceive self-deception as a passive cognitive phenomenon when it is really an intentional act on the part of the self-deceiver. The crucial element on this account is that self-deception is purposeful; it is an individual's way of protecting his self from the internal criticism generated by his own unacceptable engagements. (Chapter 3) Fingarette draws on the work of Kierkegaard, Sartre, and Freud in order to defend his view that we can intentionally ignore what we know about ourselves. Like Kierkegaard and Sartre, Fingarette claims that our beliefs can be intentional like our actions; they are an expression of our inner, existential selves, and so we are directly responsible for them. From Freud comes the view that an individual is composed of many parts, and that what we are conscious of is merely the reflection of

one of those parts. While Fingarette's views have not garnered many followers, they have generated much comment, some critical, some supportive. Some commentators have been: Gardiner, Pears, Hamlyn, Rorty, and de Sousa.

With the increase in contemporary philosophical approaches and the addition of earlier positions on self-deception, the literature became both vast and diverse. In order to bring some coherence to this variety, we need a way to categorize the various approaches to the self and its deceptions. Although somewhat forced, I think the literature can be divided into four overlapping approaches. These are: the ethical-religious, the analytic-epistemic, the sociological, and the psychological. The main proponents of the ethical-religious are Butler and the existentialists: Kierkegaard, Sartre, and Fingarette. Their accounts construe self-deception in essentially moral and religious terms. Recent philosophical approaches generally attempt to analyze self-deception with an eye to epistemic and conceptual questions. The analytic-epistemic approach begins with Russell's criticisms of Freud and is typified by Siegler's various accounts of self-deception. The usually overlooked sociological accounts of self-deception construe self-deceit as due to social determinants. Marx and Mannheim are the key figures here. Last, but most important, is the psychological approach. All writers on self-deception make some assumptions about the self and attempt some sort of psychological account

of self-deceit. (See above p.115 and below footnote 1.)

Freud's theory of defense is far and away the most theoretically developed and observationally grounded account of self-deception, and most of the psychological insights and views of others can be found in Freud. In outlining the role of self in self-deception, we shall limit ourselves to brief accounts of the positions of: Siegler, Russell, Butler, Freud, Marx-Mannheim, and Kierkegaard-Sartre-Fingarette, and attempt to elucidate the particular perspective that characterizes these accounts.

We also need to fit these accounts of self-deception into the main line of argument. This can be done if we work from the simplest view of self and self-deception through to the most complex. On the simplest account, the self is linked with desire and self-deception is a matter of wishful thinking or motivated belief. This is the position taken in the analytic-epistemic accounts of Russell and Siegler, and in one of the moral accounts of Butler. A slightly more complex view pictures the self as the source of motives and the seat of personality; the related views of self-deception go from motivated belief to mistaken beliefs we are motivated to hold although we know better. This account of self-deception is given by Freud and in Butler's second account of self-deceit; the positions of Marx and Mannheim are also closely related to this account. Finally, we have the controversial views of Kierkegaard, Sartre, and Fingarette. Here self-deception is a matter of intentionally holding a belief that "contradicts

our selves", and involves willfully going against our "true" or "knowing" selves.

By arguing against the adequacy of accounts that construe self-deception as simply wishful thinking or motivated belief, and for the plausibility of accounts of self-deception as motivated belief where we know better, we will be furthering our main objective of making philosophical sense of the most controversial accounts of self-deception. In order to do this we have divided the accounts of the role of self in self-deception into those that identify "self" with passive desire, and those that identify "self" with the more active will. When the self is viewed as desirous, self-deception appears to be merely a case of wishful thinking; when the self is viewed as willful, self-deception appears to be more a case of motivated or intentional belief where we knew better.

Because the literature is so disparate, I think it will be useful to follow up on the various ways in which self-deception is approached in the literature. And since the line of argument needs to be diffuse in order to cover the variety of positions, I think it will also help to spend a moment on the view of self as desire before outlining the accounts of self-deception that fall along this line of argument.

The Approaches to Self-Deception

The literature on self-deception is marked by a number of distinct approaches. Thus it isn't always clear if any two writers on self-deception are actually talking about the same phenomenon unless one is commenting on the other. For some self-deception is basically a moral phenomenon, for others it is a psychological fact, and for yet others it is part and parcel of a larger social reality. Philosophers are mainly interested in self-deception because of the odd things moralists, psychologists, sociologists, and other philosophers say about it. Since the philosophic approach seems to me the least informed and least informative, I think it is particularly important to make some general comments on the variety of approaches to self-deception. This will make it easier later when we outline some of the different positions, and it will also enable us to relate self-deception to some traditional categories and some current uses. But above all, the diversity of accounts and approaches to self-deception has for too long been overlooked by philosophers, and some attempt to correct this oversight is necessary.

Traditionally, philosophers have seen self-deception as a moral occurrence. And even moralists have given it scant attention.² It is from this angle that Butler, Kierkegaard, Sartre, and Fingarette approach the self and self-deception. On this view the self is seen in moral terms as basically

impartial, responsible, and good, or in religious terms as true, faithful, and unified. The result is that self-deception is always bad. Self-deception is variously conceived as giving in to base desire, a form of incontinence, a source of evil, being double-minded, impure in heart, and irresponsible to one's true self. The life of those who live in self-deceit is pictured as morally corrupt, personally unsatisfying, internally chaotic, inauthentic, and bereft of God - the self-deceiver reaps his just deserts.

When self-deception is approached in this way the objective is to help us be good or to help us save our souls. Thus Butler is concerned to point out to the unwary that self-deceit is a source of vice, and Kierkegaard's aim is to get us to see that purity of heart only comes to those who will "the Good". Self-deception is still approached through a moral or religious view of self, and it remains a topic for homily and the source of literary appeals to our better selves. Some recent additions to this line of approach are: A. C. Murphy's The Theory of Practical Reason, chapter 6; John Macmurray's Persons in Relation, en passim; and Lionel Trilling's Sincerity and Authenticity, especially chapter 6.

Many traditional and contemporary accounts of self-deception take psychology as their starting point. Freud's views are the most influential, but there are many other psychological approaches that can generate accounts of self-deception. One

possibility is to view self-deception along lines of an approach-avoidance model of conflict. (See N. Miller, "Comments on Theoretical Models", pp. 89-98; and J. Dollard and N. Miller, Personality and Psychotherapy.) Another approach construes self-deception in terms of biofeedback mechanisms. (See G. Murphy, "Experiments in over-coming self-deception"; Fingarette, appendix.) There have also been attempts by experimental psychologists to combine Freudian insight with stimulus-response and learning theory. (See Lee and Herbert, Freud and Psychology, part five; D. R. Miller and G. Swanson, Inner Conflict and Defense.) Psychological approaches to self-deception have also developed out of Sartre's existential psychoanalysis. The best known proponent of this approach, besides Sartre, is R. D. Laing. And at least one psychologist has attempted to clarify and develop a theory of self-deception by importing into his work philosophical analyses of psychological terms and some of the theory that goes with this. (See Schafer, "The Idea of Resistance", pp. 270-275.)

Many traditional philosophical attempts to analyze and explain the self are closely related to the psychological and some make passing reference to phenomena akin to self-deception. These attempts make a very long list and the tale they tell involves a large part of the history of philosophy as well as psychology. (See Brett's History of Psychology for an account of philosopher's psychology and psychology's debt to philosophy.) The patron saint of all this is Socrates. Soc-

rates saw the self (psychē) as something in us that was the source of knowledge, the basis of action, the locus of moral and religious worth, and the key to a successful society. All of the approaches to self-deception can in some way be traced back to the Socratic call to know the self. Thus a large part of philosophical psychology, the models and views of self and human nature that began with Plato and Aristotle, is a footnote to Socrates' self; and a large part of these footnotes involve the problem of akrasia, the ancestor of self-deception, which is also part of the Socratic legacy. The depth psychology of Freud, the social route to self in Marx and Mannheim, the moral psychologies of Butler, Kierkegaard, and Sartre, as well as modern philosophical analyses of psychological terms, are all in debt to Socrates' gift of self and its telltale footnotes.

What is common to all psychological approaches is their explanatory character and their reference to the individual. No matter the particular model of self, what is critical is how the self functions and interacts with external and internal conditions. When the self is approached in this way, the nature of the self and the conditions that lead to self-deceit become essential. Self-deception is the individual consequence of how the self operates and the conditions operating upon it. On this view self-deception may be bad or it may be good or useful; it may be accidental, due to fortuitous circumstance, or it may be

necessary, the outcome of the deepest aspects of our personality; it may result from desire influencing intellect, the symptomatic release of dammed up unconscious motives, or the willful expression of our personal, subjective selves. The important point is that the psychological approach attempts to give an account of how and why of self-deception in terms of the self.

A safe generalization is that all writers on self-deception make some psychological assumptions and at least try to give an explanatory account of self-deceit. In the hands of a great psychologist like Freud, the psychological approach is, perhaps, the most illuminating. In the hands of moral psychologists like Butler, Kierkegaard, and Sartre, it is very suggestive. And in the hands of some contemporary philosophers, like Demos, Hamlyn, and Fingarette, one wonders if philosophical proclivities and limited psychological insights aren't being mistaken for deep conceptual truths - a throw back to the armchair a priori psychology of philosophers who refuse to take a back seat to psychology.

Since the self has always been elusive, perhaps all this diversity cannot be helped and may even be of help. At any rate, all accounts of self-deception that are worth their salt take some psychological stand and involve some model of self and some views of human nature. Attempts to explicate our selves and unravel human nature are of perennial interest. A

recent popular attempt to uncover Freud's and Kierkegaard's views on human nature and relate this to self-deception about death is Ernest Becker's The Denial of Death.

Self-deception can also be approached in a sociological way. Although this approach may sound strange and is often overlooked, it is important. The social approach involves seeing the self in relation to other selves, and seeing that at least part of the self is merely a reflection of our social surroundings. On this approach the self's deception is the result of societal rather than personal forces. The Marxian concepts of ideology and false consciousness first called attention to this form of self-deceit. According to Marx, members of economic classes jointly deceive themselves in thinking that their ideas capture the social reality: in fact such ideologies are one-sided products of a false economic consciousness. Mannheim generalized Marx's view that societal man wears economic blinders into a grand form of perspectivism: all knowledge claims are relative to one's social background.

The common insight to the sociological approach is that a person may be only a persona and that truth may only be a social facade. This insight can be worked in many ways. Society deceives man into false ideologies - Marx; society deceives civilized man into thinking that there is objective truth - Mannheim; society requires denial of self and so our civilized discontents - Freud's social comment. The basic claim is that

our selves and our works are only social masks. Erving Goffman in The Presentation of Self in Everyday Life repeatedly shows just how much our everyday lives are part of a larger social drama. Life is a stage for man with his many masks. Oscar Wilde and Luigi Pirandello knew this: a man is only a character in search of an author, or an ignorant jack who thinks it important to be earnest. With Nietzsche the metaphor comes full circle. Nietzsche was also struck by the masks men wear, and it served to stage his most dramatic point: the "truth", as well as the "self", was only a lie that enabled men to survive. In holding that the "truth" was only useful fiction, the "self" only our present mask, Nietzsche reduced everything into the current stage of the ever evolving drama to survive. On this view the greatest self-deception is the assumption that there really is something behind the mask, or that we can see through our masks to a "universal truth". I think this is profoundly true. It may even be the one truth that no one can accept and so sets the stage for all of our other self-deceptions.

A social approach to the self and its deception can also be developed along existential lines. According to Sartre, one form of bad-faith or self-deception involves our attempt to constitute our selves in terms of how others see us. We deceive our selves by taking on the "look" of others, we become a glint in their eye. We can also deceive ourselves by

trying to become our social role, like the cafe waiter who in trying to be a cafe waiter becomes a cardboard image of a waiter. In a similar way, R. D. Laing develops his view that schizophrenia is the destruction of self that results from the impossible demands of family politics or the insane demands of society at large. The result is that the "sane" and "successful" person is found to be a divided self, only a shadow of a former, pre-social self: the "normal" man lives a life of socially required self-deception. This existential social perspective seems to be gaining ground in contemporary sociology. (See Berger, Invitation to Sociology, pp. 143-150.)

This brings us finally to contemporary philosophical approach to the self and its deceptions. This is the way of analysis and epistemology. What most catches the contemporary philosophers' eye is the epistemic problems and conceptual muddle that abound in talk of self and self-deception. There is no question that a good deal of remedial philosophical work is needed in the moral, psychological, and sociological approaches to self-deceit. What is a "true" self? Are the concepts of 'unconscious intentions' and 'willfull ignorance' coherent? Can we know or intelligibly say that society is insane, that the truth is a successful lie or that knowledge is relative? Much that is said about self-deception might not wash in philosophy's analytic bath.

And yet I think that the contemporary philosopher's approach is the least interesting. While there are problems galore in accounts of self-deception, there is also self-deception galore. Most philosophers pick up the pen because of what they find odd in what others say. Siegler, Canfield-Gustavson, and Penelhum criticize Demos; Fingarette criticizes Siegler, Canfield-Gustavson, Penelhum and Demos; Russell criticizes Freud; Hamlyn defends Freud and is criticized by Mounce, and so on and so. Self-deception gets lost in this critical shuffle.

When contemporary philosophers, as philosophers, not moralists, psychologists, or sociologists, attempt to give an account of self-deception, they hazard little and gain less. When these philosophers consider self-deception as anything more than the epistemic bare bones of unwarranted belief, they rarely get further than the emaciated view of self-deception as unwarranted belief due to desire. Philosophical problems hold them back from giving a fuller account. Self-deception is rendered philosophically unproblematic and epistemically acceptable when cut down to the procrustean bed of wishful thinking. While this approach requires only a minimum of questionable ethical, psychological, or sociological assumptions, the results are meagre: self-deception is simply a sophisticated form of cognitive misdemeanor. The baby has been dissolved by the epistemic bath.

While the "non-analytic" accounts of self-deception are often overly speculative and problematic, I think dealing with the speculative extravagances and problems worth the trouble because of the possible gain in the understanding of self deception.

As a result we have limited ourselves to only two epistemic-analytic accounts of the self in self-deception, those of Siegler and Russell. The place for the philosopher in the literature on self-deception seems to me to be that of elucidating the moral, psychological, and sociological perspective on self-deceit, and trying to unravel the philosophical problems that arise on these approaches. When contemporary philosophers try to do more than this, they tend to usurp the role of psychologist, forget the place of society, set out their own moral and religious positions, and they do all this under the guise of searching out paradox in others - a very deceptive pursuit.

What follows is an attempt to set out and explicate what are the most interesting views on self-deception. After presenting the various positions on self-deception, and arguing for the plausibility of the most philosophically controversial, I will try to deflate some of the problems that contemporary philosophers think inhere in the idea that the self is the source of its own deceptions. It is the problems, not the positions, that need to be thrown out by a philosophical bath. But first a moment's introduction to the self as desire.

The Self as Desire

Identifying the self in self-deception with desire is attractive for at least two reasons.

First, this view fits neatly with the traditional dichotomy of the mind as rational and the body as the source of irrational impulse. This common model divides man into a transparent intellect that is similar in all of us, and our individual and muddling bodily desire. This is the model of man given in Plato's Phaedo. The model is attractive because it explains deception as desire's muddling of judgment: error results when subjective, non-cognitive desires becloud our otherwise truthful understanding. This is the model often used with the problem of akrasia.³ We act against what we know to be good because desire disrupts our moral outlook so that knowledge fails to guide us to the good.

Second, this model is especially attractive for self-deception because it avoids the problem of contrary beliefs. Thus self-deception is conceived as mistakes we make when our wishes or desires distort our thinking. The deception comes from holding beliefs that desire favors instead of beliefs that the clear light of reason shows to be true. To connect the self with desire is to inherit a ready-made account of the self's deceit.

Unfortunately, the model of desirous self and rational intellect is overly simple, and the account of self-deception as wishful thinking is inadequate. Desire is not as simple

and passive as this model would have us think. Desire itself requires either the evaluation of what is desired or some understanding of the desired object and the circumstances surrounding desire's realization. In other words, there is a cognitive element to desire, and so the dichotomy of desire versus reason is a questionable model.⁴ But more important, self-deception is more than merely a case of wishful thinking. Avoiding contrary beliefs leads to an inadequate account of self-deception.

Siegler, Russell, and Butler view the self in terms of desire. The result is epistemically and conceptually non-problematic accounts of self-deception. Although this is a recommendation that is of some importance, these accounts raise the deeper problem of failing to capture the nature of self-deceit. In what follows we will outline the positions of Siegler, Russell, and Butler, and argue that all give insufficient accounts of self-deception. This leaves the more complex accounts of self-deception as motivated belief where we know better, with its attendant view of the self as willful, for the next chapter.

Siegler

Self-deception seems to require something like a belief that *p*, which is caused at least in part by a desire that *p* where there is some aversion to not-*p*'s being true, and where there was at one time at least suspicion that not-*p* is true.

Siegler, 4, pp. 160

Siegler returned to self-deception in a 1968 article, "An Analysis of Self-Deception". Although his main concern is still to give an account of self-deception that avoids the problem of contrary beliefs, this time Siegler attempts to describe the actual conditions of self-deception. His approach is to review possible examples of self-deception and to cull from them the necessary and sufficient condition for self-deceit.

The result: self-deception involves changing what we believe because of desire and fear. This account is supposed to have the merit of avoiding the problem of contrary belief and the concept of unconscious belief. In place of believing as true what we at the same time also believe is false, or in place of believing something while unconsciously knowing better, Siegler claims that all we need is change of belief because of desire and fear. Since fear and desire are the basis of the resulting belief, self-deceit involves holding an unwarranted belief. Since we changed to an unwarranted belief, self-deception also involves believing something while previously knowing better. And since self-deception involves changing from a warranted to an unwarranted belief because of desire and fear, Siegler says that the self-deceiver has good reason to know that his belief is mistaken. Siegler also assumes that distorted inference or misvaluation of evidence is likely to be involved, and that given sufficient time or evidence reason will reassert

itself and replace emotion as the basis of belief. ((4), p. 160)
 Siegler sums all this up in a set of four necessary conditions
 that together define self-deceit:

Though it might be an exaggeration, we could say that he [i.e. the self-deceiver] actually believes that not-p as a result of a desire that not-p and a fear that p. In this case the full thesis would be that A is in self-deception if and only if:

- (1) he knows (or believes) that p.
- (2) he believes that not-p as a result of desire and fear.
- (3) he believes that not-p though he has good reason to believe that p.
- (4) he misconstrues or distorts at the level of evidence and inference.

Siegler, 4, p. 161

Siegler points out that this new account avoids many of the pitfalls of the earlier account that identified self-deception with unwarranted belief. Thus Siegler says that self-deceit can now be distinguished from naive or stupid errors because of features 1 - 4. It differs from simply jumping to a conclusion because of the role of fear and feature 3. Similarly, self-deceit can be distinguished from wishful thinking because of the causal role of fear and because the self-deceiver has good reason to know better. And since the contrary beliefs of (1) and (2) are not held simultaneously, but involve change of belief over time, the problem of contrary belief doesn't arise. ((4), pp. 162-164) This is an improvement over Siegler's earlier account where self-deceit simply involved reprimands given to get someone to face the facts.

Most of the contemporary philosophical attempts to define self-deception fall along similar lines.⁵ The idea behind all of them is simple: self-deception involves desire misleading judgment into deceptive belief: desire muddling mind is the self's deceit. The consensus is that desire is needed to explain why people deceive themselves, and the explanatory role of desire is also part of the defining conditions for self-deceit. With the addition of desire, the inadequacies in the earlier definitions are thought to be corrected. Thus by holding that self-deception involves belief from desire, the problem of contrary belief is avoided, the source of unwarranted belief is explained, and the distinction between self-deceit and other forms of cognitive error is made. Desire is seen as the key to saving the epistemic-analytic approach to self-deception; the inadequacies of former definitions are removed, while epistemic and conceptual problems that supposedly inhere in treating self-deception as anything more than unwarranted belief, are avoided. Unfortunately, this view doesn't survive criticism.

Consider Siegler's version of this approach: self-deception involves changing our belief because desire and fear cause us to distort reason and evidence which leads us to believe what we have good reason not to believe (features 1 - 4). This new account has two old problems: one, it doesn't avoid the problem of contrary beliefs; two, it yields an inadequate account of self-deception.

The new definition seems to entail, not avoid, holding contrary beliefs simultaneously. Siegler says that the self-deceiver believes that p and believes that not-p (features 1 and 2), but these beliefs are separated by the time in which desire and fear lead us from belief that p to belief that not-p. Unless desired belief not-p comes exactly when or after feared belief p goes, contrary beliefs will co-habit the mind during the time our beliefs are changing. And how can Siegler be sure that desired and feared beliefs simultaneously change or never co-exist? This is replacing the bad philosophy that contrary beliefs are paradoxical with equally bad psychology. And Siegler's view that self-deception involves desire and fear seems to logically imply simultaneous contrary beliefs. To hold that belief not-p because of fear (feature 2) implies that we believe that there is something to fear. But to be afraid here seems to involve holding the belief that p. So that causal role of fear in bringing about belief not-p implies that we at the same time believe p and so have contrary beliefs. There is another reason for thinking that Siegler's account requires contrary belief. Feature 3 says that the self-deceiver believes not-p while having good reason to believe p, but if he has good reason then his reason and evidence isn't distorted (contra feature 4), and the implication here is that he either unconsciously knows better or that he does in fact have contrary beliefs.

But even if Siegler's new account can be saved from the contrary belief he wants to avoid, it will still prove inadequate. For one thing, Siegler's definition only applies to people who once believed *p* but then changed their mind because of desire and fear. This excludes the more effective case of self-deception where a man is blinded by his desire and fear so that he never believed that *p* although it was the belief that reason and evidence favored. Also, Siegler's definition does not adequately distinguish self-deception from wishful thinking. If a man gives up his belief because of desire, then he doesn't have good reason to know that his belief is questionable, and we have no reason to think that anything more than wishful thinking is involved. To say that the distinction can be made because fear as well as desire influences belief, simply shows naivete about wishful thinking rather than sophistication about self-deception. Finally, Siegler's account uses the questionable model of mind that has irrational desire disrupting normally warranted belief based on evidence and reason. Thus feature 4, that desire distorts inference and leads us to misconstrue evidence. But does desire need to distort reason in causing belief? Why can't desire and fear cause inspired inference and brilliant observation and lead to ingeniously warranted belief. Desire can change belief for the better. And even when desire leads us wrong, must it take reason along with it in order to foist mistakes on the self?

