

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. 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 material submitted.

The following explanation of techniques is provided to help you understand markings or notations 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 through an image and duplicating adjacent pages to assure you of complete continuity.
2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB ROAD, ANN ARBOR, MI 48106
18 BEDFORD ROW, LONDON WC1R 4EJ, ENGLAND

8014970

KARP, BARRIE

PERSONS AND SELF-DECEPTION

City University of New York

PH.D.

1980

University
Microfilms
International

300 N. Zeeb Road, Ann Arbor, MI 48106

18 Bedford Row, London WC1R 4EJ, England

Copyright 1979

by

Karp, Barrie

All Rights Reserved

PERSONS AND SELF-DECEPTION

by

BARRIE KARP

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.

1979

© COPYRIGHT BY
BARRIE KARP
1979

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.

2/4/50
date

David Rosenthal
Chairman of Examining
Committee

2/4/50
date

Mabel C.
Executive Officer

Gerald Myers

Margaret Atherton

Robert Schwartz
Supervisory Committee

The City University of New York

ACKNOWLEDGMENTS

Thanks to my advisors: Professor Margaret Atherton, for helping me start this project; Professor Robert Schwartz for careful criticism; and Professor Gerald Myers for patient and generous guidance in helping me finish it.

Thanks to my parents, Corky and Lenny Karp, my brother and sister, Rick and Lisa, and to Jon Sholle, for their love and support.

For discussions, advice and encouragement, and for reading and commenting on the work at various stages, thanks to Daniel Shapiro, Nora Eisenberg Halper, Professor Virginia Held, Professor John A. Fisher of the University of Colorado at Boulder, and especially, Sherryl Feinstein.

And special thanks for their generosity, understanding and friendship to Diana Peck and David Press, Ann and John Wylie, and Suzanne Anker. Finally, special thanks to Dr. Janet Wolfe for her wisdom and encouragement.

CONTENTS

	Page
ACKNOWLEDGMENTS.....	iv
Chapter	
I. INTRODUCTION.....	1
Footnotes to Chapter I	
II. SELF-DECEPTION.....	13
Introduction	
Sincere Denial and Avowal; the Subject-Matter of Self-Deception	
Conflict Between What is Denied and What is Avowed	
Deep and Shallow Senses of Having Beliefs or Feelings	
The Concept of Conflict	
Conflict Due to Error Distinguished from Self- Deceptive Conflict	
Interest	
Purpose	
Conclusion	
Footnotes to Chapter II	
III. THE INTEGRATED THEORY.....	67
Introduction	
Fingarette: Persons Defined in Terms of Consistent Avowal	
Fingarette's View of Self-Deception	
The Integrated Theory and Self-Deception	
Frankfurt: Identification--the Internal and the External	
Frankfurt: Decisiveness	
Self-Deception	
Footnotes to Chapter III	
IV. THE FRAGMENTARY THEORY.....	107
Introduction	
George Herbert Mead and the Fragmentary Theory	
William James and the Fragmentary Theory	
James: the Material and Social "Selves"	
James: the Spiritual "Self"	

Chapter	Page
James: the "I" and the "Me"--Introductory Note	
The "I" and the "Me"	
James and Mead	
"Selves" or Attitudes	
How Policy Gets Made on the Fragmentary Theory	
The Weakness of the Fragmentary Theory	
Concluding Summary	
Footnotes to Chapter IV	
V. THE FRAGMENTARY THEORY, THE INTEGRATED THEORY, AND SELF-DECEPTION.....	202
The Motivation for Self-Deception	
Another Look at the Integrated Theory and the Conflict Involved in Self-Deception	
Summary	
VI. CONCLUDING REMARKS.....	228
BIBLIOGRAPHY.....	252

Chapter I

INTRODUCTION

This thesis is concerned with certain questions about conflict in the self and how self-deception bears upon these questions. The kind of conflict to which this thesis refers is that which involves inconsistent impulses, beliefs or feelings. Such conflict raises questions about the nature of conflict and the nature of the self. One question it raises is whether the self can be split and what this means. For example, phenomena that involve conflict are self-deception and akrasia. Self-deception has been referred to as a situation in which persons are at war with themselves (Fingarette), and akrasia has been described as a situation in which persons act in spite of themselves (Watson).¹ These kinds of descriptions and the phenomena they describe seem to suggest that persons can be inconsistent in the sense that they can do incompatible things at once. Persons act against themselves, or parts of them oppose other parts in such situations. This suggests that the self can be split.²

The next question that comes to mind is whether such fragmentation of selves differs from multiple selves,

and how such fragmentary entities can be understood to be single persons. Basically, such problems can be traced back to Plato and his division of the soul into parts along the lines of its different faculties. Conflict could then be explained by reference to these different parts of the soul --reason vs. desire, etc. But this strategy does not settle certain questions about the nature of persons, and what persons' attitudes are towards conflict. We remain tempted to ask, for example, whether persons are represented more truly by desire or by reason, or whether there is some value system over and above these, which is the true person.