More criticism is possible but unnecessary. The important point is that Siegler's epistemic-analytic approach is both inadequate and fails to avoid the problem of contrary beliefs. Since the problem of contrary beliefs isn't the source of paradox that Siegler thinks it is, both the impetus for taking this approach and its great failing can be dismissed together: it isn't critical that contrary beliefs be avoided so it isn't terrible that they surreptitiously turn up. What is critical is that this necessary element in self-deception is openly left out of account, and has to sneak in through the back door. The main problem with Siegler's and other recent philosophical accounts is that paradox is seen where there is none with the result that a necessary ingredient is left out: the self-deceiver really does know better but ignores this in his deceit. The contemporary analytic philosopher's view that contrary beliefs are paradoxical plus the epistemologist's desire for simplicity and reason leads to philosophical deceptions about self-deception.

Russell

What may, with some propriety, be called self-deception arises through the operation of desires for beliefs. We desire many things which it is not in our power to achieve: that we should be universally popular and admired, that our work should be the wonder of the age, and that the universe should be so ordered as to bring ultimate happiness to all, though not to our enemies until they have repented and been purified by suffering. Such desires are too large to be achieved through our own

efforts. But it is found that a considerable portion of the satisfaction which these things would bring us if they were realized is to be achieved by the much easier operation of believing that they will be realized. This desire for beliefs, as opposed to desire for the actual facts, is a particular case of secondary desire, and, like all secondary desire its satisfaction does not lead to a complete cessation of the initial discomfort. Nevertheless, desire for beliefs, as opposed to desire for facts, is exceedingly potent both individually and socially. According to the form of belief desired, it is called vanity, optimism, or religion.

Russell, pp. 74-75

Russell deals with self-deception in The Analysis of Mind. His intention is to show that mental phenomena can be analyzed in behavioral terms, and he deals with self-deception as a particular example of how a supposedly unique mental occurrence is best understood in terms of observed behavior.

Russell is critical of those who think that we have immediate knowledge of our desires. Those who think this assume that when such knowledge is absent, self-deception is at work. Self-deceit or ignorance about our desire is supposed to involve concealing from ourselves desires that would otherwise be obvious. (p. 32) Russell thinks this is all wrong. Russell is also critical of Freud's notion of repression, for it similarly assumes that we censor or repudiate from consciousness desires which we find unacceptable in ourselves. (pp. 38-39, and pp. 59-60) For Russell both self-deception and repression mistakenly assume that we actively avoid knowledge of our desires, when in fact our desires are normally unconscious and our lack of knowledge of ourselves is the rule, not the deceptive exception.

Russell's view of self-deception develops from his analysis of desire. According to Russell, a desire is not a mental occurrence that we usually know we have. Rather 'desire' names specific cycles of behavior that are caused by the feeling of discomfort and end when the initial discomfort is alleviated and quiescence is reached. The last state of affairs before quiescence is the "purpose" of the cycle, and this state often lends its name to the whole cycle of "desire". Thus 'sex' is the name of the cycle beginning with discomfort and ending, when lucky, with sex. (pp. 63-67) On this theory desires are naturally unconscious; they become conscious when we have a true belief as to the state of affairs that will end the initial discomfort. (pp. 67-69)

According to Russell, most of what goes by the name of self-deception or unconscious desire is simply ignorance about the purpose of our behavior or ignorance about what would bring cessation of our discomfort. This ignorance may be due to simple inexperience. Civilization limits our behavior, and since we are often unable to act out our discomforts we are unable to find out what would satisfy our initial impulse. (p. 72) What is usually meant by self-deception and unconscious desire turns out to be common ignorance or mistaken belief caused by curtailing the normal cycle of behavior; if we were unrestrained with our desires our trial and error search for satisfaction might eventually lead us to find out what we wanted

by happening upon what satisfied us. Civilization is the cause of ignorance and discontent; and self-deception and unconscious desires are misnomers for the mistaken beliefs of civilized innocence, and do not refer to private foul play with personal mental occurrences.

True self-deception is more complicated than simple ignorance about primary desires. It involves attempting to satisfy our desires by believing that they are satisfied. Russell's account of this is obscure, but I believe it interesting and worth trying to reconstruct. It begins with the notion of secondary desire. Unlike primary desire, which seems to be natural, bodily tendencies that can only be satisfied by things in the real world, secondary desires are caused by beliefs and can, at least partially, be satisfied by belief. Secondary desires arise in the following way. Our body, our imagination, and our world cause us to hold beliefs about the possible objects of our discomfort. These beliefs in turn generate secondary desire. This is because we really do want what we believe will satisfy us, and because our beliefs have a reality of their own and cause us to have real enough secondary desires. It is secondary desires that eventually lead all thoughtful men to think that life is fickle, human wishes vain, and ultimate satisfaction constitutionally problematic. For even when secondary desires are satisfied, discomfort remains until all underlying primary desires are also satisfied.

(pp. 73-74) Tom is a perfect example. Tom thinks his discomfort is due to his desire to be with Sally. Actually he is hungry. At it happens, Tom's false belief that he wants to be with Sally leads to a true secondary desire for Sally's company. Lucky Tom has his way: he sees Sally. He feels better, but not that much better: a nameless yearning remains. If Sally happens to cook for Tom, Tom will soon find that this deep yearning can also be satisfied. Tom's primary desire was hunger for food, not Sally. We can imagine that when Tom's hunger is finally satisfied his budding doubts about his love for Sally, his gnawing dissatisfaction with life's pleasures, or his recurrent philosophical questioning about God's designs and human purpose, will once again subside until his next mistake about the object of his initial discomfort.

True self-deception is a special case of secondary desire. In self-deception we attempt to satisfy our desires by the simple expedient of believing that they are satisfied. This is possible in the case of some secondary desires, since they, unlike primary desires, may take beliefs as their object and so be satisfied through belief. Let's take Tom again as our example. This time Tom really does have a primary desire to be with Sally. Tom's desire leads him to think all sorts of pleasurable thoughts, and these thoughts cause Tom to have a secondary desire for Sally: the secondary desire is that he

will soon be with Sally. Unlike Tom's primary desire, this secondary desire can, at least partially, be satisfied by the simple expedient of believing that he will soon be with Sally, or the the radical expedient of believing that his phantasy has come true and so mistakes his pleasurable thoughts for the real thing. Of course all this won't do in the long run. Unless Tom remains "madly" in love, in which case he replaces reality with phantasy, and self-deceit turns into madness, Tom's discomfort will continue to gnaw at him, since his primary desire remains unsatisfied.

Russell paints a nicer picture of self-deception by using desires of a grand and abstract sort for his examples. Since such desires are by their very nature insusceptible of satisfaction in reality, satisfying our desire through belief may be as successful as we can actually be. Thus desires like vanity, religion, and optimism are apt to be satisfied only by believing them fulfilled. This seems to show that such grand desires are really of secondary sort to begin with. Vanity, religion, optimism are all desires for beliefs that are themselves based on beliefs. That is, the desire that we be well thought of, or that the world was made for us, and that things will soon go our way, are desires that are themselves caused by the belief that these things would bring satisfaction, and these belief induced desires can themselves be satisfied by belief in ourselves, religion, or the world. In this they are all unlike poor Tom's primitive desire for Sally.

Russell appears to arbitrarily limit himself to the ideal case of self-deception where there are no primary desires to leave a residue of discomfort after successful belief. The picture that Russell draws of the self-deceiver is of one who is able to satisfy his desires through belief, never suffering the failures or satisfactions of the real world. On Russell's view self-deception turns out to be similar to Marxian ideologies, or the wish-fulfilling dreams of a healthy Freudian dreamer. After all, if the desires of religion, vanity, or optimism, can be satisfied by belief, then, if realistic, we would be as truly satisfied if we never had these desires to begin with. In the end, Russell's self-deception becomes a purposeless internal mental epiphenomenon that need never touch real desire and true belief. And this is odd since it seems to show that there are non-behavioral mental occurrences, the opposite result of Russell's stated intention in writing The Analysis of Mind.

In the final analysis, Russell's account of self-deception is probably more acceptable to me than to Russell. Nevertheless, it has more failings than successes. Russell is right about desire being naturally unconscious and not the result of self-deception or repression. But with the possible exception of Descartes, it is not clear who holds the view that Russell is criticizing. Certainly not Freud. For Freud all desires or instincts are initially unconscious, some became conscious

through fitting beliefs to affects, and it is only with such conscious or pre-conscious thoughts that repression is possible.⁶ At any rate, Russell's criticism of "mentalism" with respect to desire seems more applicable to his own account of desire with its initial feeling of discomfort. Russell's account also lacks needed mental apparatus to raise his views above their simple behavioristic beginnings, since it is clear that we do not feel all unrealized desires as an undifferentiated form of discomfort.

I think Russell's view of self-deception, or rather wishful thinking, as the complex outcome of desire from belief and belief from desire is substantially correct. Russell's view that the vanity of desire is more a matter of ignorance than desire's fickleness, God's folly, or life's meaninglessness is also a point to be well taken. Often a philosophical frame of mind and an existential awareness of life's emptiness is merely a matter of ignorance about what would truly satisfy us. Tom's doubts about the value of love or the sufficiency of hedonism, as well as the existentialist's view that man is a useless passion or God a conductor to nowhere, are either the result of ignorance about our true desires or grandiosity in belief due to the desire to believe that there is something more. I suspect that Russell's simple answer would be worth a good deal if only we could be less upset about our own lack of satisfaction, and more philosophical and existential about life itself.

Russell's greatest failing is that he stops at the point of identifying self-deception with wishful thinking, and so leaves out of account the most interesting aspect of self-deception. Self-deception is not simply satisfying desire through belief, but involves the desire to hold faulty beliefs about our primary desires so that we can actually realize unacceptable desires by wishfully believing them other than they are. This is the complicating additional step that Freud took and Russell ignored. Self-deception is not just wishfully believing some desire to come true as Russell thinks. It is not simply a happy Tom wishfully believing that his phantasies of Sally will come true. Self-deception involves phantasizing what our desires are because we know that if given in their true form we cannot satisfy them. If Tom were to ignore a true account of his desire in order to avoid a conflict which might prevent him from realizing his desires, then he would really be deceiving himself. An example will help make this clear. Tom somehow realizes that Sally will reject him if he shows the true source of his desire to be with her. In order to avoid this he represses his true desire, and consciously believes that what he wants is sex when it is really Sally's cooking he is after. On this account self-deception involves wishful thinking with the important addition that the wishful believer somehow knows better. But we are jumping ahead to Butler and Freud.

What is most interesting about Russell's account is that he thinks it is a criticism of Freud. Russell must have read Freud's early Russell-like account of "self-deception" or wishful thinking. (See Freud, 1901, chapter VIII, section C, "Wish-Fulfillment", especially pp. 604-606, for a fully developed account of the sort of wishful thinking that Russell calls self-deception.) Freud went further than the view of self-deception as wish fulfillment because he realized it was only half the story. It explained why people think wishfully, but not why people deceived themselves and went contrary to what they knew in order to do this. Freud took a second step and told the later part of the story by concluding that men used wishful thinking in regard to desire because they somehow knew that success would otherwise be barred to them. It was this second step that led to the concept of repression and defense, and which required a full-fledged theory of the self's deceptions.

Russell's criticism of Freud is interesting and odd because Freud agrees with most of Russell's "criticisms" and Freud would accept most of Russell's positive views. But Russell is in even a stranger position. Not only does he criticize Freud although his own views are similar to some of Freud's, he also fails to take Freud's second step even though he seems to see a need for it. Russell appeared to realize that we use wishful thinking

to deceive our conscious selves in order to make possible the satisfaction of otherwise unacceptable desires. Thus Russell says that we tend to see the infliction of pain on those we hate as helpful punishment rather than as satisfying the morally questionable (and, perhaps, consciously unacceptable) desire of wanting to see our enemies hurt. (p. 60) Russell even gives what seems to be a Freudian account of why the existence of unconscious desires are denied; an account which could, with slight "behavioral" modification, be applied to Russell's own failure to develop the notion of repressed desire and so give a fuller account of self-deception.

The belief that a motive is normally conscious makes it easier to be mistaken as to our own motives than as to other people's. When some desire that we should be ashamed of is attributed to us, we notice that we never had it consciously, in the sense of saying to ourselves, "I wish that would happen".....It is commonly argued that people are not to blame for unconscious motives, but only for conscious ones. In order, therefore, to be wholly virtuous it is only necessary to repeat virtuous formulas. We say: "I desire to be kind to my friends, honorable in business, philanthropic towards the poor, public-spirited in politics". So long as we refuse to allow ourselves, even in the watches of the night, to avow any contrary desires, we may be bullies at home, shady in the city, skinflints in paying wages and profiteers in dealing with the public; yet, if only conscious motives are to count in moral valuation, we shall remain model characters. This is an agreeable doctrine, and it is not surprising that men are unwilling to abandon it.

Russell, pp. 31-32

Butler

Men are exceedingly prone to deceive themselves, and judge too favorably in every respect, where themselves and their own interests are concerned. Thus, though we have not the least reason to suspect that such an interested person hath any intention to deceive us, yet we of course make great allowances for his having deceived himself. If this be general, almost universal, it is prodigious that every man can think himself an exception, and that he is free from this self-partiality. The direct contrary is the truth.

Butler, X, 13*

Butler's sermon on "Self-Deceit" is the earliest attempt to give a full account of self-deception. Unlike contemporary approaches such as Russell's and Siegler's, with their eye to analytic-epistemic problems and their attempt to avoid psychological commitments, Butler approaches self-deception as a moral phenomenon, and tries to warn us about it and to explain its genesis so that we can avoid it. Butler's results are very creditable. Although he limited his view to the moral sphere, he saw how deep and distorting our deceits can be. Butler was also an astute moral psychologist, and one of his accounts comes close to construing self-deceit as a case of willful ignorance. Unfortunately Butler was unable to take the final step and see self-deceit as willfully ignoring what we know in order to do what we want. He was unable to take that last step because of a limited view of man. Butler, Bishop of Durham, saw man as basically good, and so evil was merely a matter of indiscretion

* In the case of Butler, all references are to his Fifteen Sermons Preached at the Rolls Chapel; the number of the sermon and paragraph are given as reference.

and self-deception was, ultimately, only a form of moral blindness.

According to Butler, self-deception was a particularly vicious and corrupt source of vice. Ordinary vice results when selfish desire gains the upper hand over reason and conscience. (X, 6) A thief's partiality to money leads him to steal, although his action is contrary to his conscience and involves a risk to himself. Self-deception also involves acting from selfish interests, but with this difference: we are cognizant of our partiality, and our ignorance is the result of mistaken judgment or disruption of the understanding which is due to that partiality. "Whereas, in common and ordinary wickedness, the unreasonableness, this partiality and selfishness, relates only, or chiefly, to the temper and passions, in the characters we are now considering, it reaches to the understanding, and influences the very judgment." (X, 6) Unlike the thief, the self-deceiver, being ignorant of his partiality, thinks his act right and fitting, and so steels himself to the call of reason and the proddings of conscience. In self-deception we are blind to our own vices, and can do the worst of deeds with the best of conscience. This is what bothers Butler and leads him to speak of self-deceit as the most insidious source of vice which in perverting reason and corrupting conscience "undermines the whole principle of good". (X, 16)

It is easy to understand why Butler was so concerned with self-deception. Self-deceit is a most telling vice for a religious moralist. For one thing, self-deceit is the preacher's greatest adversary. The man who nods assent to Sunday's homilies, may do Monday's evil deeds with the cherubic smile left over from Sunday's blessings. Imagine being faced with a hundred smiling, morally upright faces every Sunday, and knowing that these good people are the source of Monday's mischief. I suspect that this alone is sufficient to make any Bishop ponder, if not practice, self-deceit. But what is worse is that self-deceit is the vice that enables "good" men to do what is wrong. By blinding ourselves to our own vices, we escape the moral censure we so willingly heap on others: the evil we so clearly see in others appears a great good in ourselves. And because such "good" men act with a clear conscience, they are deaf to correction, and immune to whatever moderation knowingly doing evil may bring. But worst of all, it is perhaps the vilest of deeds that we are likely to deceive ourselves about; the sort of vicious acts that are so hard to accept in oneself are most likely to be cured through the good graces of self-deception. And I suspect that what particularly bothered Butler was that it is those who are most reviled by evil, "good men", that are most likely to practice self-deceit when they do wrong. Self-deception is something like the good man's moral escape valve.

Being religious may make things even worse. Where the common evil man does wrong, perhaps feels bad, hides, and is forgiven, the "good" evil man does wrong, feels good, acts openly, and tries to forgive others his sins. In normal circumstances bad men feel bad and know that they are the source of evil. Self-deceit implies an internal flaw in human nature, a moral loophole in God's creation. Here men do evil, feel good and stand before God and their fellow men convinced of their own virtue. Self-deceit seems to put the lie to the common and comforting view that virtue is its own reward. Thus self-deception is the one vice that a religious-moralist needs to fear from all sides: it can make preaching useless, allow good men to do the worst deeds, and calls in question man's moral nature and the perfection of God's creation; but worst of all it is the one vice that a religious-moralist is himself most apt to practice, what with his high view of things and the unpleasant questions that self-deceit raises about "good" men and God's creation.

Butler says that there are two different sources that can "prejudice and darken the understanding" in self-deceit: "being under the power of any particular passion or appetite" and "overfondness for ourselves". (X, 6) Although Butler is either unclear or uninterested in making the distinction, these two sources lead to two very different accounts of self-deception.

On the first and simplest account, self-deceit is the result of "being under the power of any particular passion or appetite". Here ignorance of our own partiality comes from not attending

to anything that may conflict with satisfying our desires. Prudence and justice are swept away by desire mindlessly pursuing its object. If a man in the heat of such a temper reflects on his actions, his thought isn't impartial, but is aimed at justifying or palliating to himself what he seeks. Butler's description is interesting:

There are instances of persons having so fixed and steady an eye upon their own interests, whatever they place it in, and the interest of those whom they consider as themselves, as in a manner to regard nothing else; their views are confined to this alone. Now we cannot be acquainted with, or in any propriety of speech be said to know anything, but what we attend to. If therefore they attend only to one side, they really will not, cannot see or know what is to be alleged on the other. Though a man hath the best eyes in the world, he cannot see any way but that which he turns them. Thus these persons, without passing over the least, the most minute thing, which can possibly be urged in favor of themselves, shall overlook entirely the plainest and most obvious thing on the other side. And whilst they are under the power of this temper, thought and consideration upon the matter before them has scarce any tendency to set them right: because they are engaged; and their deliberation concerning an action to be done, or reflection upon it afterwards, is not to see whether it be right, but to find our reasons to justify or palliate it; palliate it, not to others, but to themselves.

Butler, X, 4

Being totally involved in what interests us we do not see that we are partial or expend any thought on the possibility that we are unreasonable and unfair: the "understanding is darkened" and "conscience corrupted". The model here is a simple one: desire overpowers reason and conscience. In the headlong pursuit of desire our understanding can shed little light and conscience must take a back seat and wait for a

quieter, better moment. We have all seen people in such a temper, and may have suffered from it ourselves. Think of the times when desire burned bright and all seemed pale beside whatever kindled our flaming interest. It is hard to find a place for prudence and conscience in such a picture.

But such times and tempers surely don't capture what Butler meant by self-deceit. Butler's description is of someone so absorbed that he is oblivious to what is going on around him or in him, except if it pertains to the object of his immediate interest. Such a one is, as we say, "lost to the world". Thus it sometimes happens that I burn my dinner because I get involved in a novel; or I may fail to notice that the Queen of Sheba has passed by because I am all attentive to my present companion; and by concentrating on defending my present position I may fail to realize that your objections undermine it. It is true that by being so attentive to our interests we may fail to notice that we are wrongly selfish. But it is equally true that by being attentive we may fail to notice some illicit desire, and so fail to do what is wrong. This seems to be the idea behind the Police Athletic League, government sponsored summertime jobs, and public education. Sometimes total involvement is good and sometimes it isn't. But whether such involvement is good or bad will depend on the consequences of being in such a state, and not something about the inherent goodness or badness of being in a certain "temper".

Butler, ever the moralist, may have considered only the adverse consequences of having our attention dominated by a single desire, and so been wrongly led to equate the ignorance of total involvement with the somewhat different ignorance of self-deceit. But, at any rate, it is clear that ignorance of partiality due to attending only to what interests us is not self-deceit. For if self-deception was simply lack of attention due to desire, then burning my dinner as well as not seeing that my desire leads to vice would both involve self-deceit.

Butler's second account of self-deceit is more illuminating. On this account self-deception is due to "over-fondness for ourselves". By "over-fondness" Butler is referring to rationalized desire or self-love which is in conflict with conscience. Here self-deceit involves actions done from desire and seconded by reason, prudence or self-love, but which are against conscience, our morally true and better part. "Over-fondness" is undue regard for our appetitive side at the expense of impartial reason and adjudicating conscience. Since reason sides with desire against conscience, we do wrong, are ignorant that we do wrong, because understanding is darkened by desire, and so deceive ourselves. On this second account, self-deceit involves desire and reason betraying conscience, and this explains Butler's contention that self-deception is essentially immoral, involving both a partial use of reason and an abuse of conscience.

The psychology of this view is also more complex than that of desire overrunning reason and conscience. Here desire is on one side, conscience on the other, and reason in the middle. When reason becomes partial to desire we have overly rationalized desire, immoderate self-love, a betrayed conscience, and self-deceit. When reason sides with conscience we have impartial reason, proper self-love, conscientious action: desire is kept in its place and reason and conscience harmonized with the right amount of appetite. This second picture is of man at his best. The truly good man is one who fully realizes his nature: he realizes those desires that fit with prudent self-love and agree with conscience. This is man the way God and the good bishop meant him to be.⁷

It is with this model of human nature and the second account of self-deceit that Butler shows his keen insight into his fellow man. Toward the end of his sermon Butler speaks of the "falseness of heart" and "internal hypocrisy" involved in self-deceit. When a man in such a temper is exposed to criticism concerning one of his failings, he is apt to respond by wholeheartedly condemning all men with such faults, and somehow find a way to persuade himself that he is not such a man. And yet Butler feels that at bottom there is always at least a suspicion that all is not right.

There are doubtless many instances of the ambitious, the revengeful, the covetous, and those whom with too great indulgence we only call the

men of pleasure, who will not allow themselves to think how guilty they are, who explain and argue away their guilt to themselves: and though they do really impose upon themselves in some measure, yet there are none of them but have, if not proper knowledge, yet at least an implicit suspicion, where the weakness lies, and what part of their behavior they have reason to wish unknown or forgotten for ever.

Butler, X, 11

Unlike the earlier account of self-deceit as total involvement with one's desires, this account implies that the man really does know his circumstances and that he knows too that what he does is wrong. Self-deception allows him to ignore this. By using reason to disarm and avoid the morally unpleasant facts, self-deceit allows him to satisfy the appetites that conscience forbids. This is self-deception in its fullest sense.

Butler's account of such a man is equally astute. He says that the self-deceiver's particular ignorance of his misdeeds is really the result of his general knowledge of them: the man, sensing that something is wrong, knows enough not to look too far into things.

It is much the same as if we should suppose a man to have a general view of some scene, enough to satisfy himself that it was very disagreeable, and then shut his eyes, that he might not have a particular or distinct view of its several deformities. It is as easy to close the eyes of the mind as those of the body: and the former is more frequently done with willfulness, and yet not attended to, than the latter; the actions of the mind being more quick and transient, than those of the senses.

Butler, X, 11

And Butler is equally illuminating when he traces the willful ignorance of self-deceit to its cause in self-love or overfondness for ourselves:

Men find the survey of themselves, their own heart and temper, their own life and behavior, doth not afford them satisfaction: things are not as they should be: therefore they turn away, will not go over particulars, or look deeper, lest they should find more amiss. For who would choose to be put out of humor with himself?

Butler, X, 11

This is brilliant moral psychology. It combines keen psychological insight with a cool understanding of moral man. On this second view, self-deceit involves our willfully ignoring reason and conscience in order to pursue vicious appetites. But Butler has problems of his own in explaining what he sees so well.

One of the main aims of Butler's moral philosophy was to show that egoism was not a threat to morality. He wanted to disarm Hobbesian self-love of its implication that man cannot be moral. Thus Butler argues that egoism, in the sense of self-love, is not inconsistent with morality, for self-love although "confined to the interest of the present world, does in general perfectly coincide with virtue". (III, 8) Butler seemed to enjoy being both iconoclast and moralist in turning the tables on egoism with his remark that: "The thing to be lamented is not that men have so great regard to their own good or interest in the present world, for they have not enough". (Preface, 40) As the moral to all this Butler recommends pursuing a life of

prudent self-interest or self-love, for while seeking happiness is not particularly moral, "it would manifestly prevent numberless follies and vices".

The trouble with all this is that self-love isn't always consistent with conscience and virtue. After all Butler also says that the most vicious source of vice is "immoderate self-love" which, far from perfectly coinciding with virtue and conscience, can "undermine the whole principle of good....and corrupt conscience". One cannot hold both that self-love coincides with virtue and that men have not enough of it, and at the same time hold that self-love can be immoderate and the cause of our greatest vices, without some deceit about the self and the love of self.

Butler has similar difficulties with conscience. On the one hand, reason and conscience are the supreme principles in human nature. Here conscience and reason are man's God given guides; they are the authorities that show us the way to the right and eschews the way to the wrong. (III, 5) But Butler also speaks of reason as being partial and conscience corrupt, in which case they cannot be the infallible and supreme principles of morality, nor God's perfect gift to man. And if conscience and reflection are somehow saved, and really are the supreme, impartial, and incorruptible bastions of morality that they were meant to be, then immoderate self-love is impossible and so too self-deceit. This too seems to involve a case of self-deceit that is being paradoxically denied.

I think the problem here is basic and not just a matter of terminological confusion. Butler thinks that human nature is basically good, that if we rightly follow our nature, if conscience and reason rule, all will be well. (See footnote 7 below.) This is why Butler sees self-deception as avoidable and limits it to moral contexts. Unfortunately, self-deception shows that such sanguine hopes are totally unjustified. Reason and conscience can be corrupt: they can lead us wrong as well as right. Butler is overly optimistic in his assumption that self-love and conscience are consistent, and that conscience is a sufficient basis of morality. Basing a moral theory on the call and proddings of conscience is like trying to stop immoral acts with a whispered 'no'. To trust morality to conscience is either to abandon the fort, to live in a time where values are relatively clear, unchanging and agreed to by all, or to treat moral failings as the indiscretions of otherwise good men. And to think that self-love and conscience are for the most part in agreement seems possible only if men are seen at their Sunday best, and perhaps even then only from the perspective of the pulpit.

Self-deception shows that men are inherently neither good nor bad, that human nature can go either way, and that there are no infallible signs or faculties to keep us on the right track. Butler's basic unstated assumption is that if we truly realize our nature and follow reason and conscience, we will all do

well.⁸ But if nature is basically good and ideally infallible, then self-deceit must in the end be unaccountable, a momentary aberration due to desire or weakness of intellect. Butler seemed to have a suspicion that something was wrong with his view. He was much more sure of the fact that self-deception occurred than that he had accounted for it. Thus Butler says that "there is not anything, relating to men and characters, more surprising and unaccountable, than this partiality to themselves, which is observable in many; as there is nothing of more melancholy reflection, respecting morality, virtue, and religion".

(X, 2) If one is to account for self-deception then the question of the fallibility of reason and the goodness of human nature must arise. Butler was unable to see this given his perspective. Perhaps the good Bishop's singular concern for man's salvation led him to see only the good in man. Perhaps Butler was aware of the darker side of things but dared not harp particulars fearing that he could not then be a good Bishop and lose his happy, self-loving flock. Or, perhaps Bishop Butler had a suspicion that all was not well, but couldn't look into particulars fearing that there was no good or God and so lose his own self-loving salvation.

At any rate, while clearly aware of the faults and follies that we "naturally" perpetrate on ourselves in self-deceit, Butler didn't look into the deeper questions about the goodness of human nature and the fallibility of reason. It is with Freud

and Mannheim that questions about the goodness and rationality of human nature are brought into accounts of self-deception. While Butler did consider self-deception a form of willful ignorance that allowed us to pursue otherwise questionable goals, he did not look further into things. It is interesting that some modern moralists, the existentialists, seem to go in a direction directly opposite to Butler. They, like Butler, seem to think that if we are untrue to what is deepest in ourselves, then our lives will be inauthentic, without value and meaning. But, unlike Butler, they seem to conclude that unless carefully watched men will willingly throw moral responsibility and self away and live happily in the dark, amoral side of life, rather than seek the light and the good with only a moment's self-deceitful exceptions. It is to these positions that we now turn.

Footnotes to Chapter IV

¹ The reason I say that these positions are "possible exceptions" is as follows. As we have seen, Siegler's, Canfield-Gustavson's, and Penelhum's view that self-deceit is merely unwarranted belief is clearly an inadequate position. And this even seems to be realized by these writers. Thus Canfield-Gustavson says that there may be other senses to 'self-deception'. (p. 35) Siegler agrees. (1, p. 475) Since the only plausibility to these minimal accounts of self-deception is that anything more would be paradoxical, the above admissions amount to a denial of the adequacy of self-deceit as unwarranted belief. Siegler even includes desire for belief as part of the truth condition for the application of 'self-deception' in a later article analyzing self-deceit. Similarly, Penelhum, commenting on his earlier view, said that when he did not include motives as a necessary condition for self-deception it wasn't because motives weren't present, but because "I considered the motive for which something is done not a proper part of an analysis of what is done". (2, p. 152) In general, when these and other writers tend to exclude motives, desires, etc., from their analyses of self-deceit it is because they hold that only conceptual conditions are appropriate elements in the analyses of concepts, causal conditions being appropriate only for explanation (e.g. see Rorty, 1, p. 398).

Sartre excepts motives, desires, etc., from self-deceit because he holds that there are no causal factors for our actions. This seems to me so implausible that it is no real objection. Sartre wants to emphasize that we often use causal explanations as a way of disavowing responsibility for our actions. This is no doubt true. To deny this over-used avenue for excuse, Sartre claims that there are no causal determinants for action. This is absurd. (See Being and Nothingness, pp. 39 - 42)

² For the most part, reference to self-deceit in the traditional philosophical literature involves pointing out that we may be mistaken about our motives and intentions. This is usually assumed to be the result of undue love for ourselves. Self-deceit is also used to explain how a man may fail to follow the ethical dictates to which he subscribes. Some examples follow.

Our predominant motive or intention is, indeed, frequently concealed from ourselves when it is mingled and compounded with other motives which the mind, from vanity or self-conceit, is desirous of supposing more prevalent.

Hume, An Inquiry Concerning the Principles of Morals, Appendix ii, p. 299

Footnotes to Chapter IV - Continued

It is in fact absolutely impossible by experience to discern with complete certainty a single case in which the maxim of an action, however much it may conform to duty, rested solely on moral grounds and on the conception of one's duty. It sometimes happens that in the most searching self-examination we can find nothing except the moral ground of duty which could have been powerful enough to move us to this or that good action and to such great sacrifice. But from this we cannot by any means conclude with certainty that a secret impulse of self-love, falsely appearing as the idea of duty, was not actually the true determining cause of the will. For we like to flatter ourselves with a pretended nobler motive, while in fact even the strictest examination can never lead us entirely behind the secret incentives, for, when moral worth is in question, it is not a matter of action which one sees but of their inner principles which one does not see.

Kant, Foundations of the Metaphysics of Morals, section ii, p. 23

There is no moral creed which does not temper the rigidity of its laws by giving certain latitude, under the moral responsibility of the agent, for accommodation to peculiarities of circumstances, and under every creed, at the opening thus made, self-deception and dishonest casuistry get in.

Mill, Utilitarianism, p. 33

If a man does what he says he ought not do do, though perfectly able to resist temptation to do it, then there is something wrong with what he says, as well as with what he does. In the simplest case it is insincerity; he is not saying what he really thinks. In other cases, it is self-deception; he thinks that he thinks that he ought, but he escaped his own notice using 'ought' in an off-color way.

Hare, Freedom and Reason, pp. 82-83

3 While Aristotle's account(s) of akrasia or moral weakness are open to interpretation, the most common view is that he thought it involved ignorance in the particular case

Footnoted to Chapter IV - Continued

because of desire. For Aristotle's view see *Nicomachean Ethics*, 1146b - 1147b 15, for commentary see Hardie and Milo, pp. 81 - 102. Aristotle's account of moral weakness has been reused countless times. For some current cases see Demos, p. 593, and Hare, chapter 5; and for comment on this general approach to *akrasia* see Rorty, 2, pp. 56-57, and Davidson, pp. 99 - 100. It should also be remarked that a large part of the problem of weakness of will comes from seeing knowledge as akin to desire in the first place, and assuming that we want to do what we see is good. See Plato's *Protagoras*, 351b - 358d, and for commentary here see Santas, and Rorty, 2, 50 - 55.

4 This is a traditional criticism to the traditional model of man as emotional vs. man as rational. A modern version of this criticism is made against Ryle's view of the emotions as dispositions. (*Concept of Mind*, chapter IV) The criticisms against Ryle was that this view does not adequately capture the role of social context and values in having and expressing emotions (Bedford, "Emotions"), or that this view does not adequately capture the place of cognition in interpreting a situation as emotional or evaluating our object as desirable. (See Kenny, *Action, Emotion and Will*, chapter IV, and Peters, "motivation, Emotion, and the Conceptual Schemes of Common Sense".)

5 Some examples of definitions similar to Siegler's are:

- X is self-deceived.....if:
- (1) X believes p and p is false.
 - (2) X knows the evidence which counts against the truth of p.
 - (3) X has some motive for discounting the evidence.
 - (4) If the motive were lacking X would see that p is false and its negation true.
 - (5) If the motive were made clear to X he would see that it provided no legitimate grounds for his belief.
 - (6) X is free to discern the character of his motives.

Paluch, p. 276

To say that A is self-deceived in believing P is true is to say that the reason why A believes that P is true is that A wants to believe that P is true. This means that in any case of self-deception, the following conditions must be fulfilled: (i) A wants to believe that P is true; (ii) A holds the indefensible belief that P is true; (iii) A acts on his desire to believe that P is true in doing whatever he does in connection with putative evidence relating to P.

Shea, pp. 90 - 91

Footnotes to Chapter IV- Continued

6 Freud in fact has a very elaborate view of how we became conscious of our desires and instincts, and when and why we became conscious of them. His basic notion is really very similar to Russell's, although, of course, much more developed. Through experience of the world and the symbolic use of words we learn to label certain bodily urgings as desires for certain objects. Exactly which desires can be labeled in this way, and which, if labeled, can become conscious, involves a complex of processes that range from what behavior and conceptual apparatus society allow us to our selectively attending to the manifold of inner urgings or discomforts, where these processes themselves refer to other desires and beliefs. A large part of this involves Freud's theory of cathexis or view of conscious awareness. Freud speculated about this question over the long course of his work, but a fair presentation of Freud's very interesting, difficult, and overlooked views are given in 1901, chapter VII, section e, "The Primary and Secondary Processes - Repression", and in his 1915 articles "Repression", and "The Unconscious".

7 Butler's psychology is an up-dated version of Plato's tri-partite soul. (The Republic, IV, 434d-441c) Like Plato, Butler's model of human nature develops from an analysis of conflicting motives; and also like Plato, when human nature functions well and all parts are in unison, it also functions best and justice rules. Butler can also be thought of as reclaiming Aristotle for the Church of England, and, as we shall shortly see, much of Butler's second account of self-deceit parallels Aristotle's account of akrasia. But either way, Butler's views are interesting.

According to Butler human nature is composed of three distinct but inter-related parts: appetites, self-love and benevolence, and conscience. Appetite or any particular desires are "a direct and simple tendency toward such and such objects, without distinction of the means by which they are obtained". (II, 13) To act from an appetite is to be moved by blind desire, without reflecting whether our action will obtain what is desired or whether it will involve injuries to ourselves or others. Of the various appetites or passions inherent in human nature some tend to be good for us, some indifferent, and some bad; some are good in relation to others, some indifferent, and some are bad for others; and, of course, there are all sorts of combinations of these possibilities. To act from self-love involves reflecting on our appetites and following the rational principle of only realizing these appetites which are good for us and tend to our happiness in the long run (XI, 5) To act from benevolence involves reflecting on our appetites and following the rational principle of only realizing those appetites which are helpful to

Footnotes to Chapter IV- Continued

others, or, at least, not harmful to them. (XI, 11) Thus both self-love and benevolence involve reflection and the use of rational principles to guide us in satisfying our desires. Conscience is the third part of our nature: conscience is the principle of morality that is stamped by God into human nature. It is because of conscience that we intuitively know and approve what is right and know and disapprove what is wrong. (I, 8)

Conscience, as well as self-love and benevolence, is, along with desire, a source of action. As the supreme principle in human nature, conscience adjudicates conflicts between self-love and benevolence, where self-love and benevolence are thought to generally coincide with what conscience dictates. Self-love and benevolence, as sources of action, are themselves superior to the appetites: prudence or self-love, and concern for others or benevolence, are, so to speak, rationalized desires. Appetites came last; they are good, bad, and indifferent; self-love, benevolence, and conscience being the only way to test their worth.

What we actually do is the result of which part of our nature is the most forceful: in the moral man conscience rules; in the rational man it is prudence and benevolence; and in the thief appetite gains the day. In the ideal man all three parts function together, the "lower" parts of human nature being in accord with the "higher" parts. The good man is the true work of God, he realizes his desires (appetite), which are best for him (self-love), and at least not harmful to others (benevolence), and all this goes on under the smiling, watchful eye of a godly conscience.

8 The assumption that human nature is basically good, and that reason and conscience are the source of morality, may explain why Butler failed to distinguish his two accounts of self-deceit. If reason and conscience are the source of morality, then whether we are ignorant because we single-mindedly follow some desire, or whether we close our mind's eye and let desire rule and conscience follow, the immoral results will be the same. Thus while single-mindedly following our interest doesn't entail that reason has become partial and conscience corrupt, it practically has the same effect. For if, as Butler holds, virtuous acts are impartial acts due to reason and conscience, then a consequence of single-mindedly following some desire is that we cannot be virtuous since we do not follow reason or conscience. Since Butler holds that we best realize our nature when acting from reason and conscience, following desire undermines what is best in us. It is conceivable that because of this Butler may have held that following desire is a case of self-deceit. Here self-deceit would involve being motivated by what is worst in us, our deceiving our "true"

Footnotes to Chapter IV - Continued

or better selves. And so impartial reason and corrupt conscience need not actually appear for there to be self-deceit. This, as we shall see, is the view of Kierkegaard; self-deception involves a failure to realize our inner moral or religious worth, and is not a failure to know what is right and good.

Chapter V: The Self in Self-Deception - 2

With this chapter we finish our review of the accounts of self in self-deception. We have saved the best for last. Unlike the over-simplified and timid views self-deception covered in the last chapter, this chapter deals with the most complex and speculative of views. We have finally come to the interesting ideas of Freud, Marx and Mannheim, and Kierkegaard, Sartre, and Fingarette. Since what is recounted here involves very complex notions of self, and raises very large and barely manageable questions, such as whether human nature is good, knowledge possible, truth valuable, and life meaningful, much that is of interest must be left unsaid or only hinted at. Even with countless limitations, what follows is not as brief or as clear as I wished it to be.

The Self as Will

The easiest way to get our bearings with the accounts of this chapter is to outline the general view of the self upon which they rest. The common perspective here is that the self is seen as willful, the source of purposeful and intentional acts. We have already reconnoitered the views of self-deception that developed from seeing the self as desirous. A comparison of the two views will help clear the ground.

Siegler, Russell, and Butler, in his first account of self-deceit, all linked the self with desire. The result is that these accounts share some general characteristics. They all agree that self-deceit is something that happens to us; it is something suffered or undergone. According to this, men do not commit acts of self-deception the way, say, they commit acts of sabotage or tell lies. This is related to a rather simple view of self that all these accounts of self-deception share. The self is divided in two. One part is desire, the other, and better part, is intellect. To be active in the full sense is always a matter of using one's head; to be moved by the passions, such as desire or the emotions, is to find impetus in our baser parts. The dichotomy of desirous self and rational mind is the link to a very satisfying intellectual account of the mind's errors and the body's deceptions: when desire or emotion influence mind the result is a mudding of an otherwise transparent intellect with its clear-headed judgment. The result: self-deception is simply ignorance due to desire's effect on mind. Unlike some myths about mind, this one has a happy, interpersonal ending. Self-deception, as merely a matter of wishful thinking, is curable, needing only the salve of a little veritas, and self-deceit doesn't involve the problems of contrary beliefs. Unfortunately self-deceit has other problems on this account, it doesn't appear able to survive being reduced to wishful thinking.

This led us to the more sophisticated of Butler's two accounts. Here self-deception involved closing the mind's eye to what we don't want to see. The self-deceiver is like the extravagant man who is surprised by both creditors and bankruptcy at the end of every enjoyable spree. Desire, having won reason to its side and by-passed conscience, uses partial reason to justify unconscionable acts, although we know enough to know better. Unlike the earlier view of self-deceit as wishful thinking, this account involves more than a pellucid mind muddled by base desire. Reason now justifies desire, it is an active element in self-deceit, and it is now conscience that suffers the deceits of the unholy alliance of desire and reason. On this view we have a tri-partite version of the self, and a more active view of self-deceit as erroneous belief while in some way knowing better. But while this almost captures the full sense of self-deception, it has an insurmountable problem. For Butler reason and conscience were both good and the ruling elements in man, but this makes "partial reason" and "corrupt conscience" unaccountable. In the end Butler is unable to say what he so clearly sees, that man can do wrong willfully and while knowing better. The two aspects of knowing better and being willing to be wrong, which are the basic ingredients in self-deceit, are both but darkly understood because of Butler's moral view of human nature.

All the views in this chapter pick up where Butler's second account left off. Self-deceit is no longer the beguiling snake of appetite deceiving men about their heavenly selves: self-

deception is now as natural as human nature. Men, for the first time, are seen in their proper dress; without the moralist's and epistemologist's rosy glasses of the good, the true, and the beautiful, men emerge as their naked, truly earthy selves. Self-deceit has at last come to roost in human nature. We ourselves are responsible for our own deceptions. Life is no garden of Eden, but then men are not angels, only men. And men as men are the willful, active, knowing source of their own deceptions: there is a snake in the heart of all of us.

With this conversion in the view of man and the resulting insight into our seamier side, self-deception appears fully fleshed and with a wardrobe to dress all the vagaries of mankind. Self-deceit has finally come to its self, and it emerges as the real life phenomenon that it is. There have been many accounts of self-deception as purposeful and natural. According to Freud we need to deceive ourselves about an unacceptable psychic reality; for Marx and Mannheim thinking is often deceptive in its nature and tends to reflect our personal needs and outlooks; Kierkegaard, Sartre and Fingarette see men as deceiving themselves about their own responsibility for the lives they lead. Self-deceit is at last seen to be alive and well and living in us all. Butler's deceitful extravagant man is seen for what he really is: one, small, unextravagant example of the self-deceptions of man: in self-deceit men willfully suppress what they know in order to do what they want.

The single unifying view to the accounts of this chapter is that self-deceit is natural to the self. All else follows from this. Thus self-deceit is no longer an epistemic anomaly; the result of rambunctious desire, partial reason, or momentarily blinded conscious; nor is it necessarily bad for the self. Self-deception when placed within the context of human nature becomes an integral function of the self, a useful device for its practitioners. In the case of Freud, Mannheim, and Sartre, self-deception is not only natural, it is also necessary; while for Marx, Kierkegaard, and Fingarette, who share a moral point of view, it is avoidable, but only with great effort and the good graces of history, God, or personal identity. But all see self-deceit as a willful act. For Freud self-deceit is necessary for the preservation of self. For Marx and Mannheim it is the natural outcome of intellectual processes. For Kierkegaard, Sartre, and Fingarette it is part of the flight from selfhood and responsibility that we all sometimes take. Self-deception has become a purposeful and expressive act by and for the self; it is no longer passive, unusual, and completely dysfunctional.