Finally, the situation of conflict in the self raises a question about the attitudes persons have towards conflict. This thesis is focussed around this question, and the issues mentioned above will be taken up in order to elucidate this question. This question is simultaneously about the nature of conflict and the nature of persons. In finding out what kinds of attitudes towards conflict persons are capable of having, we arrive at a clearer picture of the nature of persons. For example, questions are raised such as whether persons are consistent systems or whether they constantly strive to become consistent systems. If the latter is the case, persons are identified with a value for consistency and all other values are subordinate to this value. A third possibility is that persons sometimes

strive for consistency, but, that they can have other attitudes towards conflict too, such as acceptance, cultivation, or even purposeful creation of it.

The first possibility, that persons are consistent systems, is ruled out empirically. We are only too familiar with inconsistency of all sorts in ourselves and others. The problem, then, is to understand persons via the problem of conflict. This question suggests two major kinds of theories about the nature of persons vis à vis conflict. One theory is what I call the integrated theory of the self. This theory puts forth the claim that persons always strive to eliminate conflict. Such a view is a likely suggestion for the solution to the multiple-self type of problem that conflict tends to raise. Perhaps conflict creates a kind of tension such that persons try to eliminate it. Tension exists because conflict pulls a single self in different directions. A single self is seen, in this view, as having a constant consistency impulse. In this view, then, anything that persons do in respect to conflict must ultimately be seen in terms of their constant motivation to eliminate it.

Another possible view about persons' attitudes towards conflict is represented by what I call the fragmentary theory of the self. In this view, persons can and do take a variety of attitudes towards conflict. They need not want to or try to eliminate it. A view

like this appeals to intuitions we might tend to have as the result of being impressed by the enormous role that conflict plays in our lives. The ubiquity of conflict suggests that conflict is natural to us and even that we do not object to it, we can accept it, create it, and cultivate it. The fragmentary theory also seems to do justice to the way in which the nature of the self is complicated by conflict, whereas there may seem to be some question about the integrated theory's ability to do this. The integrated theory tends to suggest that there is only one impulse with which persons are primarily identified, and that is their impulse for consistency and whatever consistent resolutions issue from it.

Thus, in that view, if persons are conflicted--say they want to sleep and to play, or they believe they are loved and that they are not loved, or they love and they hate--only one of these impulses, beliefs, or feelings of each pair (or some consistent transformation of all of them) is the true impulse, belief or feeling of these persons. The integrated theory-- via the posited consistency impulse-- tends to suggest a way of tying together apparent fragments of conflicted persons, but it also seems to do this in a way that eliminates the fact of conflict. Conflict seems to mean that persons are truly identified with inconsistent impulses, beliefs or feelings. But the integrated theory tends to suggest a monolithic understanding of the nature of persons, vis à vis the

assumed constant consistency impulse, that eliminates the phenomenon of conflict we were originally trying to explain.

For this reason, the fragmentary theory seems like a good alternative. Another reason in its favor is that the integrated theory tries to explain self-deception in terms of a consistency impulse. But it is not immediately clear how it can account for the fact that self-deception creates or perpetuates conflict, and that it thus seems to involve an attitude of acceptance of conflict. But the fragmentary theory raises other serious problems about the nature of conflict and the possibility of unity in the self. For, although it more emphatically preserves the idea about conflict that persons in conflict truly hold both inconsistent impulses, beliefs, or feelings of any pair, it calls into question the idea that conflict need be accompanied by tension and that such tensions lead persons to want to eliminate it. Thus, we are unclear about how we should construe "inconsistent" in the fragmentary theory, and why we should call "conflict" anything that occurs in persons that is not accompanied by such tension and impulses for consistency. So the fragmentary theory must clarify what conflict is.

Another problem, related to this one, is that of unity in the self. How is it that single persons can make policies if they are composed of these different

conflicted in a sense other than just having beliefs or feelings they do not want to have. The self-deceptive denial actually creates (or, in some cases, it perpetuates) more conflict in persons. In self-deception, persons believe or feel both what they deny and the denial (what they claim). Thus, they love and hate, or believe someone is guilty and not guilty, and so on.

It is also the view of this thesis that self-deception is purposeful. Self-deceivers have interests in denying what they believe or feel, and they do so because of these interests. Since self-deception involves conflict and is purposeful, we should find it useful to look at it in an effort to shed light on the nature of persons' attitudes towards and motivations with respect to conflict.

What we shall find in this thesis is that some of the ideas that seemed to be suggested by the integrated theory need not belong to it after all. The integrated theory need not claim that there is a true self or system of values, just because there is a constant impulse for consistency in persons. Thus, it need not eliminate the phenomenon of conflict when it attempts to explain it. It need not deny that conflict can exist in persons, in order to claim that persons have an aversion to it. It turns out that the integrated theory gives an acceptable explanation for the motivation for self-deception and that the fragmentary theory is only descriptive of

a certain phase of self-deception. We are unable to find a motivation for self-deception that could be suggested by the fragmentary theory. All possible motivations seem to appeal to a consistency impulse. The integrated theory explains not only the initial situation that gives rise to self-deception, but also the self-deceptive denial. Part of that denial is that it serves to mask from persons, not only that they have certain beliefs or feelings, but that they are conflicted about having them. Discomfort about being conflicted about one's beliefs or feelings is the likely explanation for self-deceptive denial. Otherwise, if one simply had beliefs or feelings that one did not want to have, one could accept this as an unfortunate situation. But the integrated theory posits that persons have an aversion to conflict and that this is what explains self-deception.