Along with the new, natural view of self-deceit goes a new image of the self. Because of the many additions, this image is rather ungainly and complex. In fact, I do not think that there is either a common or straight-forward model of the self on any of the above views. Freud can be thought to have updated the Platonic tri-partite model of self into id, ego, and super-ego, but I believe this is illusory. While this model does construe

the self as the dynamic outcome of three interrelated functional parts, this model can also be interpreted in descriptive terms, where relation to consciousness is basic; economic terms, where relation to instinctual energy is basic; or developmental terms, where the causal relations to the past is basic. A large part of the fruitfulness and value of psychoanalytic theory comes from its versatility and breadth, but this also leaves the self as somewhat elusive, the common complaint of psychologists and philosophers. Nor is there a clear image of the self in Marx or Mannheim. As with Freud's psychoanalysis, in Marx's ideology and Mannheim's false consciousness, the self remains ever elusive, although all three share the view that the self, whatever it finally turns out to be, is causally accountable.

This is not so for the existentialists. They stress the complexity of self from an internal point of view. Here self is treated in an holistic fashion and is identified with the "individual" or the "personal" in us. Choice, will, self-knowledge, values, rather than instincts or social background, are the key to the self. Subjective experience, existential reality, and personal identity are all basic to the self and its deceptions. Sartre is so emphatic in holding that this involves a unique approach that he denies the possibility of a causal or scientific account of self.

Perhaps the only generalization true of this new image of self is that it has an immense internal complexity. From this follows the common rejection of the transparency model of the

last few chapters. All of the present writers concur in the view that it involves complex and only vaguely understood processes, and that the self may always be opaque to itself and elusive to others.

This leads to another important feature of the present views. Not only is the mind opaque, and self-deceit inherent in man, but men are also irrational. And the irrationality here cuts deep; it is not a matter of unwitting error or the mind's momentary aberrations; the repudiation of reason is part of the very fiber of man. To see the self as grounded on instincts, primitive, and willful, is to see man as having a heart of darkness. This is the opposite of the transparency model of mind. Knowledge, as well as self-deceit, is due to an act of will. This is a far cry from the view that self-deceit is simply unwarranted belief or cognitive error. Now all beliefs, all intellectual functions, are seen to develop out of an emotional or willful core. Man is ever an animal; rationality is society's, culture's, or evolution's veneer: deep down there is still an irrational self.

All of the positions in this chapter see the self and its surroundings as ultimately the result of blind, heedless forces; in a man this is the will, in the world at large it is the contingent laws of nature and history. At its very heart this view is antithetic to mind, reason, and an essential goodness or meaning to things. In Freud, the rational processes are secondary, and are themselves hard won accretions to primary,

primitive processes. In Marx and Mannheim, ideas and beliefs are always the end-product of dumb material processes, superstructures built on a non-cognitive economic or social base. For the existentialists men are not essentially thinking things, but more a matter of feeling and willing. Mind is inherently misleading: it objectifies our subjective selves, and gives a distorted picture of men and their existential reality; it also structures the absurd, unintelligible world of things, replacing amorphous reality with a neatly categorized world of facts. In the end, as in the beginning, the world is irreducibly contingent; there is nothing essential, or sensible, or systematic about people and things: we make of the world and ourselves what we will of it.

There is a common intellectual heritage behind all this. It begins with Kant, and the assumption that the mind plays an active role in making experience intelligible. And it has so far come to a head with Nietzsche's view that our concepts and understanding put a lie to experience in making it manageable, and Sartre's view that man and the world are in themselves absurd and without meaning. In between beginning and current end comes Hegel's assumption that there are social and historical determinants for the mind's categories, and Schopenhauer's idea that consciousness is often the fall guy of will and that the sexual impulse is the focus of the will. These ideas run through the works of Freud, Marx, Mannheim, and Kierkegaard. The similarities between Freud, Schopenhauer and Nietzsche are legend. That

Kierkegaard and Marx rebelled against their common teacher, Hegel, is history. There is a rich and varied intellectual tradition that gave rise to all the positions in this chapter, although most of these men came to their ideas on their own. Self-deception, it seems, was an idea whose time had come; it had a long gestation period, and when it was finally born it arrived as a family of ideas, each sibling developing in its own way.

It is interesting to compare this tradition with the one that gave rise to the accounts in the last chapter. Russell, Butler, and Siegler all share in what may broadly be called an analytic-epistemic tradition. Their approach is to understand and gain knowledge through analyzing the larger whole of experience. Norms and criteria of analysis become the standard for reality. Thus their common psychological approach is divide and conquer, as in Siegler's desire vs. intellect, and atomize and associate, as in Butler's reason siding with desire against conscience. What does not fit the bounds of analysis is lost to understanding, and is sometimes given up as lost to reality. This is what happened to self-deception: it seemed to disappear in the midst of the necessary and sufficient conditions for its occurrence. The tradition that treats the self as will is of a very different temper. Here the watchword is synthesis. Self, knowledge, value, mind, history, etc., are all swirled together in the hopes of building up the whole of reality. It is our concepts and knowledge that are always on the brink of being thrown out as inaccurate and unreal, experience and reality are

the irreducible givens which always seem to slip the bonds of understanding. The common psychological approach on this view is to "understand" (verstehen) the whole man, the person. On this approach things tend to look as they really are, all the real life complexities are captured in theory, and just because of this explanations often seem no more understandable than the real life phenomena.

Such grand perspectives on philosophical traditions and tempers are of limited worth. Broad overviews lose sight of the important particulars and when it comes to particulars, men with different perspectives often see the same thing. Thus Nietzsche's ideas seem to be the mirror image of Hume's; they both saw, from their different perspectives, the problems inherent in the notion of physical and mental substance or things-in-themselves. Butler seemed to see things with both the critical eye of Freud, and the religious eye of Kierkegaard. But even so, I think these general differences in approach are important. The analytic-epistemic tradition tends to see little, but sees it clearly and says it; the synthetic tradition tends to see much, but it is often hard to understand and decipher the accounts given from such grand, synoptic heights. The ideal would be a rapprochement of these traditions, but I suspect this is either impossible or would lead to disaster. Impossible because one perspective often interferes with and cancels out the other, leaving nothing; disaster because when virtues are matched in order to overcome adjoining faults, the issue is often all

fault with no redeeming virtue.

So much for general perspective. We now turn to the particular accounts.

Freud

It is astonishing (and this generally meets with too little acceptance) how easily and frequently intelligent people give reactions of partial feeble-mindedness under psychological constellations; anyone who is not too conceited may observe this in himself as often as he wishes, and especially when some of the thought-processes concerned are connected with unconscious or repressed motives....It is rather an every day occurrence, even in normal people, that they are deceived about the motives of their actions and do not become conscious of them until afterward, when a conflict of several emotional occurrents establishes for them the causality of such confusion.

Freud, 1907,* p. 94 and
p. 88

Freud never mentions 'self-deception'. But it is with Freudian theory that we find the fullest account of the self and its deceptions. According to Freud, a large variety of mental occurrences, including lapses of memory, rationalization of moral failings, intellectual errors and feeble-mindedness, as well as character traits, neurotic symptoms and psychotic delusions, can all be traced to our need to deny, ignore or avoid what is unacceptable to ourselves. Much of what we do also involves a willful denial of what we find unaccommodating

* Following the usual convention, date of publication, instead of number, will be used to designate the particular works of Freud.

in ourselves or reality: slips of the tongue, bungled actions, art, religious practices, even civilization itself, all rest, in part, on self-deception. Freud never spoke of self-deceit, and rarely spoke explicitly of its Freudian analog, defense mechanisms, because he did not consider self-deception a unique, remarkable occurrence, but saw it as basic to both self and society. Along with the concept of unconscious mental phenomena and the theory of the instincts, the theory of defense is one of the founding constructs upon which the large edifice of Freud's thought rests, and it is an almost inextricable part of Freud's work. (1922, p. 244) "Looked at from this angle, [psycho]analysis can be termed a systematic study of self-deception and its motivations." (Hartmann, p. 20)

Psychoanalysis begins with the hypothesis that much of what we think, feel, and do can only be explained by reference to desires, wishes, motives and intentions of which we are ignorant. We can act purposefully and not know it. Freud's insight was that we are unaware, and could never become aware, of a large portion of an underlying psychic reality. We are all constitutionally condemned to ignorance of what is basic in ourselves. This general lack of self-knowledge is aided and abetted by our desire to think well of ourselves, and so we often think of ourselves what we want, not what we are. But unlike Siegler, Russell, and Butler, Freud did not stop here. Self-deception is not simply the result of our natural ignorance nor the consequence of mistaken belief due to wishful thinking

or self-love. According to Freud, our ignorance and mistaken belief is often required by our conscious and knowing selves. Men need to defend themselves from the truth of an unacceptable psychic reality.

Our deceptions are not accidents. They serve the needed function of protecting us from otherwise unmanageable impulses, irresolvable conflicts, and unbearable frustrations. As strange as it may seem, self-deception is basic to the lives we lead and the people we are. Much of what we think, what we feel, what we do, reflects the self's need to hide, mask, or transform its own nature. Men cannot live on knowledge, truth, and reason; they also need their illusions and deceptions. For good and bad, in sickness and in health, self-deception is both natural and necessary; it is the source of error and madness, but it is also the basis of character and culture. Like helpless, disillusioned King Lear, contending with the storm from within and without, natural man, man without his defenses and deceptions, would be no more than a poor, bare, unaccommodated animal at the mercy of nature's means.

Although Freud distinguished a large number of defensive mechanisms and used these in a variety of ways to explain such diverse phenomena as slips of the tongue, psychosis and culture, the basic logic of 'defense' is clear and straightforward. All cases of defense are essentially attempts to deal with impulses that conflict and are unacceptable to our selves. Such impulses evoke "unpleasure" or anxiety, and so we try to rid ourselves of them by repressing, isolating, or denying them, or we try to

make them appear more acceptable by rationalizing or masking them to our "knowing" selves. The various mechanisms of defense all serve a common function: helping us to muddle along in a life otherwise intolerable, and irreconcilable to our conscious, rational selves. Our defenses and deceptions are the "solutions" to our conflicts that are not amenable to more rational processes nor consciously tolerable as unsolved. As such defense mechanisms are themselves neither rational nor conscious, and when they are effective there is no tell-tale sign of either the initial conflict nor its solution.

Both the offending impulse and our defensive reaction to it are unconscious mental occurrences; they are not directly experienced by us nor are they directly observable in others. Freud was led to hypothesize both an underlying psychic reality and our defense against it as a way of explaining otherwise unaccountable aspects of our conscious lives. Cases where an obvious belief is absent or in error, where an expected reaction does not materialize or unexpected behavior is manifest, and where our emotions are misplaced, misdirected, or simply missing, can all be explained as the result of our unconscious attempts to protect ourselves from the onslaught of unacceptable impulses. Successful explanations of otherwise unaccountable occurrences serve as indirect support for the initial assumption of unconscious impulse and our defensive reaction to it. Some examples will make this clear.

That children are sexual creatures is now an accepted fact. It wasn't always so. Prior to Freud's work, and not until after the displeasure with which it was greeted had subsided, infantile sexuality was neither obvious nor acceptable. Yet every parent and nurse maid since the dawn of human history must have observed the all too obvious sexual behavior of children, and it could only have been with great effort that this fact was so long overlooked. The lack of this belief and the initial rejection of it can be explained as the common attempt on the part of adults to ignore their own unacceptable sexual impulses and propensities. (1925, p. 260) For the realization that children manifest perverse sexual propensities and are sexually attracted to their parents might evoke an answering note from similar impulses in ourselves, impulses that we have kept hidden since our own childhood past. In this case it was the absence of belief that led Freud to hypothesize the adult's successful defense against his own unacceptable sexuality.

There are many cases where our defenses are not so successful, and so unconscious impulses or intentions are allowed to come to the surface. This is the case in parapraxis - slips of the tongue, failures of memory, bungled acts, and errors of judgment. Freud often used parapraxis to introduce the concepts of psychoanalysis, for such phenomena are normal, everyday occurrences, where the unconscious impulse or intention and the ensuing conflict with our knowing selves can easily be brought to light and seen for what it is. In parapraxis

an unacceptable impulse, motive, or intention is unsuccessfully defended against; it "slips out" and interferes with an ongoing psychical process. (1901, and 1917, lectures II-IV). In slips of the tongue what we would have liked to have left unsaid asserts itself by intruding on what we would like to have said; and so we admit what we think but didn't realize or want to say. A simplistic and amusing example is the story Freud tells of how a controlling wife, who, in relating that the doctor hadn't placed any restrictions on what her sick husband could eat, let slip that "he can eat and drink what I want". (1901, p. 70)

This same explanatory pattern also fits dreams, neurotic symptoms, psychosis, character traits, and our artistic, religious, and civilizing creations. Dreams are a compromise that allows unfulfilled wishes an avenue of expression while preserving for ourselves the quiet calm of sleep. Neurotic symptoms are the obtrusions into an otherwise normal waking life of feelings, thoughts or actions which express alien, unacceptable impulses that we are unable to successfully defend ourselves against. Psychosis is similar, but more extensive. An unendurable idea that is defended against comes back to haunt us, leading us to hallucinate a reality that better mirrors our frustrated wishes, like the mother who incessantly rocks the log of wood that has replaced her lost baby, or the jilted bride, arrayed in all her finery, who waits for years for her betrothed. (1894, p. 80, 1924a, pp. 185-187) Orderliness,

parsimony, and obstinacy are habitual reactions, formed through toilet training, that serve to protect us and others from our inherently opposite desires. (1908, p. 29) Even bodily attitudes, such as stiffness and rigidity, and personal peculiarities, such as fixed smiles and arrogant behavior, can be construed as the residue of past defenses that have become the protective, armor-plating of character that serves as an ever present form of defense. (A. Freud, p. 35)

The same processes writ large became the foundation for art and civilization. Thus civilization is only possible if mankind protects and conceals from itself its aggressive and incestuous impulses. Like Oedipus, each of us struggles to suppress our unacceptable, uncivilized tendencies. The extent of our success is whether we avoid being a plague to ourselves and others, and can make do with the release that art allows us in seeing our would-be tragedy played out by others. It is the other side of this, civilization's need to hide our animal, asocial beginnings that led to the communal oversight of infantile sexuality with which we began.

Freud also set out some of the specific mechanisms that underlie the many defensive phenomena. The list is long. Repression, isolation, sublimation, negation, inhibition, and rationalization, are but some examples. In sublimation, a very useful form of defense, the object or aim of an instinctual need is modified, allowing for the expression of an otherwise unacceptable need: a nascent Don Juan may become a great gyne-

cologist. (1923b, p. 181) Inhibition involves renouncing some normal function in order to avoid conflict with our unconscious selves; this often results in a stifled limited personality. (1926, pp. 14-15) Many of the eccentricities and follies of mankind result from our ability to forgo synthesizing our thoughts and actions. Opposite tendencies can be realized without conscious conflict simply by isolating or failing to unify our inconsistent thoughts and feelings: we forfeit unity of mind in order to indulge what would otherwise be conflicting and irreconcilable. (1924a, pp. 188-189) Rationalization captures Butler's "unaccountable" notion of self-deceit. By disguising otherwise unacceptable motives in whatever moral and reasonable clothes will fit, we enable ourselves to do what we want while avoiding conflict with our better selves. (1923 a, p. 46) Repression is the most important of all defense mechanisms. It involves rejecting from consciousness what is in conflict with our accepted view of ourselves. Freud often spoke of defense and repression as identical, and to the extent that all defenses are unconscious and involve the rejection of unwelcome additions to how we picture ourselves, all defenses do involve repression. (1896, p. 151; 1915a, p. 91; 1915b, p. 105; 1926, p. 17 and p. 111; and 1937, p. 318)

When all goes well we need not be defensive. Instinctual needs, reality, values, and our conscious image of ourselves are integrated into a happy, reasonable, unified self. But needs, reality, and values, do not often make a happy whole,

and when we cannot reasonably or emotionally deal with this our defenses come into play and defensive phenomena arise. This is often. Those defenses that enable us to deal with otherwise unbearable frustration, and that do not adversely affect more rational ways of realizing our needs, are good, or, at least, the best we can manage, while defenses that do not successfully avoid excessive pain, or put us at the mercy of our fantasies and distort reality beyond recognition are, of course, bad. Unfortunately life is such that often bad defenses are the best that we can manage. Since all defenses, whether good or bad, are not open to conscious control, and involve internal division that can only lead to limited, frustrated selves, the more we can be rational and open, the better off we are.

The practice of psychoanalysis is intended to help us deal more effectively with our problems. In some cases this will involve increasing our knowledge of ourselves, allowing us to consciously tolerate more frustration and enabling us to rationally approach our conflicts. In other cases this is impossible, and the best that psychoanalysis can offer is to help us replace bad defenses with more economical and effective ones. In his later writings, Freud applied the same techniques to civilization as a whole. He recommended replacing the illusions of religion with the rational approach of science as a means of dealing with life's problems. (1927) And he tried to make others realize that while discontent was a natural aspect of civilization, we would make things a bit better by replacing

an overly repressive sexual morality with one more fitting our needs. (1929)

Freud's theory of defense seems to me the most interesting and illuminating account of self-deception. It brings a common light to the different types and levels of the self's deceptions. These form a sliding scale that move from the normal and obvious errors of parapraxis, to abnormal, complex delusional world of the psychotic. Some of our deceptions are superficial, easy to cure, and avoid; others are essential in making us what we are, such as the repression of incestuous and aggressive impulses that makes civilization possible, and the defensive reactions that form our character. Failures in our defenses can lead from slips of the tongue to irreparable destruction of our psychic selves. The theory shows all of this and more to be part of a common whole.

This may be why our every day notion of self-deception seems pale besides the fuller Freudian concept. The variety of forms, aspects, and uses of the self's deception leads to an extended, imprecise and seemingly overblown view of self-deception. But I think Freud's great insight was that what seems an anomalous, strange, and fortuitous occurrence is really a common part of a greater whole. The open-textured, complex, highly theoretical Freudian view of self-deception leads to a theory that is hard to clarify, systematize, and test. As it stands, the Freudian notion of defense seems to be a practical, working concept that is incredibly incisive and valuable when

used by men of skill and vision, and easily absurd and misunderstood when in the hands of men with lesser insight.²

Philosophically the most interesting aspect of Freud's theory of defense or self-deception is the model of self upon which it rests and the view of human nature which it entails. Freud saw man as an instinctual animal who gradually gains control over the deep, unthinking forces that move him, and so develops into a partially rational, conscious person. Freud's insight into the nature and necessity of self-deceit came from the realization that the self was nothing substantial, merely a precarious balance of opposing elements, and that reason was just one way in which we solved our problems. This complex view of the mind, with its rational and instinctual sides, its need for deception as well as knowledge, was what enabled Freud to see our deceptions for what they were, part of the very fiber with which each of us weaves the complex fabric of our lives.

According to Freud, instincts, needs, wishes, or desires are the impetus behind all psychic processes. The instincts are the purposive promptings of the body which blindly seek satisfaction through the route of least resistance. Our basic instinctual self, the id, is pictured as a seething cauldron of excitations where time does not exist, reality is unknown, and the laws of logic don't apply. Contrary impulses exist side by side without cancelling each other out, the object of one instinct can be displaced on to another, opposite ideas can condense and form compromises. Unknown and unknowable forces

that blindly and irrationally seek satisfaction forever remain at the heart of psychic reality. (See 1905, p. 60; 1911, p. 22; 1915a, pp. 87-88; 1915c, p. 126; 1923b, p. 181, and 1933, pp. 73-74)

Life would not last long for an animal that constantly sought direct expression of its instinctual needs. Experience teaches patience, and frustration domesticates our instinctual selves. The satisfaction and pleasure of the instincts is still sought, but now under the limits set by awareness of the world and the rule of logic. We learn to accept reality and apply rational principles because life teaches that this in the long run is the surest way to pleasure. The transformed, unified instincts and our beliefs about the world and ourselves together make up our personalities or egos, our more or less conscious view of things. Being rational and realistic is simply a matter of learning to manage one's instincts, linking them to appropriate objects, and relating them to one another and the world. The measure of one's rationality, self-knowledge, and self-control, is the extent to which instincts have become moderated by experience and values, and integrated into a manageable whole. (See 1911, pp. 25-27; 1923a, chapter II; 1926, chapter III; 1929, pp. 12-16; 1933, pp. 75-77; 1940, p. 15)

The super-ego represents the ideals of the ego, the morals and inhibitions that we learn at an early age. Conscience, for example, comes from identifying with the authoritative parent, and internalizing and introjecting our parent's anger and repro-

bations on ourselves. With the acquisition of the super-ego comes the psychical limits that tend to control the ego's thoughts and actions, it sets the standards of how we represent our instinctual drives to our conscious, knowing selves. (See 1923a, pp. 18-19; 1929, pp. 83-91; and 1933, pp. 58-69)

This model of mind is very suggestive. Reason is simply reality modifying the instincts on their road to pleasure. Our conscious self, ego, is a half-way house between chaotic instincts and limiting values. It is a delicate balance of needs, beliefs, and values, and consists in integrating all three into a coherent picture of things. Whatever does not fit this picture is necessarily unconscious; not open to rational control, and not part of a unified self. It is even possible to imagine a number of different ego-like structures of integrated needs, values, and beliefs, that form around different loci of consciousness, but which are not themselves integrated. When all goes well we can have a clear, orderly picture of ourselves; conflicting desires and beliefs that do not fit our view of ourselves are left out of the picture and must be defended against. Consciousness, with its reasonable ways, may at any time yield to blind impulse, may compromise its beliefs to fit its needs, or deny what it cannot manage. There is no one true, rational self that we can know and be; what can be unified and controlled forms the limited, rational picture that we have of our selves.

Besides allowing for a general theory of self-deceit, this view of self and its attendant model of mind sheds light on the darker, paradoxical aspects of self-deception. On this view the "paradox" of contrary beliefs is no problem at all. Our beliefs needn't be integrated or unified. And since our conscious view of ourselves is limited to what we can synthesize into a coherent whole, not only the existence of contrary beliefs, but also our inability to see our own deceptions and defenses becomes tolerably clear.

The role of reason in the self's deceptions also begins to make sense. Self-deception is not a totally irrational act. In many cases our deceptions show the mark of skill, knowledge, and cunning. It is the clever man, with wide knowledge and sharp intellect, who is most apt to succeed in rationalizing his basest acts. Our deceptions also seem to show that we already know what we want to keep from consciousness, for how else could we be so effective in overlooking just what we want to keep hidden? All this is possible because the basic function of intellect or mind is that of seeking pleasure or mediating conflict, not that of acquiring knowledge and truth. Reason is not necessarily truth seeking or truth preserving, it is only a process that seeks consistency and success. When we can attain satisfaction through unifying our desires, beliefs, and values, we have knowledge and truth. When this is not possible and desire guides our thinking, reason becomes rationalizing, logic and knowledge are used so that conflicting ideas or impulses are made to appear consistent with the rest of our

selves. Reason leads to truly consistent thoughts and actions when it pays, and it leads to deception when conflicts are not truly remedial.

The irrational, instinctual component of our nature helps explain how self-deception can be a solution to our problems. On a surface view of things the process is rather strange. The self-deceiver's method of solving his conflicts is like the ostrich who hides its head in the sand. But how can closing our mind's eye to our inner conflicts help us to be rid of them? And how can a piece of irrationality help us to solve them? The process seems as odd and incomprehensible as the driver who closes his eyes in order to avoid a collision. Closing our mind's eye or playing peek-a-boo with ourselves can be successful as long as our constitution or reality doesn't intervene with renewed frustration or demanding impulse. For our conflicts are not rational in nature, they involve the inability to satisfy incompatible aspects of ourselves. As long as we can deny one aspect of the conflict or make a compromise between our incompatible needs and so satisfy ourselves, our conflict is resolved. Repressing sexual urges can be a solution to a problem, for the collision we need to protect ourselves against is that of the incompatibility of various aspects of our psychic selves.

Freud's view of the self also gives the semblance of sense to the fact that while reason and consciousness makes self-knowledge possible and so uncovers some of our ignorance

and deceptions, yet the very possibility of consciousness and knowledge itself seems to require a form of self-deceit. To be reasonable and conscious is to apply the strictures of consistency to what we think and do. Rational, conscious thought involves synthesizing the various aspects of ourselves into a unified, coherent whole. This unity forces exclusion of inconsistent needs, beliefs, and values. And by unifying separate ideas, values, and needs we are able to act more effectively, creating and testing possible actions through anticipatory thought. When thinking along these lines the ideal is total consciousness and the rule of rationality. This is why there is a natural pull toward knowledge and self-consistency, and away from self-deceit. Rationality and conscious control is a goal because the more integrated we are the better chance we have of realizing some needs and are better able to deal with the frustration of others.

But the natural pull toward unity and the ideal of rationality and consciousness also makes self-deception both possible and necessary. Since nature is not logical in its needs or demands, simply impulsive and assertive, the very possibility of consciousness and rationality requires that we be able to ignore or avoid what conflicts are inconsistent with how we represent or picture things to ourselves. The call of consistency can go too far and force us to exclude satisfying necessary but incompatible desires. This is where some form of self-deception becomes required. Thus through isolation, where

internally consistent but mutually exclusive sets of beliefs, desires, and values can exist side by side, we can expand the adequacy of what we think and do in specific areas without ultimately excluding conflicting needs, contrary beliefs, and incompatible values. Reason and consciousness are only viable to a creature who uses them sparingly, who separates conflicting desires by keeping some from consciousness and unifies otherwise incompatible beliefs, desires, and values in separate conscious acts.

Freud did not think that the ideal of rationality and self-knowledge was either possible or practical because there was no natural harmony to the internal psychic spheres. Consciousness and reason are only a sometimes goal for an illogical, conflicting nature. It is a goal that is itself made possible by a form of self-deceit. For creatures such as us, deceit is necessary for consciousness and reason to be possible. In some cases consciousness and rationality work well making our particular deceptions our worst enemies; in other cases our deceptions can be our best friends, making reason and consciousness possible and life a bit more manageable.

Perhaps Freud's greatest achievement was to help us to see ourselves as we really are: willful animals with a thin veneer of reason and civilization. The picture isn't pretty. Man is not as he appears to his own mind's eye: we are not just rational, conscious selves seeking the good, the true, and the beautiful. Freud sometimes described his view as a psychological

blow to man's self-love and rational image of himself, a blow comparable to Copernicus' and Darwin's onslaught on mankind's narcissism. (1917b, and 1925)

Freud's views are really more revolutionary than this. Freud's theory is not just one more step on the road to the one true picture of things. In the last analysis, it is really impossible to fully see ourselves as we really are. Man's greatest deceit is that ultimately we can fully know ourselves, that reason will rule the day, and that ignorance and self-deception will be banished to the night of dreams and madmen.

What enabled Freud to give a general account of self-deception was that he did not limit himself to the conscious mind's rational picture of itself. It is not just dreams and madness that can be traced to the irrational in us; reason and consciousness are also the fruits of an underground, psychic reality. What ultimately makes the self's deceit comprehensible is the fact that it shares a common source with reason and consciousness. Our nature is such that we require that satisfaction of conflicting needs and desires. This is the source of self-deception. It is also the reason why we cannot be totally integrated and rational in our beliefs.

Marx and Mannheim

With Marx and Mannheim we return to self-deception as an intellectual rather than a psychological phenomenon. But as with Freud's theory of defense, self-deception is now seen in a very broad perspective. The Marxian cognate for self-decep-

tion is 'ideology' and it has its cause in the correlative notion of 'false consciousness'. In false consciousness the thinking of a whole class of people within a society, or the thinking of a whole society, can be deceptive. Consciousness may be false in its very foundation; our ideas may reflect varying degrees of one grand, collective illusion. Whether such thinking is due to our selves, and so self-deceptive, depends on the interpretation that personal interest plays in Marx and that belief plays in Mannheim.

Marx used the notion of ideological, self-deceptive thought to disparage the ideas of his adversaries. Self-deception was limited to a non-Marxian view of things. Economic has-beens deceive themselves by thinking that their economic reality is what has to be. Mannheim draws out the general epistemological consequences of Marx's thought. All thinking is ideological; there is no objective truth. At the heart of Mannheim's view is the idea that belief itself is ideological in nature, the result of the individual's and society's need to organize and relate what we see into a coherent, useful whole. Knowledge becomes a sociological tool whose "truth" is determined by its pragmatic effectiveness.

Mannheim's sociology of knowledge is an epistemological analog of Freud's theory of defense. Knowledge is a socially shared perspective that we adopt in order to control and make sense of experience. In place of Freud's denial of a true self, Mannheim denies that there is an objective reality that we

gradually come to through knowledge. For Mannheim as for Freud, there is ultimately no way to know if our ideas truly reflect the way things really are. The best we can do is to try to uncover the individual, conscious determinants for our thoughts, and to come to a generally workable agreement on the nature of things.

Marx

Ideology is a process accomplished by the so-called thinker consciously, it is true, but with a false consciousness. The real motive force impelling him remains unknown to him; otherwise it simply would not be an ideological process. Hence he imagines false or seeming motive forces. Because it is a process of thought, he derives its form as well as its content from pure thought, either his own or his predecessors. He works with mere thought material, which he accepts without examination as the product of thought, and does not investigate further for a more remote source independent of thought.

Engels, 1, p. 408

While it is in Engel's 1893 letter to Mehring that we find the locus classicus on ideology and false consciousness, it was Marx who first developed these notions. (See Engels, 1, p. 407, and Marx, 1, pp. 43-44) According to Marx, the economic conditions of a thinker can lead to mistaken thinking that results in ideological beliefs. Ideologies are, in turn, fictitious intellectual creations that unconsciously reflect a thinker's economic interests, and tend to perpetuate those interests.

There are at least three ways in which consciousness can be false and belief deceptive. On the simplest level, consciousness is false because it deceives us about the nature of consciousness.

Thus it is natural for us to think that our thoughts are purely intellectual products, having their source, if any, in other thoughts. This is false. Thought is always a product of unconscious determinants that are not themselves thoughts. This follows from Marx's materialism, the view that all processes, including mental ones, are ultimately material in origin. Specifically, Marx thought that the real foundation for what we think is our place in the economic structure of society, and that a man's thought is a product that reflects his economic manipulation of his environment. (Marx, 1, p. 43; Marx and Engels, 1, p. 247)

This leads to another more important sense of false consciousness. In this sense our thoughts not only reflect unconscious, material determinants, they directly express our economic interests. This is Marx's economic determinism. The farmer, say, thinks that what is good for the farm is good for the country, and the banker thinks that a land of happy bankers is a happy land. It is not just that each group tries to realize what is in their interest, which is not deceptive or mistaken at all; but that each group wears economic blinders and thinks only in terms of its own interest and is ignorant of this, assuming all along that it is guided by obvious and unbiased motives. (Marx and Engels, 1, p. 247) False consciousness, as being influenced in what we think by our interests and being blind to this, is very much like Butler's notion of self-deception.

There is a third sense in which ideological thinking can be false, and it is this sense which is new and essential to the Marxian view of false consciousness and the attendant notion of self-deception. Consciousness is false and ideologies mistaken when what is thought does not adequately represent the underlying economic reality. Since thought, according to Marx, naturally has unconscious, economic determinants, we may naturally be mistaken about why we think as we do, but we needn't also be mistaken in what we think. Although we may be mistaken about both the nature of consciousness and the role of economic interest, what we think may yet be true: the banker may after all be right in thinking that healthy bankers are a sign of a healthy economy. Our consciousness is false and our thought ideological in this third sense when what we think is economically mistaken. Marxian 'false consciousness' applies to men's thought about their present situation, where these thoughts are due to unconscious, economic determinants, and are mistaken in representing the present economic conditions or material aspects of life.

Ideologies are intellectual theories and accounts that tend to perpetuate an economic condition that is on the wane, and in this they serve their creator's economic interests and are the result of unconscious forces. Ideologies, as the products of a false consciousness of economic reality, serve as "fettters" or "bonds" to true economic growth, and are a "specter" of the passing economic forces. This is the source of their falsity

and why they are dangerous. (Marx and Engels, 2, p. 26; Marx and Engels, 1, p. 253; Marx, 1. p. 44)

The role of ideology in Marxian theory is interesting and complicated. Originally ideologies would seem to be the beliefs or intellectual accounts that reflect the material conditions of a specific class. These ideas are about the "super-structure" of society: the values, laws, and beliefs that structure the lives of a given class in terms of its relation to the economic means of production. (Marx, 1, p. 43) The belief in the importance of savings and the sanctity of the family might be the ideological beliefs of a developing middle-class society of shop keepers. Since these beliefs initially reflect economic reality, they lead to success for their economic believers, and might be called "true ideologies". These ideas in turn cause other ideas. Belief in statehood might be the clear beginnings of an ideology that leads in steps to further ideological beliefs about the importance of law and ends with beliefs about religion and philosophy; beliefs which have only a complex and obscure connection with the original economic determinants. (Engels, 2, pp. 236-237)

Eventually the beliefs men have about themselves and their societies no longer reflect the underlying economic reality. When this happens the ruling classes' "super-structure" and ideology serve as fetters to economic development, they become economically false beliefs, and new ideologies naturally develop that more adequately reflect the changed economic foundations.

(Marx, 1, p. 44) Thus while ideologies start out as true, they all end up as out-dated and unrealistic. The only way this can be avoided is if all members of a society share equally in the various means of production. This is the Marxian ideal of a classless society, where all thought truly reflects the ongoing economic reality.

There are two possible explanations for the supposed inevitability of false consciousness and ideology. First, the farther ideological thought develops and gets from the original material reality, the less likely it is to correspond with reality. Second, because of the influence of thought on other thoughts, there might be a lag between what we think and the reality behind the thought. (There is an implicit anti-intellectualism that is built into economic determinism.) At any rate, the consequence is clear. Once realistic and economically fruitful thinking about men and their relation to society eventually becomes outmoded ideology. Marx saw the clash between the older ideological beliefs that tend to perpetuate economic has-beens and the new, realistic beliefs of the economic vanguard as a class struggle.

Marx thought that we could scientifically determine which beliefs were ideological. By determining what are the actual economic conditions of production, and comparing this with the practical consequences of our beliefs, we could see whether our beliefs would help or hinder economic growth. Those beliefs which would hurt the economic means of production are econom-

ically false and need to be discarded as ideological. (Marx, 1, p. 44; Marx, 2, p. 243)

Marx, of course, was no optimist in thinking that economic thought and reality would come together through peaceful economic evolution. People do not easily give up their beliefs in order to fit the economic facts rather than their economic needs. But Marx also believed that history marched inexorably on, and with history economics, and with economics social justice. What inevitably leads to ideology, the march of time and with it change in the economic foundations of society, also inevitably leads to a cure. For it is economic reality, not ideas, that really counts, and a change in the means of production eventually leads to a change in ideologies and successful revolution. Historical and economic progress are what must be, and our thinking and our society will eventually correct itself and reflect the underlying reality. Thus in the long run ideologies will always be false, but also in the long run it doesn't matter, for thought will always come to reflect the changed reality.

Marx used the notion of false consciousness in a limited way. He used it to explain why social thinking was often narrow and self-serving and why it was so hard to point this out to the mistaken believers. He did not think that thought was necessarily biased, that consciousness was inherently deceptive. This was the grander view of Mannheim.

Mannheim

The particular conception of ideology is implied when the term denotes that we are skeptical of the ideas and representations advanced by our opponents. They are regarded as more or less conscious disguises of the real nature of a situation, the true recognition of which would not be in accord with his interests. These distortions range all the way from conscious lies to half-conscious unwitting disguises; from calculated attempts to dupe others to self-deception.

The particular conception of ideology merges with the total.... Herewith a new and perhaps most decisive stage in the history of modes of thought has been reached.... The total conception of ideology raises a problem which has frequently been adumbrated before, but which now for the first time acquires broader significance, namely the problem of how such a thing as "false consciousness"....could ever have arisen. It is the awareness that our total outlook as distinguished from its details may be distorted, which lends to the total conception of ideology a special significance and relevance for the understanding of our social life.

Mannheim, p. 55 and
pp. 70-71*

Mannheim does not limit ideology to the influence of economic factors, nor does he assume that the works of false consciousness will inevitably be corrected. The result is that a "totally distorted mind which falsifies everything which comes within its range" (p. 70) is literally possible. It just may be that our minds have a Cartesian evil demon planted within, and so we may deceive ourselves about everything and not be able to know it. The workings of the mind may be inherently self-deceptive.

* All references are to Mannheim's Ideology and Utopia.

Mannheim is led to this extended view of false consciousness because he holds that there are "existential" or "situational" determinants for all thought. All beliefs must be interpreted against the background and circumstances of the thinker, and not, as in Marx, just economically related thoughts. Besides the particular, existential determinants behind every individual's thoughts, there are also shared situational or "social" determinants. Social determinants affect the thinking of people as a group, rather than just the thought of particular individuals. Shared "thought-structures" are reflected in a common "perspective" or "outlook" for a group, class, or a society as a whole. (pp. 55-59) Since all thought reflects existential or social factors there is no objective, presuppositional thinking. Science and even mathematics and logic are seen to involve a shared perspective, although one with a very low common denominator so that it is little open to influence by particular situational determinants, and so relatively "impartial". (p. 79)

The result of there being existential and social, thought-structures for all belief is a "relational view of knowledge". No belief or knowledge claim is a totally objective, impartial claim about what is. There isn't one right way to see the world. Belief and thought must always be seen in context. For Mannheim there is no pure, theoretical knowledge, no "true" account of things, but only different "representationally valid" views. (p. 87) The most knowledgeable among us will realize this, realize that there are different

personal determinants and social perspectives behind one's own and others' thinking, and view all belief as relative to its appropriate background. The most likely to have such a broad non-dogmatic view of "knowledge" are the "intelligentsia". These are men who, because of their social background and education, are the least likely to see their perspective as the only valid one, and so realize the relational nature of belief. (pp. 153-164)

For Mannheim all thought is ideological; there are only differences of degree. He termed the study of the existential and social determinants of thought the sociology of knowledge. And following his distinction between existential or individual determinants and social or shared determinants, Mannheim distinguished two types of ideological thought. In the "particular" conception of ideology our life-situation influences the context of our thought. Since the influence is psychological and only skin deep, it is possible to uncover and to overcome. A common criterion for validity is possible, and so agreement and understanding is, at least, possible. Marx's notion of ideology is of this sort.

The "total" conception of ideology involves the very categories and frames of reference in which we think. Here differences are not just skin deep, but reflect fundamentally different "outlooks", "perspectives", or "Weltanschauungen". Total ideology gives a general caste to our beliefs; it involves

fundamentally different ways of interpreting and understanding things. Since it lends a perspective or outlook to all our thoughts, there is no way to get behind and see our own perspective on things, and so we cannot tell if our thoughts distort or truly reflect how things are. When there are differences on this level of thought shared agreement and understanding is not possible. (pp. 55-58)

Self-deception would seem to fall within the class of particular ideological distortions, for it appears to involve our particular background leading us to see things in a certain light. (p. 55, and footnote pp. 265-266) In fact things are not so clear, and self-deception is just as much a matter of total ideology. The reason for this is that psychological and social determinants tend to merge, and with this gives the distinction between particular and total ideology. (See footnote p. 61, and pp. 69-70.) Mannheim gives two reasons for this. First, "human affairs cannot be understood by an isolation of their elements", (p. 68) so presumably there is always an interdependence of psychological and sociological determinants. Second, what we take to be a particular or isolated element is itself a consequence of social determinants. (p. 69) Whatever we may think of this reasoning, Mannheim's conclusion is clear. Our deceptions needn't be limited to psychological factors distorting what we think: all thought and belief can be deceptive.

Mannheim's own use of 'self-deception' reflects the ambiguous relation between personal "particular ideology", and impersonal, "total ideology". (See p. 15, p. 124, and p. 138) Sometimes deceptive belief is due to personal or situational factors, and sometimes it is due to impersonal, social factors. I think Mannheim's lack of clarity can be explained in terms of his general view of belief and thought.

All beliefs and cognitive processes are the result of the "cooperative process of group life". (p. 29) Thus our beliefs develop in a framework of common activity and common problems. It is also the case that all belief involves an element of will, and so reflects our personal interests and strivings. (p. 261) Thought and belief reflect determinants from both community and self, and this may be why Mannheim thinks that the personal and total views of ideology merge. But besides being the result of internal and external determinants, our beliefs also have a life of their own. Belief functions to relate ourselves to our community, and to relate our own experiences to ourselves. It is through belief that we organize and understand both ourselves and our surroundings. Our beliefs and "thought-structures" themselves influence and organize both the internal and external determinants of experience. Belief or thought serves to filter, modify, and focus personal and social determinants. The result of this melange of personal and impersonal factors, and the self-organizing and focusing role of our beliefs, is our intellectual "perspective" or "outlook".

Given this general view of the nature and role of belief, Mannheim's view of ideology and self-deception becomes tolerably clear. I think Mannheim's view of ideology and self-deception comes to something like the following. The necessary condition for deceptive belief is that a belief is held although the underlying external and internal determinants that caused this belief no longer obtain. Since there are no "objective" truths for Mannheim, such beliefs are only "pragmatically" false. Anyone acting on them is bound to act ineffectively. Mannheim would, I think, consider deceptive beliefs as cases of self-deception only if we continued to hold the belief because of the functional, integrative role that the belief played in our own thoughts. In self-deception we hold pragmatically false beliefs because they play a central role in organizing and focusing our other thoughts, and because they form our perspective or way of looking at ourselves and others. All belief can be self-deceptive since it is possible that our organizing perspective, our way of thinking, is misleading or distorting. Since we need to focus and organize what we believe so that it accords with our interests, and since this organization itself involves a general perspective, Mannheim can hold that thinking is self-deceptive in both a particular and a total ideological way.

Mannheim makes a number of interesting remarks which accord with this interpretation. Thus according to Mannheim

there is a "subconscious reluctance to think out the implications of a concretely formulated insight to a point where the theoretical formulations latent in it would be clear enough to have a disquieting effect on one's own position". (p. 277) This, I think, is explained by the organizing character of beliefs, what Mannheim calls "the meta-theoretical pragmatic nature of the human mind". (p. 275) We tend to maintain our beliefs because of the practical role they play; "true beliefs" are simply those that are more useful, and relatively false beliefs may be useful if they are the best we can manage. It is this that makes sense of Mannheim's claim that exposure or awareness of the unrealistic aspects of our beliefs will only destroy those beliefs "with which we are ourselves not too intimately identified" (p. 88), and that we tend to close our eyes and hide from the irrational elements in our outlook so that we need not change it. (pp. 42-44)

Mannheim's point is that self-deception and ideology are basic aspects of thought. Thought is ideological because it basically serves a personal and social function. And since the organizing and integrating role we require of beliefs is basic, we may lead ourselves into beliefs that are relatively ineffective and even irrational. We do not believe what we do because it is "reasonable" or accords with the evidence. We believe something because it fits our "outlook" or "perspective", and because it accords with our interests. Given the

role of personal interest and the organizing element of our perspective, there cannot be truly rational, objective belief. Belief and thought are by their very nature limited to serving a pragmatic, functional role, and there are inherent limits to what we can believe because of this. The best we can do is to become aware of our limits; there is no cure for the possible deceptions of thought.

Mannheim's view that "truth" is always a matter of degree and that beliefs are relative to the context in which they arise is a disturbing one. So too is the view that we can only understand and agree with those who share our perspective and that there is no guarantee that our perspective is even relatively right. That others agree with us may simply mean that they share our mistakes and our blindness, and that others seem mistaken may also be only a matter of perspective.

But while both relativism in regard to truth and the possibly deceptive quality of thought is disturbing, I do not think that they are logically incoherent nor easy to dismiss. The idea that truth and falsity are relative differences is sufficient to deal with the logical arguments against the incoherence of total skepticism. And it often does seem to be true that we misinterpret and misunderstand each other because we do not share a common outlook. In cases where we do agree, whether in science, ethics, or anywhere else, the ultimate basis of argument is shared assumptions or intuitions, not self-certi-

fyng reason. While there is little or no reason to think that we are totally deceived or self-deceived in what we think, there is no way to deny that this is possible.

Popper thinks that Mannheim's views tend to "destroy the basis of rational discussion, and they must lead, ultimately to anti-rationalism and mysticism". (p. 216) He is particularly critical of the view that there are unconscious determinants behind what we think, and that acceptance of a belief may be merely a matter of perspective. He argues that if this was generally accepted then all claims would become questionable, but if this is so then Mannheim's original claim of relativism needn't be accepted. He gives similar criticisms of the Freudian concept of defense; it too can be used to dismiss any claim as, say, a rationalization, and this can be used to dismiss Freud's own claims. (p. 215) Popper is right; such general claims are self-stultifying and useless. But neither Freud nor Mannheim ever in fact make such a general claim. Popper's own guarantee for objective truth as what survives the criticisms of others is clearly inadequate for undermining Mannheim's notion of total ideology, for if a whole community shares a common perspective criticism will not help to uncover it.

The issues raised here are not easily dismissed. In a way, the positions taken on irrationalism vs. rationalism,

skepticism vs. objective truth, seem to boil down to a matter of philosophical "perspectives", broad overviews on the nature of knowledge and reason. I do not see how they can be finally settled. And while this doesn't justify Mannheim's views, it at least shows that they are not without some redeeming features.

Kierkegaard, Sartre, and Fingarette

All of the positions considered so far treat self-deception as the consequence of desire, fear, unconscious motives, or the inherently distorting properties of mind, causally affecting what we believe. For Siegler, Russell, and Butler this was construed as an anomalous, disruptive occurrence. For Freud and Mannheim self-deceit was a natural outcome of how the mind functioned, and it served some useful purpose. On the positions to be considered now, self-deception is seen in a more personal light. For Kierkegaard, Sartre, and Fingarette, our deceptions are a sign that all is not well with ourselves; self-deception is not so much an intellectual error as a matter of mistaken choice, a flight from responsibility, or a lack of spiritual courage. In self-deception we are literally unwilling to be honest with ourselves; we refuse to see ourselves as we truly are; we live a life of lies.

The common approach here is a moral and subjective one. Self-knowledge, choice, responsibility, meaningfulness, and value are all seen as intimately linked together. The assump-

tion seems to be that it is only if we choose to uncover what is basic in ourselves and so come to know ourselves, that we can act responsibly and find meaning and value in life. The idea that there is a "true self" that each of us needs to know and be in order to be wise, good, and happy, seems to be a deep, philosophical truth. It is embodied in old Polonius' advice:

This above all: to thine own self be true
And it doth follow, as the night the day,
Thou canst not then be false to any man.

While the idea and importance of a true self appears to hold some deep moral and psychological insight, it is not at all clear just what this is. Since Socrates first badgered man to know themselves and live by their own inner moral light, many have muddied the waters trying to uncover the nature and source of this insight. I cannot help but feel that there is something deep below all the muddy water and that it is worth trying to salvage.

Kierkegaard

There is an ignorance about one's own life that is equally tragic for the learned and for the simple, for both are bound by the same responsibility. This ignorance is called self-deceit. There is an ignorance that by degrees, as more and more is learned, gradually changes into knowledge. But there is only one thing that can remove the other ignorance which is self-deception. And to be ignorant of the fact that there is one and only thing, and that only one thing is necessary, is still to be in self-deception.... The ignorant man can gradually acquire wisdom and knowledge, but the self-deceived one if he won "the one thing needful" would have won purity of heart.

Kierkegaard, 1, p. 52

According to Kierkegaard men deceive themselves by ignoring and failing to realize what is in their heart. The result of such self-deception is that we are not "at one" with our selves, but are in internal conflict or "double-minded". Kierkegaard is interested in double-mindedness because it shows a moral or religious failing, a failure to value or will what is most important in life. The cure for self-deception is "purity of heart", and this only comes to those who are open and honest with themselves. In coming to know our true selves we realize what is basic in life, "the Good"; and it is only by single-mindedly willing the Good that we can achieve a meaningful life.

Kierkegaard writes of this in Purity of Heart Is To Will One Thing. The book is a religious sermon that attempts to prepare the initiated for confession by helping them to see that the point of confession is to uncover what is "usually kept hidden in darkness". Through gaining knowledge of what is "true" and usually hidden in ourselves we can come to see that what is important in life is to will the Good and be committed to a religious way of life. It is ignorance of self borne of double-mindedness that prevents us from finding salvation in our true, inner selves.

Kierkegaard lists three forms of double-mindedness, ways in which men deceive themselves in confession and so fail to know themselves. The first type of self-deceit involves willing

something that is multiple or constantly changing. Since what is willed is diverse and unstable, as in the case of pleasure, it is impossible to realize what is willed. (1, pp. 55-57)

In the second form of double-mindedness we act with conflicting motives, as in acting with the desire to be good and also the desire to gain reward and avoid punishment. The result of acting with conflicting motives is that we cannot fully will what we do, and so must be unsuccessful. (1, pp. 74-77) And, lastly, we are double-minded when what is willed does not involve a total commitment. The lack of commitment shows ambivalence. It shows that what we do is not all important to us, and so we do not act with a singleness of mind or a pureness of heart. (1, pp. 104-105, p. 117, and pp. 132-133)

It is clear from this that self-deception is more an error of will than a matter of ignorance or mistaken belief. To deceive oneself is to will more than one thing, it is literally to be of two minds, and this is shown either by what we will (e.g. pleasure), how we will (e.g. the good for the sake of reward), and by the extent of our commitment (the good up to a point). Thus take the case of being double-minded because we act with two motives. Imagine a boy-scout who desires or wills to help others for its own sake, and at the same time wills or hopes for a reward. Kierkegaard would say that having as an ulterior motive the concern for reward makes it impossible to realize the central motive of acting out of con-

cern and care for others. The boy-scout's hoped for reward falsifies his care and concern: it prevents him from being a good boy-scout and so from deserving the reward. Acts done from mixed motives are seen to be errors of choice or will because they lead to internal conflict and frustration. In such cases, we are not "self-consistent" or true to ourselves and so what we do is personally impossible and cannot be successful.

I think Kierkegaard has a real insight into "the subjective reality of the individual". Although language and logic make it difficult to describe, there really is a difference between our public and private selves. Kierkegaard is trying to describe and bring out the importance of our internal, private, psychological reality, the importance of choice, consistency, and personal commitment. Cases where "inner" and "outer" selves clash, as in attempting to apologize for something although we feel we are right, trying to be spontaneous or humble, show that there is a subjective or existential reality. (See 2, pp. 176-178) According to Kierkegaard the act of willing or choosing comes close to being a direct expression of this inner reality. (3, p. 167, and p. 171) It is only if we choose what is deeply part of ourselves, express and realize what is in our heart, that we can be true to ourselves, find inner peace, purpose and value in life. Unfortunately men do not know themselves, their choices often mirror the moment's pleas-

ure or are only a matter of degree, and so we feel despair and lack inner fulfillment. Like the man who tries to be humble or apologize for what he does not feel, most of us are double-minded, we will to be other than we are, with the result that we are frustrated, in conflict, untrue to our selves.

But Kierkegaard wants to say more than this. Double-mindedness or self-deceit is not simply psychological, the source of personal confusion and frustration. Kierkegaard also thinks that the failure to know and realize one's true self is a moral and religious failure. Like Kant, Kierkegaard thought that what can be willed without conflict or contradiction is what is morally or religiously good, and not just what is good for us. To will without conflict, to act with complete commitment, is necessarily to will one thing, "the Good". Those who do not know themselves and cannot will completely or consistently are not just personally frustrated, they are also deceived as to what is of value in life. Instead of willing what has moral and religious worth, they will the moment's pleasures, or they will the Good but only to a degree, and thus they show that they do not know what is really of value. The man who does not come to terms with his true, inner self is unable to will the Good, forever at odds with himself he lives a life of personal and moral confusion and frustration. Not knowing ourselves and not choosing to be what we truly are, our lives lack inner worth and value.³

Self-deception is the ultimate personal sin. Since the meaning and value of life is to be found from within ourselves, our inability to commit our selves, our lack of self-knowledge, our not being true to our inner selves, all lead to frustrating, confused, meaningless lives.

Sartre

Good faith seeks to flee the inner disintegration of my being in the direction of the in-itself which it should be and is not. Bad faith seeks to flee the in-itself by means of the inner disintegration of my being. But it denies this very disintegration and it denies that it is itself in bad faith. Bad faith seeks by means of "not-being-what-one-is" to escape from the in-itself which I am not in the mode of being what one is not. It denies itself as bad faith and aims at the in-itself which I am not in the mode of "not-being-what-one-is-not". If bad faith is possible, it is because it is an immediate, permanent threat to every project of the human being; it is because consciousness conceals in its being a permanent risk of bad faith. The origin of this risk is the fact that the nature of human consciousness simultaneously is to be what it is not and not to be what it is.

Sartre, p. 70

The hardest thing about Sartre's view of self-deception is to understand it. Sartre seems to teeter on the brink of paradox, incoherence, and dark nay-sayings. Sartre's writing captures some of the being and nothingness that he writes about. One gets the feeling that he tries to relate in his use of language the ambiguity that he sees as part of reality.

It is as if he intentionally spoke in confusing ways, thinking that this was the way to get a clear picture of a confusing reality. But even so, some of what Sartre says is interesting, the part that I think I understand.

Sartre considers bad-faith or self-deception endemic to human existence. Because of the nature of consciousness and its inherent duplicity, human existence is something of a half-way house between "being and nothingness". The inner duplicity of human beings makes it possible for men to deny their own factual existence or being and also to deny their nothingness or freedom to transcend their existence. In either case, by denying part of their human reality, men live in bad-faith, they deceive themselves about themselves and so lead an inauthentic form of life.

Sartre develops his view of the nature of consciousness and bad-faith by comparing them to the relation between knowing subject and appearing object. Consciousness is a "transphenomenal medium" or "structure" that connects subject and object. Consciousness is itself "transparent", it has no content and it is nothing in itself. Like light, consciousness illuminates external, opaque objects and it makes knowledge of things and personal existence possible.

Since consciousness is itself clear and contentless, all particular acts of consciousness necessarily involve some object that consciousness illuminates. Sartre calls the reality of

objects when considered in themselves as having "being in-itself" or "facticity". Facts are true of such things: they have a past, causal relations, etc. Consciousness, on the other hand, is always of something, and so it is "transcendent" or a "being for-itself". Its existence is defined not by what it is but by what it chooses to be about.

In our ordinary commerce with the world we are immediately conscious of objects as, say, in seeing a table. This is "pre-reflective consciousness". In self-consciousness we reflect on pre-reflective consciousness and so become conscious that we are conscious. In self-consciousness we posit what is given in pre-reflective consciousness as an object, as something having facticity. As conscious beings, human beings have a dual and ambiguous reality, they have both facticity - are something that they are aware of - and transcendence - they are an immaterial something that can be aware of things. Take being sad as an example. In part we really are sad: sadness is part of our being; it is something we are aware of and so is really true of us. But to become conscious that we are sad involves choice, it involves focusing our consciousness on our selves, and since this is up to us, what we are is not so much given as chosen. (pp. l-lvii) If this makes sense then so does Sartre's view of self-deception.

Human consciousness is the source of both man's being and his nothingness. Since we are what we are conscious of, and

since what we are conscious of always involves choice, human existence teeters on the brink of incoherence. In bad-faith or self-deception we try to deny this inner ambiguity. We can do this in two ways. Either we deny the role of choice in making us what we are, and so identify ourselves with what we are aware of, or we can identify ourselves with our ability to choose, and so deny that we really are anything. Sartre generally speaks of the denial of the importance of choice, our responsibility in our making ourselves what we are, as the basic form of bad-faith.

Men as conscious beings are basically being for-itself. Men are what they choose to be and are not anything in themselves. When we become aware that we are the result of our choices we feel anxiety. The source of anxiety is the realization that we are free to do or be anything. The anxiety that we are responsible for ourselves, that we are what we choose to be, is unbearable, and so we tend to seek security by thinking of ourselves as determined. We see our "choices" as the result of our past, our personalities, our circumstances, or our obedience to God. The attempt to avoid responsibility for ourselves by denying our inherent freedom, transcendence, or being for-itself, leads us to see ourselves as limited, being in-itself. This is a form of bad-faith, that is endemic to human nature.

Thus men deceive themselves in thinking that their acts are required, that their lives happen to them, and they deceive themselves because they do not want to accept responsibility for what they are. And so we tend to think that the choices that give structure to our lives - our jobs, marriages, values - are all external givens. We secure for ourselves solid, closed, unanguished existences by making our choices into what has to be. (pp. 31-45) There is some truth to this, but I think it is a psychological, not an ontological truth. Men seem to feel a pull toward essentialism as a way of justifying their particular form of life. We are often freer than we care to admit, and we often seek divine justification for our personal existence.

Sartre also speaks of the denial of our facticity as a form of bad-faith or self-deception. The motive in both cases is similar, we want to avoid responsibility. By denying our facticity and identifying ourselves with consciousness and choice, we attempt to escape responsibility for choices already made. The woman who leaves her hand in the hand of her companion, becoming all thought, she disassociates herself from the sexual implications of her act, is an example of this form of bad-faith. By identifying ourselves with our transcendent, conscious selves, we try to leave the factual world of our deeds and doings behind, being all mind we escape responsibility for our bodies, being only our conscious choices we disown

responsibility for their consequences. (pp. 55-58) Sartre also sees sincerity in a similar way. Sincerity consists in consciously identifying ourselves with our acts. The waiter who thinks he is a waiter, who tries to be just what he thinks he is, no more and no less, denies the transcendent reality of consciousness. He becomes a caricature of a waiter, a cardboard image of himself. (pp. 58-59, pp. 63-65)

Like Kierkegaard, Sartre captures some very interesting aspects of consciousness and subjective experience. The dialectic of being and nothingness, of facticity and transcendence, does lead to interesting psychological descriptions. But it remains unclear to me why Sartre thinks that the duplicity he describes is necessary to human existence, and why he thinks that it is always moral in nature.

Fingarette

Each of us has biases which affect the formation of his beliefs and attitudes without his knowing or intending that they should. If having a passionate wish that something should be so brings it about unintentionally that one believes it so, this may with some propriety be called self-deception, though I believe we are much more likely to call it prejudice or wishful thinking. To the extent that we suppose the belief to be intentionally cultivated because of the wish to believe we distinguish the case as one of self-deception, a less innocent and philosophically more interesting kind of wishful thinking.

The self-deceiver is one who is in some way engaged in the world but who disavows the engagement, who will not acknowledge it even to himself as his. That it, self-deception turns upon the personal identity one accepts rather than the belief one has.

Fingarette, p. 19
and p. 67

The crucial element in self-deception according to Fingarette is that it is a purposeful act of ours. Fingarette argues against those who, like Siegler and Penelhum, construe self-deception as basically a cognitive error involving the holding of a belief contrary to evidence. Self-deceit is never simply a failure in knowledge, a case of happening to be ignorant or mistaken in our beliefs. In self-deception we intentionally come to hold our beliefs. Volition is basic to self-deceit; evidence and ignorance are second. With the shift to a volitional view of self-deception, the problem of willful ignorance becomes central, replacing the problem of contrary beliefs. Fingarette combines ideas from Kierkegaard and Sartre, along with those of Freud, in developing his account of how self-deception is possible and why it occurs.

Fingarette develops a model of consciousness that is volitional and active, rather than intellectual and passive, in order to account for intentional ignorance. On this model to become explicitly conscious of something involves a choice or act on our part. This act is one of "spelling out". It is modeled on the use of linguistic skill, and involves putting into words how we see our "engagements in the world". To be

explicitly conscious of something is to be able to give a subjective description of how we see ourselves and our surroundings. (pp. 38-43) Usually we do not spell out to ourselves what we are doing. We only do this when there is special reason, as in assessing a particularly important or unusual situation. We can also avoid becoming aware of engagements that we have reason to hide, and so willful ignorance is possible. "In general, that person in self-deception is a person of whom it is a patent characteristic that even when normally appropriate he persistently avoids spelling out some feature of his engagement in the world." (p. 47)

In order to be successful in our self-imposed ignorance, we must not be aware that we have chosen not to spell out some engagement, and so we do not spell this out. We tend to adopt some "cover-story" to hide from awareness the resulting gap in our histories, thus hiding our self-deceits from discovery. (p. 49) Like Kierkegaard and Sartre, Fingarette uses the notion of consciousness to stress the role of existential or subjective experience in how we see the world and in how we deceive ourselves.

Fingarette also tries to explain why we deceive ourselves, why we refuse to spell out or acknowledge our engagements to ourselves. Now it sometimes happens that some of our engagements conflict with our conscious image of ourselves. Avowing or spelling out such engagements will lead to internal rifts in

our self-image. By refusing to avow these engagements we preserve our view of ourselves intact, we maintain our personal identities. Self-deception or the refusal to avow unacceptable engagements, enables us to avoid conflicts by isolating our unacceptable engagements. The price we pay for this is lack of personal unity. We are reduced to a disparate community of unintegrated, "primitive selves", instead of an independent, self-aware, responsible person, whose life is the result of his conscious responsible choices.⁴

Fingarette, like Kierkegaard and Sartre, calls needed attention to what we are and how we see ourselves, and the importance of choice to all this. He also holds that inability to see ourselves for what we are involves a form of spiritual weakness, a flight from moral responsibility, a refusal to own up to what is reprehensible in us.

The notion common to Kierkegaard, Sartre, and Fingarette in that self-deception as linked to our personal identity is interesting. That self-deception is always related to one's view of oneself, one's values and concerns is a useful psychological point. But this is then linked to a moral point of view. It is assumed that not knowing ourselves or ignoring some engagement is always a source of moral confusion, personal irresponsibility, and inner dishonesty. This leads to the common moral tone taken by these writers. We are told that our lives are without value, our selves dishonest, our ignorance the source

of personal disaster and a crisis of identity. The justification for this moral point of view is not given other than to assume that there is a "true", moral self behind the facade of our everyday deceptions. The moral perspective on self-deception is no more warranted here than it was with Bishop Butler. No doubt our deceit and ignorance often lead us wrong, but I see no evidence that self-knowledge, choice, acceptance of what we do, will help lead any but the already well intentioned to do better. For many, more knowledge, acceptance and conscious commitment to what is in their hearts would quickly lead to the worst of deeds.

But perhaps what is meant by being untrue to ourselves is that we all act in ways contrary to what we know to be right, and that we deny and deceive ourselves about this. And this I think is true. We do act irresponsibly to others, cheat our children, hide our love, deceive our wives, fail ourselves, and turn from strangers. I do these things and so do you. Just think, and you'll find that lie that you live. What is it? The book not written? The child not wanted? The love that is long gone? And it is also true that we are responsible. The existentialists are surely right in all this.

But they add to this an unnecessary, additional deceit: that we didn't know this already, and that it is moral and valuable to point it all out. Well, maybe sometimes it is. Sometimes the moralist, novelist, playwright, and preacher do help us to

see our personal and moral blindnesses. But what then? Things, of course, go on much as before. We still cheat our children, desert our friends, and fail in our works. The great deceit of existentialism is the moral tone of it all, the view that there really is a "true" self and that we can be badgered into being better. It is a deceit and a tone we can well do without.

Without the presumption that we have a good, coherent self and that we should be true and responsible to it, the present account appears as similar to the psychological account of Freud. Self-deception involves denying, ignoring, or mistaking something in order to do what we want. Being "untrue to oneself" involves being in a conflict about what we do and think. The added idea that such a conflict is a sign of moral irresponsibility, divided identity, or mistaken values, seems to follow from the shared moral perspective or outlook of these writers. They all presume that there is a true, moral self behind all the deceit, and it is this self that we deny and ignore in self-deception.

Conclusion

This completes the views taken of the self in self-deceit. All these accounts share the common assumption that self-deception involves self-induced belief. Self-deceit is not just a matter of unwarranted belief, it is unwarranted belief that we somehow foist on ourselves. The variety of accounts mark the

wide divergence on how and why we induce ourselves into holding our beliefs.

The similarities and differences between these accounts are many. It would be an endless and perhaps thankless task to try to sort it all out. Sometimes our deceptions are seen to arise from uneducated appetite, sometimes from educated intellect; sometimes our deceptions are seen as immoral and unnecessary, sometimes they are natural and necessary. The character of self-deception is open and varied. We have limited ourselves to two general points of comparison. The first was the type of approach taken by the various accounts. For Butler, Kierkegaard and company, self-deception was seen in moral and religious terms; for Marx and Mannheim, self-deception was basically an economic and sociological phenomenon; with Freud it was psychological; and with Russell and Siegler the approach was essentially analytic and epistemological.

The second point of comparison was whether the self's deceptions were purposeful or whether they were simply the consequence of desire's effect on belief. The view that self-deception was unwarranted belief due to desire was considered in the last chapter. The more complex view that self-deception is in some way purposeful was considered in this chapter. It seems to me that self-deception involves purposefully ignoring or avoiding what conflicts with what we want, and that Freud's psychological approach gives the most illuminating account of this.

The view that self-deception involves believing something because of desire, need, motive, or intention, raises a number of problems. The notion that we can induce ourselves into holding our beliefs appears to be as "paradoxical" as holding contrary beliefs. These problems are dealt with next.

Footnotes to Chapter V

1 To document the growth and evolution of this tradition over the last two hundred years would, I suspect, take the next two hundred years. The best I can do is mention some of the works that seemed to me especially interesting in placing particular authors into the broad scheme of intellectual history: Gardiner's Schopenhauer, esp. 174-179; Rieff's Freud: The Mind of the Moralist; Hook's From Hegel to Marx: Studies in the Intellectual Development of Karl Marx; Merton's "Karl Mannheim and Sociology of Knowledge", pp. 490-491; Olafson's Principles and Persons; and Lowith's From Hegel to Nietzsche.

2 The incompleteness, imprecision, and speculative character of Freud's work is legend.

Freud was possessed by what I think was the deepest most telling insight into the psychological nature of man. It is an insight which needs to be studied and developed, but I do not think it likely that it will ever be surpassed. It is as if Freud saw too much, and could only sketch out the broad outlines of his vision, leaving to others the job of working out and systematizing the ideas that he left strewn about. (For some general criticisms of Freud's work see Nagel, "Methodological Issues in Psychoanalytic Theory", and Flew, "Motives and the Unconscious"; for a clarification of some of the issues and problems see Edel, "The Concept of the Unconscious: Some Analytic Preliminaries"; and for a defense of Freud's theory see Waelder, Basic Theory of Psychoanalysis, pp. 3-31.)

In the case of the theory of defense, the work of clarification and development was begun by Anna Freud in The Ego and the Mechanisms of Defense. Her book attempts to put together the scattered remarks that Freud made on the nature of the defenses, and sets out a general classification of the defenses according to the type of conflict the defense attempts to solve. Other works that try to piece together Freud's theory of defense are Brenner's "The Nature and Development of the Concept of Repression in Freud's Writings", and Madison's Freud's Concept of Repression and Defense; Its Theoretical and Observational Language.

There have also been a number of attempts to simplify and test various aspects of Freud's notion of defense. While the resulting experiments tend to corroborate Freud's theory, they seem singularly unilluminating. The gain in clarity and empirical justification seems to be less valuable than what is lost in psychological insight and speculative scope. For some examples of theory and experiment along these lines see Miller and Swanson, Inner Conflict and Defense, and articles by Sears, Mowrer, and others in Lee and Herbert's Freud and Psychology.

It seems possible that, because of the complexity of the theory, the imprecision of the terminology, the extensive use of theoretical constructs, the metaphysical assumptions

Footnotes to Chapter V - Continued

behind the theory of instincts, the lack of clear distinction between description and hypothesis, as well as the complexity of the phenomena being explained and the difficulties involved in creating appropriate experiments, the theory may never be adequately clarified and tested. But even so, the theory of defense is useful, valuable, suggestive, empirical, and, perhaps, for the most part, true.

³ For more on this aspect of Kierkegaard's thought, see the interesting work of Dilman and Phillips, Sense and Delusion, and Winch's "Can a Good Man be Harmed?".

⁴ Fingarette's book has generated much commentary and criticism. See de Sousa, "Self-Deception"; Mrinal Miri, "Self-Deception"; King-Farlow, "Self-Deception, Sincerity, and Willing to Believe"; King-Farlow and Gide, "Mauvaise Foi, Sincérité et Espoir", and Szabados, "The Morality of Self-Deception", and "Wishful Thinking and Self-Deception".

Chapter VI: The "Paradox" of Self-Induced Belief

It is clear that desire or will plays a role in self-deceit. Without some subjective source, self-deception is no more than a mysterious case of unwarranted belief. This is especially so if self-deception is held to involve unwarranted belief while knowing better. And so we were led to the role of self in self-deception as a way of distinguishing self-deception from other forms of cognitive error or to explain why we make such errors. By construing self-deceit as a case of self-made error we seem able to distinguish self-deceit from run-of-the-mill errors and to explain why such errors occur. All roads lead to the self-induced character of self-deceit.

But like believing something while knowing better, believing what we desire or will has the air of paradox about it. In the first place, holding a belief on subjective grounds conflicts with the cognitive status of belief. Believing something, after all, entails thinking that it is true. As much as we may desire or will things otherwise, we need to wait upon experience to see what the truth of the matter is. Thus the idea that our beliefs can be self-made, rather than the effect of the world working its way on us, strikes a conceptually odd note. If we should find a way around this oddity the problem becomes that of explaining how we can induce ourselves into holding our beliefs. That is, once the conceptual problem is removed we are

faced with the question of how we make ourselves believe something. We can't just believe something because we desire or will it to be so: thus try to imagine how someone might go about getting himself to believe that President Lincoln was a woman. Whatever the answer here, it seems that our ability to influence what we believe must take place in a behind the scenes, surreptitious fashion. Getting ourselves to believe something is not likely to involve any obvious procedure, for then it would be all too easy to unearth and wouldn't be effective. But if our ability to manipulate our beliefs is by its very nature difficult to detect, how can we ever be sure that it in fact takes place. By explaining the mysterious nature of self-deception in terms of the self-induced character of our unwarranted belief, we have, it seems, just answered one mystery with another.

That belief is effected by our subjective selves is a bit too common to make these problems seem truly paradoxical. By now the reason for both the oddity and its answer is clear. The mind is not in fact as clear as it appears to be. The prospective believer sees himself as self-consciously evaluating all would-be beliefs. With this view in mind, the idea that desire or will can be the source of belief appears odd. But no one is in fact an ideal believer; the picture of ourselves as self-conscious, rational believers is not borne out by the facts. Like the "paradox" of contrary beliefs, the "paradox" of the self-induced character of our beliefs is due to mistaken assumptions about the workings of the mind. In particular, we again seem

led by our view of ourselves to form a transparent model of the mind. On this model the irrationality of self-induced beliefs is clear and so it is strange to think that we could ever dupe ourselves into our beliefs.

Belief, Desire and Will

At first sight the oddity that belief may be due to desire or will appears to be conceptual in nature. This is because of belief's connection with truth. We can only believe what we think is true. Desire and will, on the other hand, are not connected with truth. We can will or desire what we know isn't so, and unfortunately we often do. Since truth is so important to us we have a natural tendency to put all truth-related occurrence in one category, and everything else in a separate category. We tend to think that belief, unlike desire and will, is involuntary, the passive result of the truth impressing itself on us.

Thus it may even seem strange to say that we desire or will to have a certain belief. The object of desire and will is usually some thing. We desire to eat; we will to have our way. Belief, on the other hand, takes propositions for its object, where propositions are truth bearing vehicles. We believe that dinner will be ready soon or that we will succeed. There really is a difference here. And perhaps if we always got our way we wouldn't bother to think about or believe in things; understanding reality would be an unnecessary luxury,

But it is clear that we can will or desire to hold certain beliefs and that holding a belief may be the object of desire or will. The dreamer, narcissist, and idealist are sometimes happiest when they can believe that a desire is fulfilled rather than actually fulfilling it. In such cases it can be said that what they desire is the belief that they are great, loved, and know how things should be, for it is the belief, not the reality, that satisfies them. Men seem to be higher than animals because they can gain at least some sustenance from the ethereal world of the mind. Our "higher" desires and concerns may take beliefs as their objects. In the case of remorse it can be the belief that we have done wrong and so blemished our self-image, rather than actually having acted wrongly, that may upset us. It is only, as Russell says, in the case of primary desires that the world alone will do.

What is basic to belief is the notion of truth. Thus whatever is presumed to be true, and whatever leads to the truth, are the natural objects and sources of belief. Whatever experience, reason and evidence teaches we are supposed to believe, and desire and will are not supposed to play a part in this. If evidence or reason show some statement to be true, we are naturally led to believe it. But if evidence or reason leaves the truth in doubt, then it is questionable whether we should or even could believe it. Desire and will are either insufficient or otiose in the case of belief. If evidence or reason are sufficient then we will believe, and have no need of will or desire. And if evidence or reason are not adequate for belief, then our belief is either wrong or out of place.

At the heart of all this is the assumption that our beliefs and our will or desire connect us with the world in different ways. Belief, as cognitive, is seen to be the natural outcome of our commerce with the world. Desire, will, intention, etc., are all out of place in regard to belief for they reflect the active aspects of mind that are relevant to action, but irrelevant to truth and belief. The traditional faculty psychology that divided mind into passive, intellectual or cognitive functions, and active, emotional or conative functions, is based on this. The view that belief is the effect of experience's influence on the mind, and that desire and will are the source of our attempts to influence the world formed the traditional faculty psychology that divided that mind into intellectual and emotional functions.

All of this merely represents in psychological terms what is basic to our notion of belief or cognition. The model of mind as passive does nothing to explain or elucidate the nature of belief or the psychological functions relevant to belief. What it does is portray in obvious metaphorical terms the idea that belief is conceptually connected with truth, and this is no more than to restate in a picture that belief presumes truth. It does not preclude that will or desire can lead us to see something as true and so believe it.

This point can be made by comparing the two traditionally conflicting accounts of belief given by Descartes and Hume. For Descartes belief was always voluntary. For Hume it was

always involuntary. As we shall see, although these two approaches appear to take opposite stands on the relation of desire or will and belief, both accounts leave room for a subjective element in belief.

For Hume, our beliefs are the direct consequence of circumstances and are not in our control. The mind is pictured as passive in belief. Thus Hume's often quoted pronouncement that "belief depends not on will, but must arise from certain determinate causes and principles, of which we are not masters". (1, p. 624) We have all experienced the quality that led Hume to speak of belief as a feeling or automatic response to circumstance. To see grey skies and slanting lines of rain is to believe, without doubt or reflection, that it is raining out. And try as we like, no desire for the sun can will this belief away.

For Descartes all beliefs are due to judgments, where judgments involve acts of will. According to Descartes it is the understanding that passively leads us to conceive certain ideas. To believe in these ideas involves a judgement that either affirms or denies the idea suggested by the understanding. The will is free and unlimited so that our judgments are completely up to us. Thus our beliefs always involve an act of will, and so for Descartes we are responsible for all our mistaken beliefs since they result from our "not making a proper use of free will". (p. 57)

Descartes is not, of course, suggesting that we can believe whatever we want, but he is saying that whenever we believe something there is an implicit choice on our part. There are situations that seem to fit Descartes' model of belief very well. These are situations where we first are appraised of the evidence for belief and perhaps deliberate about it, and then come to a conclusion about what to believe. Descartes is suggesting that we take such circumstances as the model for coming to our beliefs. Ideally we should self-consciously consider the evidence for our beliefs and only believe what is justified by the evidence. The presumption is that if we are careful and only affirm what we have considered in this way, we can always avoid mistaken beliefs. The errors we make are due not to circumstance but to ourselves. It is in this sense that Descartes said that error was due to an improper use of free will and not to our intellects.

According to Hume belief is never voluntary. According to Descartes belief always involves an act of assent or denial. For Descartes belief comes about through the influence of our will, and for Hume there is no room for such influence.

It is clear that both of these views are exaggerations. In some cases belief does seem to be the natural outcome of circumstance, and talk of a distinct act of assent is at best highly artificial. Similarly, in some cases, notably those involving deliberation, it does seem as if self-conscious choice is the basis of belief.

In fact, Descartes and Hume are not so far apart as it at first appears. Thus Hume, instead of holding, as we would expect, that all belief is beyond our control, gives advice as to what should be the basis of our beliefs. Hume's council is "mitigated" skepticism, to believe in matters of fact only what experience shows us, and with the exception of the abstract sciences of quantity and number, to ignore as sophistry and illusion all arguments and reasons that appear to give us certain truths. (2, section XII, part III) To give such council on what are the proper grounds of belief clearly implies that we can follow such a policy, and so we appear to have some choice, even if it is of a very general policy making sort, about holding beliefs.

And Descartes, for his part, says that the will is naturally inclined to affirm what is clear and distinct to the understanding. In matters where we clearly see that something is true we do not deliberate but spontaneously give our assent. Such an "automatic" act of will is similar to Hume's view that we are passive in coming to our beliefs, for in both cases belief seems to be the result of "seeing" that something is true.

But not only do Descartes and Hume seem closer on their views of belief and will than one would at first expect, their conflicting theories seem to be generated from similar beginnings. Both Descartes and Hume, and everyone else for that matter, agree that to see something as true is to believe it. The common insight here is that belief and the assumption of

truth are conceptually connected, and both construe this connection on the intuitively plausible psychological model of "seeing" as believing. What seems at first to be the key difference in their respective theories of belief and its relationship to will, is really the result of secondary considerations that are used in filling out the same conceptual approach to belief.

Hume speaks of belief only in cases of contingent matters of fact. Unlike knowledge, which involves relations between ideas that are "discoverable at first sight" (1, p. 70), beliefs are about the probable existence of things. Belief consists in the manner in which we conceive our ideas. We believe those ideas that have the force and vitality of reality for we naturally feel that they are true. Ideas that have been constantly conjoined with experience tend to have the force and feel of experience itself. It is as if repeated experience has forced us to see and come to believe that the world is a certain way. When we see grey skies and slanting lines of rain the idea that it is wet out is conceived in so lively and strong a fashion that we naturally believe that it is so. The constant conjunction of such experience in the past with wet streets, leads us now to believe that the streets will again be linked with the impression of rain.

All that Hume is doing is explaining in terms of the natural history of our experience the mental occurrence of belief. Desire and will never appear in this account because Hume's

psychology of belief only deals with ideas, impressions, and their relations. Hume's model makes belief into a natural consequence of experience, making belief "nothing but a species of sensation" which may be augmented "in some measure by reflexion and general rules". (1, p. 103, and p. 631) Hume has simply given a psychological account that fleshes out the conceptual connection of belief and presumed truth. There is nothing in this that precludes belief being the result of inner determinants that make certain ideas appear forceful, vivacious, and having all the feel of reality. Hume simply does not deal with this for he is interested in giving a plausible psychological account that relates belief and truth under ideal epistemic conditions. So Hume's denial that belief is not a matter of will is not a conceptual point at all, but rather a way of psychologically picturing the conceptual point that belief is related to truth.

Descartes' view that an act of will is involved in all cases of belief can also be traced to psychological assumptions. Unlike Hume, Descartes was interested in beliefs that were guaranteed true. This led him to emphasize the connection between belief and truth. All ideas that we see as true we automatically affirm and believe. Descartes gives a psychological criterion for determining truth; all ideas that are clear and distinct are necessarily true. (Meditation III) Since we are conscious or "see" which beliefs are true, and since we automatically affirm true beliefs, we have our infallible basis for true belief.

But this leads to a problem with mistaken belief. Why should we ever be mistaken in our beliefs if we can tell which beliefs are true by inspection, and if all beliefs except clear and distinct ones require a conscious, non-spontaneous act of assent? If true belief is to be guaranteed, then we have to work at making errors. Thus Descartes' intellectually infallible basis of true belief requires a counter-balanced psychological faculty to account for error. This is the faculty of will. Thus his view that all error is avoidable or due to an act of will is a consequence of his psychological account of knowledge; that we must believe whatever we see as true. It is not that we will to believe what is false, but since truth is guaranteed on this account of the intellect, we need some activity on our part to generate error. We err if we do not confine our assent to what is clear and distinct, and this lack of constraint involves an act of will.

Descartes' view of belief is not so different from Hume's. Here too we believe what we "see" is true. The same psychological metaphor links belief with truth in both cases. But since seeing is now a guaranteed insight, not a feeling based on experience, something additional is needed to account for error. And so Descartes is led to bring in the will. The relationship of belief and will is the consequence of the psychological model that Descartes uses to flesh out the conceptual relationship of belief and truth.

In general, I do not think that any theory that deals only with the conceptual requirement that we can only believe what we presume to be true, either requires or excludes reference to subjective conditions for belief, conditions such as desire or will. These factors are brought in or excluded only as the result of additional psychological assumptions that are used to picture the relationship of belief to truth. If belief can be infallibly and automatically true, as in Descartes, then will is needed to account for error. If belief is simply the result of the natural history of experience, and may or may not be true, then experience alone is sufficient to account for our beliefs and for their presumed truth. In neither case is there an a priori argument relating or rejecting belief and will or desire. The connection between the two is dependent on the context or the assumptions surrounding belief. Some situations will lead us to emphasize the role of will or desire in belief, and others will lead us to reject it. Some belief may involve desire or will on our part, and others will not. Descartes and Hume, because of their psychological and epistemic assumptions, each give a one-sided account and so tell only half the story.¹

Belief and Attention

The fact that belief is conceptually linked to truth does not preclude that subjective elements lead us to hold beliefs.

What it does show is that these elements, such as desire and will, can only induce belief by leading us to presume the truth of what we believe. But desire and will are not appropriate grounds for thinking something true. This leads to a common caveat in the philosophical literature on self-deception. Evidence and reason, which is the proper source of belief, must in some way play a role in the beliefs we hold. At most desire and will can play an indirect role in forming our beliefs. The role they play is supposedly that of disrupting the normal cognitive process that usually leads to belief. They do this either by directing what evidence we attend to or by misleading us in evaluating that evidence.

Thus Butler points out that a man under the sway of a particular passion is likely to be keenly aware of what furthers that passion, and may simply fail to notice constraining influences. (Butler, X, 4) Similarly, Siegler argues that we do not believe something on bad or irrelevant evidence because we want to believe what we realize isn't so. It is rather that we are so anxious to find evidence for a desired belief that we mistakenly think inadequate or irrelevant evidence to be better than it is. (2, pp. 34-35) And in cases where we do notice evidence against a desired belief, we can reduce its importance to the vanishing point by considering how many slight but telling ways our situation is different, and so the evidence doesn't count. (Gardiner, 1; and Rorty, 1) We may also make use of an

even simpler expedient and fail to realize that the evidence is a good reason for not believing what we want. (Demos and Pugmire)

In all these cases, desire effects belief through the indirect path of guiding attention to evidence favorable for a belief and away from evidence against the belief. Thus there is always some evidence, reason, or justification that leads us to presume the truth of what we believe. Desire remains in the background as the source of this "evidence", "reason", and "justification".

Will can effect belief in a more direct, "up-front" fashion, although it too supposedly always works through attention. We may come to think that our upbringing has led us to be prejudiced, that our past experience has induced us into a general distrust of others, and that our circle of friends has inculcated in us a harsh, unfeeling outlook on life. And while we may not be able to immediately lose our prejudice, distrust, or cynicism, we may in time change our way of thinking by noticing when we are prejudiced, by acting as if we trusted others, and by associating with people who have a more genial outlook. We can control our beliefs by being aware of them and by putting ourselves in circumstances that naturally tend to modify them.

Price seems to think that all cases of cultivated belief involve the willful control of our attention. By attending to the evidence that favors the belief we will to have, and by

ignoring adverse evidence, we can indirectly, although voluntarily, come to change our beliefs. (Price, 2, p. 107) Thus Victorian ladies were supposedly able to believe in their husbands' fidelity by ignoring the gossip and attending instead to all the signs of their husbands' family virtues. Price also considers a second way in which we may voluntarily come to hold a belief through controlling our attention. This involves imagining circumstances or situations which would obtain if a certain belief were true. By fixing our attention on imagined situations we can see more clearly what is involved in a belief and come to take the belief seriously. (Price, 2, p. 109) This method is often recommended for developing and acting on our ideals, as in creating a belief that peace is possible by imagining what life would be like if it were to obtain. At any rate, it seems clear that there are cases where we can voluntarily come to hold a belief, although we cannot do this directly as a matter of will.

I do not think that the subjective elements in belief need to be limited in this way. The view that we can only induce ourselves into holding a belief by manipulating evidence and reason assumes that evidence and reason are always involved in holding beliefs. Caught by the conceptual fact that we can believe only what we think is true, it is assumed that what justifiably allows for presumed truth is the natural basis of belief. Desire and will are thus seen as always working in a

behind the scenes way to manipulate cognitive processes. The picture here is that of a self-conscious, rational believer who in seeking the truth is misled by subjective factors.

But why can't subjective factors be the primary source of our beliefs? Why can't we believe something first and find evidence and reason for our belief second? I think that this in fact is what often happens in belief. Subjective factors as well as evidence and reason form the basis of most beliefs. It is only in special times that evidence and reason alone are the source of belief. More often than not justification for belief is a matter of hindsight. Desire and will along with evidence and reason form our beliefs.

Countless times each day our beliefs more truly mirror subjective factors than they do rational, objective ones. When waiting for a bus people seem to naturally believe that the first bus they see is the bus they want. I suspect that people who buy lottery tickets often believe that they will win, not just that they have a chance of winning, and they believe this although they know that their expectations are unwarranted. Much of what we believe is merely the passive acceptance of what others say, the end result of a common non-rational lesson. And even when we are self-conscious about our beliefs and seek advice as to what to believe, our choice of advisors often shows that we have already made up our minds. Our beliefs are usually not held on evidence and reason, but are due to personal commitment, social pressure, or because they just feel right to us.

(See Peirce's "The Fixation of Belief".)

Often when we do seek evidence and reason for what we believe it comes after having formed the belief in question. Consider the case of post-hypnotic suggestion. A man is hypnotized and told to open an umbrella at a certain time. He does this and is asked why. Invariably such a person gives some reason, such as he wanted to see if the umbrella works, believing his act his own and his reason its explanation. In many cases we realize that others will not find our beliefs acceptable, and so seek reasons to justify those beliefs for others, the belief being acceptable to us on less rational grounds. Much of philosophy is like this. It is, as Bradley said, the attempt to give rational grounds for our most primitive beliefs that are normally formed and held without reason or evidence. (See Bradley, p. xiv.)

The view that reason and evidence always initiate our beliefs, and that subjective factors only distort beliefs by manipulating cognitive processes is a much too rational picture of things. Desire and will can induce us into holding a belief and they can also induce us into thinking that there is reason or evidence for these beliefs. Subjective factors can directly lead to belief, and need not travel the more circuitous route of influencing belief through effecting our attention.

Conclusion

The reason why the self-induced character of our beliefs seems paradoxical is that when we think about beliefs we do so in a self-conscious, rational fashion. From this perspective the subjective factors appear as an anomaly and of only minor importance. We all know that our beliefs should be held on rational grounds and so it seems odd that desire or will can play a role in the formation of our beliefs. And when we are forced to admit that they do play such a role, we naturally assume that they do so in the indirect way of manipulating our more rational faculties.

The philosophical view of man, unlike the psychological or sociological, takes this rational perspective and builds it into a general model of mind. Belief becomes identified with rational belief. The result is that the subjective factors that lead to belief are delegated to a minor role. It is this view that makes the positions of Freud and Mannheim appear to be overblown. It is also this view that tends to limit accounts of self-deception to aberrations in our normal cognitive processes. But we are not the self-conscious rational believers that we see ourselves as being.

To be self-consciously rational about our beliefs is something of an achievement. It takes a good deal of sophistication about the appropriate criteria for belief and a good deal of effort to weigh evidence and reason out the relationship

between the evidence and our other beliefs. Most beliefs probably fall in between a self-conscious critical evaluation of whether a specific belief is justified and a totally non-critical response to belief. But at any rate, the notion that we always buck reason when we come to believe something from subjective factors is to judge belief from a rather one-sided, self-conscious perspective. Desire and will as well as reason make our beliefs credible to us. And no matter how well experience or logic may have taught us that belief commensurate with evidence and reason is best, it is not likely that we are always able to be critical and in control of our subjective tendencies to hold beliefs. So even though we know better, desire and will, as well as reason, can directly father our beliefs.

Footnotes to Chapter VI

¹ For more on Hume's and especially Descartes' views on the relationship of belief and will see: Price, Belief, series I, lecture 10; O'Hear, "Belief and the Will", and Evans, "Error and the Will".

Chapter VII: Conclusion

Men can believe what they know isn't so. We are able to hold beliefs that conflict with what we know but do not want to see. It is a common occurrence even in the lives of the most rational of men. And yet the idea that we can deceive ourselves into holding unwarranted beliefs seems paradoxical.

It is the seemingly paradoxical aspects of self-deception that have led philosophers to attempt to define and explain it without characterizing the self-deceiver as someone who knows that his beliefs are mistaken. We have considered two basic attempts to account for self-deception along these lines. The first of these was the view that self-deception was simply a case of holding an unwarranted belief. This was the position of Siegler (1) and (2), Canfield-Gustavson, and Penelhum. According to them, self-deception involves mistaken belief when we should know better. No reference is made to why the self-deceiver believes what he does nor why his belief is unwarranted for him. The problem with this view is that self-deception is not distinguished from other forms of cognitive error. The second attempt construes self-deception as a case of motivated belief or wishful thinking. Desire or interest misleads the self-deceiver to see only what he wants to believe. This was the position of Siegler (4), Russell and Butler. While this account does explain why the self-deceiver holds his belief, it does not explain why his belief is deceptive nor why he

holds his belief in just those circumstances that tell against it.

On both of these accounts self-deception is seen as basically a case of cognitive error. The self-deceiver is pictured as someone who is rational and wants to know, but who is somehow misled into holding an unwarranted belief. He does not know that his belief is mistaken nor does he try to induce himself into believing what he wants, his deception is simply the result of a snag in otherwise truth-oriented rational processes that normally lead to warranted belief. These accounts leave out what is central to self-deception. The self-deceiver isn't someone who wants to know and happens to be mistaken, he does know and he wants to avoid this knowledge. His deceptive belief is the result of irrational processes that enable him to ignore the truth that he doesn't want to see. The cuckold sees that his wife is unfaithful, but he cannot accept this. If he consciously believed that his wife was having an affair he would have to question his masculinity, feel compelled to break up his marriage, accept his friends' jests and his family's scorn. It is because he knows this that he is anxious to avoid seeing that he is being cuckolded. He finds excuses for his wife's behavior, he stays at work until late in order to avoid having to deal with her evening absences, he persuades himself that his friends are malicious and he stays away from his family.

This is the view of self-deception given by Freud, Sartre

and Fingarette among others. The self-deceiver holds an unwarranted belief in order to avoid the unpleasant truth. What makes this account seem paradoxical is the view that our beliefs are rational and must reflect what we see is true. This is the assumption that leads to the view that self-deception is merely a case of unwarranted belief. Basic to the account of self-deception given by Freud and others is the denial that belief is always the result of rational processes. Although these accounts differ in how they explain how self-deception occurs, they all involve a sufficiently complicated view of the mind that allows for the irrationality of belief. Once the inherently irrational character of belief is accepted, deceiving ourselves into believing what we want although contrary to what we know ceases to be paradoxical.

The seemingly paradoxical character of self-deception is the result of assumptions made about the nature of the mind's workings and the attendant view of belief. If the mind is seen as transparent in belief so that our beliefs are held to coincide with what we see is true, then self-deception is impossible. This is the model of the mind that leads to the view that self-deception is merely a matter of unwarranted belief. In order to develop an adequate account of self-deception a complex view of the mind that allows for the fact that we can hold beliefs while knowing better is required. There is no way to avoid this conclusion without denying that there is self-

deception. The ultimate source of paradox in self-deception is an unwillingness to give up the cherished view that we are always rational in coming to our beliefs. This is just what must be denied if self-deception is to be seen for what it really is.

List of Works Cited

- G.E.M. Amscombe, "Pretending," in Philosophy of Mind, ed. by S. Hampshire, New York: Harper & Row, 1966.
- J. L. Austin, (1) "Pretending," in Philosophical Papers, ed. by J. O. Uninson and G. J. Warnock, Oxford: Oxford University Press, 1961.
- _____, (2) Sense and Sensibilia, reconstructed from manuscript notes by G. J. Warnock, New York: Oxford University Press, 1964.
- Y. Bar-Hillel, "Analysis of 'Correct' Language," in Mind, 55 (1946).
- Ernest Becker, The Denial of Death, New York: Free Press, 1973.
- Errol Bedford, "Emotions," in Essays in Philosophical Psychology, ed. by D. F. Gustavson, Garden City: Doubleday Anchor, 1964.
- Peter L. Berger, Invitation to Sociology, Garden City: Doubleday Anchor, 1963.
- Max Black, "Saying and Disbelieving," in Analysis, 13 (1952).
- F. H. Bradley, Appearance and Reality, New York: Oxford University Press, 1969.
- R. B. Braithwaite, "Half-Belief," in Proceedings of the Aristotelian Society, Supp. Vol. 38, (1964).
- Charles Brenner, "The Nature and Development of the Concept of Repression in Freud's Writings," in The Psychoanalytic Study of the Child, 12 (1957).
- Brett's History of Psychology, ed. and abridged by R. S. Peters, Cambridge, Mass: M.I.T., 1967.
- Joseph Butler, The Works of Joseph Butler, vols. I and II, prefaced by Samuel Halifax, New York: Carter and Bros.
- John Canfield and Don F. Gustavson, "Self-Deception," in Analysis, 23 (1962).
- John Canfield and Patrick McNally, "Paradoxes of Self-Deception," in Analysis, 21 (1961).

List of Works Cited - Continued

- Stanley Cavell, "The Avoidance of Love: A Reading of King Lear," in Must We Mean What We Say, New York: Scribner's and Sons, 1969.
- Gerald A. Cohen, "Belief and Roles," in Proceedings of the Aristotelian Society, 67 ns (1966-67).
- Neil Cooper, "Consistency," in Philosophical Quarterly, 16 (1966).
- Dante, The Divine Comedy 1, Hell, trans. by D.C. Sayers, Baltimore: Penguin, 1968.
- Donald Davidson, "How is Weakness of Will Possible," in Moral Concepts, ed. by J. Feinberg, New York: Oxford University Press, 1970.
- Ronald B. de Sousa, "Self-Deception,," in Inquiry, 13 (1970).
- Raphael Demos, "Lying to Oneself," in The Journal of Philosophy, 57 (1960).
- Rene Descartes, Meditations on First Philosophy, trans. by L. J. Lafleur, Indianapolis: Bobbs-Merrill, 1960.
- Charles Dickens, Great Expectations, New York: New American Library, 1963.
- Ilham Dilman and D.Z. Phillips, Sense and Delusion, New York: Humanities Press, 1971.
- John Dollard and Neal E. Miller, Personality and Psychotherapy, New York: McGraw Hill, 1966.
- Fyodor Dostoevsky, Notes From Underground, trans. by R.E. Matlaw, New York: Dulton, 1960.
- Marcel Eck, Lies and Truth, trans. by B. Murchland, London: Macmillan, 1970.
- Abraham Edel, "The Concept of the Unconscious: Some Analytic Preliminaries," in Journal of General Psychology, 59 (1954).
- Friedrich Engels, (1) "Letter to Franz Mehring," in Marx and Engels, ed. by F.S. Feuer, Graden City: Doubleday Anchor, 1959.

List of Works Cited - Continued

Friedrich Engels, (2) "Ludwig Feuerbach and the End of Classical German Philosophy," in Marx and Engels, ed. by F.S. Feuer, Garden City: Doubleday Anchor, 1959.

J.L. Evans, "Error and the Will," in Philosophy, 38 (1963).

Herbert Fingarette, Self-Deception, London: Routledge and Kegan Paul, 1969.

Anthony Flew, "Motives and the Unconscious," in The Minnesota Studies of the Philosophy of Science I, ed. by H. Feigl and M. Scriven, Minneapolis: University of Minneapolis Press, 1956.

Anna Freud, The Ego and The Mechanisms of Defense, trans. by C. Barnes, New York: International Universities Press, 1973.

Sigmund Freud, (1894) "The Defense Neuro-Psychosis," trans. by J. Rickman, in Early Psychological Writings, ed. by R. Rieff, New York: Macmillan, 1969.

_____, (1896) "Further Remarks on the Defense Neuro-Psychosis," trans. by J. Rickman, in Early Psychological Writings, ed. by P. Rieff, New York: Macmillan, 1969.

_____, (1900) The Interpretation of Dreams, trans. by J. Strachey, New York: Avon, 1971.

_____, (1901) The Psychopathology of Everyday Life, trans. by A. Tyson, ed. by J. Strachey, New York: Norton, 1965.

_____, (1905) Three Essays on the Theory of Sexuality, trans. and ed. by J. Strachey, New York: Avon, 1965.

_____, (1907) "Delusion and Dream," trans. by H. Zohn, in Delusion and Dream, ed. by P. Rieff, Boston: Beacon Press, 1968.

_____, (1908) "Character and Anal Erotism," trans. by R. C. McWatters, in Character and Culture, ed. by P. Rieff, New York: Macmillan, 1970.

List of Works Cited - Continued

- _____, (1911) "Formulations Regarding the Two Principles in Mental Functioning," trans. by M.N. Searl, in General Psychological Theory, ed. P. Rieff, New York: Macmillan, 1970.
- _____, (1915a) "Instincts and Their Vicissitudes," trans. by C. Barnes, in General Psychological Theory, ed. by P. Rieff, New York: Macmillan, 1970.
- _____, (1915b) "Repression," trans. by C. Barnes, in General Psychological Theory, ed. by P. Rieff, New York: Macmillan, 1970.
- _____, (1915c) "The Unconscious," trans. by C. Barnes, in General Psychological Theory, ed. by P. Rieff, New York: Macmillan, 1970.
- _____, (1917a) A General Introduction to Psychoanalysis, trans. by J. Riviere, New York: Simon and Schuster, 1969.
- _____, (1917b) "One of the Difficulties of Psychoanalysis," trans. by J. Riviere, in Character and Culture, ed. by P. Rieff, New York: Macmillan, 1970.
- _____, (1922) "Psychoanalysis," trans. by J. Strachey, in Character and Culture, ed. by P. Rieff, New York: Macmillan, 1970.
- _____, (1923a) The Ego and The Id, trans. by J. Riviere, ed. and revised by J. Strachey, New York: Norton, 1962.
- _____, (1923b) "The Libido Theory," trans. by J. Riviere, in General Psychological Theory, ed. by P. Rieff, New York: Macmillan, 1970.
- _____, (1924a) "Neurosis and Psychosis," trans. by J. Riviere, in General Psychological Theory, ed. by P. Rieff, New York: Macmillan, 1970.
- _____, (1924b) "The Loss of Reality in Neurosis and Psychosis," trans. by J. Riviere, in General Psychological Theory, ed. by P. Rieff, New York: Macmillan, 1970.

List of Works Cited - Continued

- Sigmund Freud, (1925) "The Resistance to Psychoanalysis," trans. by J. Strachey, in Character and Culture, ed. by P. Rieff, New York: Macmillan, 1970.
- _____, (1926) The Problem of Anxiety, trans. by H. Bunker, New York: Norton, 1965.
- _____, (1927) The Future of An Illusion, trans. by W.D. Robson-Scott, re. and ed. by J. Strachey, Garden City: Doubleday Anchor, 1964.
- _____, (1929) Civilization and Its Discontents, trans. and ed. by J. Strachey, New York: Norton, 1962.
- _____, (1933) New Introductory Lectures on Psychoanalysis, trans. and ed. by J. Strachey, New York: Norton, 1965.
- _____, (1937) "Disturbance of Memory on the Acropolis," trans. by J. Strachey, in Character and Culture, ed. P. Rieff, New York: Macmillan, 1970.
- _____, (1940) An Outline of Psychoanalysis, trans. by J. Strachey, New York: Norton, 1963.
- Patrick Gardiner, (1) "Error, Faith and Self-Deception," in Proceedings of the Aristotelian Society, 68 ns. (1969-70).
- _____, (2) Schopenhauer, Baltimore: Penguin, 1971.
- Mitchell Ginsberg, Mind and Belief, New York: Humanities Press, 1972.
- Erving Goffman, The Presentation of Self in Everyday Life, Garden City: Doubleday Anchor, 1959.
- A. Phillips Griffiths, "On Belief," in Knowledge and Belief, ed. by A. P. Griffiths, Oxford: Oxford University Press, 1967.
- D. W. Hamlyn, "Symposium: Self-Deception," in Proceedings of the Aristotelian Society, Supp. Vol. 45 (1971).

List of Works Cited

- W.F.R. Hardie, "Aristotle on Moral Weakness," in Weakness of Will, ed. by G.W. Mortimore, London: St. Martin's Press, 1971.
- R.M. Hare, Freedom and Reason, Oxford: Clarendon Press, 1964.
- Heinz Hartmann, "Psychoanalysis as a Scientific Theory," in Psychoanalysis: Scientific Method and Philosophy, ed. by S. Hook, New York: New York University Press, 1964.
- Carl G. Hempel, "Rational Action," in Readings in the Theory of Action, ed. by N.S. Care and C. Landesman, Bloomington: Indiana University Press, 1968.
- Jaakko Hintikka, Knowledge and Belief, Ithaca: Cornell University Press, 1964.
- Sidney Hook, From Hegel to Marx: Studies in the Intellectual Development of Karl Marx, Ann Arbor: University of Michigan Press, 1962.
- David Hume, (1) A Treatise of Human Nature, ed. by L.A. Selby-Bigge, Oxford: Clarendon Press, 1965.
- _____, (2) Enquiry Concerning the Human Understanding, in Hume's Enquiries, sec. ed., ed. by L.A. Selby-Bigge, Oxford: Clarendon Press, 1962.
- _____, (3) Enquiry Concerning the Principles of Morals, in Hume's Enquiries, sec. ed., ed. by L.A. Selby-Bigge, Oxford: Clarendon Press, 1962.
- Arnold Isenberg, "Oeontology and the Ethics of Lying," in Ethics, ed. by J.J. Thomson and G. Dworkin, New York: Harper and Row, 1968.
- Immanuel Kant, Foundations of the Metaphysics of Morals, trans. with an intro. by L.W. Beck, Indianapolis: Bobbs-Merrill, 1959.
- Anthony Kenny, Action, Emotion and Will, London: Routledge and Kegan Paul, 1963.

List of Works Cited

John King-Farlow, (1) "Self-Deceivers and Sartrian Seducers," in Analysis, 23 (1963).

_____, (2) "Self-Deception, Sincerity and Willing to Believe, (unpublished paper)

_____ and Lorraine Gide, "Bonne Foi, Mauvaise Foi, Sincerite et Esprit," (unpublished paper).

Soren Kierkegaard, (1) Purity of Heart, New York: Harper and Row, 1956.

_____, (2) Concluding Unscientific Postscript, Princeton: Princeton University Press, 1968.

_____, (3) Either/Or, vol. II, Garden City: Doubleday Anchor, 1959.

Thomas Kuhn, The Structure of Scientific Revolutions, second edition, enlarged, Chicago: Chicago University Press, 1972.

R.D. Laing, The Divided Self, Baltimore: Penguin, 1965.

_____, The Politics of the Family, Toronto: CBC Publications, 1969.

Charles Landesman, Discourse and Its Presuppositions, New Haven: Yale University Press, 1972.

C. H. Langford, "Notions of Analysis in Moore's Philosophy," in The Philosophy of G.E. Moore, ed. by P.A. Schilpp, Evanston: Northwestern University Press, 1942.

S.G.M. Lee and Martin Herbert, eds., Freud and Psychology, Baltimore: Penguin Books, 1970.

T. Foster Lindley, "Lying and Falsity," in Australasian Journal of Philosophy, 49 (1971).

Leonard Linsky, "Deception," in Inquiry, 6 (1963).

Karl Lowith, From Hegel to Nietzsche, Garden City: Doubleday Anchor, 1967.

List of Works Cited

- A.M. MacIver, "Some Questions About 'Know' and 'Think'," in Analysis, 5 (1938).
- John Macmurray, Persons in Relation, London: Faber and Faber, 1961.
- Peter Madison, Freud's Concept of Repression and Defense, Minneapolis: University of Minneapolis Press, 1961.
- Karl Mannheim, Ideology and Utopia, New York: Harcourt, Brace and World, 1936.
- D.S. Mannison, "Lying and Lies," in Australasian Journal of Philosophy, 47 (1969).
- Karl Marx, (1) "Excerpts from 'A Contribution to the Critique of Political Economy,'" in Marx and Engels, ed. by F.S. Feuer, Garden City: Doubleday Anchor, 1959.
- _____, (2) "Theses on Feuerbach," in Marx and Engels, ed. by F.S. Feuer, Garden City: Doubleday Anchor, 1959.
- _____ and Friedrich Engels, (1) "Excerpts from The German Ideology," in Marx and Engels, ed. by F.S. Feuer, Garden City: Doubleday Anchor, 1959.
- _____ and Friedrich Engels, (2) "Manifesto of the Communist Party," in Marx and Engels, ed. by F.S. Feuer, Garden City: Doubleday Anchor, 1959.
- Robert Merton, "Karl Mannheim and the Sociology of Knowledge," in Social Theory and Social Structure, New York: Free Press, 1957.
- John Stuart Mill, Utilitarianism, Indianapolis: Bobbs-Merrill, 1957.
- Daniel R. Miller and Guy Swanson, Inner Conflict and Defense, New York: Schocken Books, 1966.
- Neal E. Miller, "Comments on Theoretical Models Illustrated by the Development of a Theory of Conflict," in Journal of Personality, 20 (1951-52).

List of Works Cited - Continued

- Ronald D. Milo, Aristotle on Practical Knowledge and Weakness of Will, Hague: Mouton, 1966.
- Mrinal Miri, "Self-Deception," in Philosophy and Phenomenological Research, 34 (1974).
- G.E. Moore, "A Reply to My Critics," in The Philosophy of G.E. Moore, ed. by P.A. Schilpp. Evanston: Northwestern University Press, 1942.
- _____, "Russell's Theory of Descriptions," in The Philosophy of Bertrand Russell, ed. by P. A. Schilpp. Evanston: Northwestern University Press, 1944.
- N. O. Mounce, "Symposium: Self-Deception," in Proceedings of the Aristotelian Society, Supp. Vol. 45 (1971).
- A.E. Murphy, The Theory of Practical Reason, La Salle: Open Court, 1965.
- G. Murphy, "Experiments in over-coming self-deception," in Biofeedback and Self-Control 1970, ed. by T. Barber et. al., New York: Anodyne, 1971.
- Ernest Nagel, "Methodological Issues in Psychoanalytic Theory," in Psychoanalysis, Scientific Method and Philosophy, ed. by S. Hook, New York: New York University Press, 1964.
- Friedrich Nietzsche, Beyond Good and Evil, trans. with a commentary by Walter Kaufmann, New York: Vintage Books, 1966.
- D.J. O'Connor, "Beliefs, Dispositions and Actions," in Proceedings of the Aristotelian Society, 49 ns (1969).
- Anthony O'Hear, "Belief and the Will," in Philosophy, 47 (1972).
- Frederick A. Olafson, Principles and Persons, Baltimore: John Hopkins University Press, 1967.
- Stanley Paluch, "Self-Deception," in Inquiry, 10 (1967).

List of Works Cited - Continued

- Arthur Pap, Semantics and Necessary Truth, New Haven: Yale University Press, 1968.
- David Pears, "Freud, Sartre and Self-Deception," in Freud: A Collection of Critical Essays, ed. by R. Wollheim, Garden City: Doubleday Anchor, 1966.
- Charles Peirce, "The Fixation of Beliefs," in Collected Papers, Vol. 5, eds. by C. Haulshorne and P. Weiss, Cambridge: Harvard University Press, 1960.
- Terence Penelhum, (1) "Pleasure and Falsity," in Philosophy of Mind, ed. by S. Hampshire, New York: Harper and Row, 1966.
- _____, (2) "Appendix A: Self-Deception," in Problems of Religious Knowledge, London: Macmillan, 1971.
- R. S. Peters, "Motivation, Emotion and the Conceptual Schemes of Common Sense," in Human Action, ed. by T. Mischel, New York: Academic Press, 1969.
- Karl Popper, The Open Society and Its Enemies, vol. II, New York: Harper and Row, 1967.
- H. H. Price, (1) Belief, London: George Allen and Unwin, 1969.
- _____, (2) "Belief and Will," in Philosophy of Mind, ed. by S. Hampshire, New York: Harper and Row, 1966.
- _____, (3) "Half-Belief," in Proceedings of the Aristotelian Society, Supp. Vol. 38 (1964).
- David Pugmire, "'Strong' Self-Deception," in Inquiry, 12 (1969).
- W. V. Quine, "The Ways of Paradox," in The Ways of Paradox and Other Essays, New York: Random House, 1966.
- Philip Rieff, Freud: The Mind of a Moralist, Garden City: Doubleday Anchor, 1961.
- Amelie Oksenberg Rorty, (1) "Belief and Self-Deception," in Inquiry, 15 (1972).
- _____, (2) "Plato and Aristotle on Belief, Habit, and Akrasia," in American Philosophical Quarterly, 7 (1970).

List of Works Cited

- Bertrand Russell, The Analysis of Mind, London: George Allen Unwin, 1968.
- Gilbert Ryle, Dilemmas, Cambridge: Cambridge University Press, 1964.
- _____, The Concept of Mind, New York: Barnes and Noble, 1965.
- Gerasimos Santas, "Plato's Protagoras and explanations of weakness," in Weakness of Will, ed. by G. W. Mortimore, London: St. Martin's Press, 1971.
- Jean-Paul Sartre, Being and Nothingness, trans. by H. Barnes, New York, Philosophical Library, 1956.
- Roy Schafer, "The Idea of Resistance," in The International Journal of Psychoanalysis, 54 (1973).
- James Shea, Jr., Self-Deception, Ann Arbor: University Microfilms, 1966.
- F. A. Siegler, (1) "Demos on Lying to Oneself," in The Journal of Philosophy, 59 (1962).
- _____, (2) "Self-Deception," in Australasian Journal of Philosophy, 41 (1963).
- _____, (3) "Lying," in American Philosophical Quarterly, 3 (1966).
- _____, (4) "An Analysis of Self-Deception," in Nous, 2, (1968).
- Laurent Stern, "On Make-Believe," in Philosophy and Phenomenological Research, 28 (1967).
- Bela Szabados, "Wishful Thinking and Self-Deception," in Analysis, 33 (1973).
- _____, "The Morality of Self-Deception," in Dialogue, 13 (1974).
- _____, "Rorty on Belief and Self-Deception," in Inquiry, 17 (1974).
- Thomas Szasz, The Myth of Mental Illness, New York: Delta, 1968.
- Leo Tolstoy, War and Peace, vol. 1, Baltimore Penguin Books, 1972.

List of Works Cited - Continued

- Lionel Trilling, Sincerity and Authenticity, Cambridge, Mass.:
Harvard University Press, 1973.
- Robert Waelder, Basic Theory of Psychoanalysis, New York:
Schocken Books, 1971.
- Bernard Williams, "Ethical Consistency," in Proceedings of the
Aristotelian Society, Supp. Vol. 39 (1965).
- Peter Winch, "Can a Good Man Be Harmed," in Proceedings of the
Aristotelian Society, 44 ns (1965-66).
- John Wisdom, "Paradox and Discovery," in Paradox and Discovery,
Oxford: Basil Blackwell, 1965.
- Ludwig Wittgenstein, Philosophical Investigations, trans. by
G.E.M. Anscombe, New York: Macmillan,
1965.