It has not been established, however, that the integrated theory is universally explanatory of persons' attitudes towards conflict. Technically, the integrated theory states that persons have a constant impulse for consistency. In showing that impulses for consistency explain self-deception we have not shown that this is the nature of conflict or of persons' attitudes towards it. Thus, it has not been shown that the integrated theory can explain some instances of acceptance or creation of conflict. But, on the other hand, since we have not been able to make sense of the notion of con-

flict without tension and aversion to it, the integrated theory might be implied in its full sense. Then, it would follow that acceptance of conflict is always a byproduct of not being able to resolve it and not a primary motivation. Persons might try to learn to accept conflict in spite of their aversion to it, and not just in spite of society's disapproval of it or to the influence of currently accepted notions of responsibility, as the fragmentary theory tried to suggest.

Another thing that has been clarified as the result of this inquiry is the nature of self-deception. We see the source of the misconception that self-deception involves impulses to create or perpetuate conflict and that it implies a fragmentary theory of the self.

In Chapter II, the elements of self-deception are described. There, it is argued that self-deception involves conflict and that the kind of conflict it involves is about the self, rather than being about something else. The significance of this is that in the kind of conflict involved, the beliefs, feelings, or impulses all belong to the person, even though the depth of identification may vary. Self-deception is distinguished from mere error in which persons do not have conflicting beliefs or feelings, or, where if they do, the denial that they do not is purposeful, and their identification with one of them is simply erroneous. Self-deception, although it is motivated by interests, is distinguished

from mere wishful thinking, in that it involves conflict and it is purposeful in a stronger sense. The interests wishful thinkers have in holding certain beliefs or feelings cause these beliefs or feelings unwittingly and mistakenly. But the role such interests play in self-deception is not simply error. In this chapter, a view of self-deception is presented that acknowledges the kind of conflict that initially seems to be more readily allowed by the fragmentary theory. But it is not clear whether the motivation for self-deception is that suggested by the integrated theory or that suggested by the fragmentary theory. Both seem to play a role and that seems paradoxical.

In Chapter III, the integrated theory is described. Some of the views about persons of Fingarette, Watson, and Frankfurt are presented along with some aspects of Fingarette's view of self-deception.³ These philosophers have views that are related to and that tend to suggest the integrated theory. Some of the problems of such views, if they are interpreted as integrated theories, are pointed out. They tend to suggest that conflict is not possible because they tend to rule out that dissonant beliefs or feelings could be identified with persons. The criteria they use for determining what motivates persons seem to be arbitrary. The result is that the question is raised as to whether the integrated theory can account for the kind of conflict involved in self-

deception and, since self-deception perpetuates and/or creates conflict, the motivation for self-deception. In the integrated theory, it seems that persons in self-deception are either not really represented by all the beliefs and feelings that conflict, and so self-deception does not involve the kind of conflict described in Chapter II, or self-deception is not purposeful. The integrated theory suggests a way that conflicted persons can be unified--via the consistency impulse. But it then must explain self-deception in terms of motivation for consistency, and this does not seem to jive with the claim of Chapter II, that self-deception creates or perpetuates conflict and thereby seems to involve attitudes of acceptance towards conflict on the part of persons.

In Chapter IV, we turn to a description of the fragmentary theory. There we describe the views of Mead and James at length because they provide a very full picture of what persons are like in terms of plurality and the likelihood of conflict.⁴ These philosophers are sensitive to the complexity of the self. James's notions of the "I" and the "me" are suggestive of a way of seeing unity in pluralistic, complex, and often conflicted persons. A possible answer to the problem of how policy gets made on a fragmentary theory is described. But ultimately, it is rejected, as it fails to finally show how persons can deal with conflict without being motivated by consistency impulses.

Finally, in Chapter V, the motivation for self-deception is again discussed. The fragmentary theory was shown to have difficulty in explaining persons' policies vis à vis conflict without referring to a consistency impulse, but it is not entitled to appeal to a consistency impulse. Earlier, it seemed that self-deception could not be explained by appeal to a consistency impulse. New ideas about the motivation for self-deception are explored at this point. It turns out that one thing that is needed in explaining self-deception is an account for the masking of beliefs or feelings from themselves that self-deceivers engage in. It appears that the masking can be explained by an appeal to persons' aversions to conflict. Thus, self-deception can be explained by the integrated theory after all. In this section, two different views about the motivation for self-deception are discussed. The fragmentary theory claims that persons in self-deception want to deny certain beliefs or feelings. The integrated theory claims that self-deceivers want to deny the conflict between these beliefs or feelings and the discomfort they have about these beliefs or feelings. It turns out that the idea of the integrated theory--that self-deception is motivated by an aversion to conflict--is acceptable after all.

Footnotes to Chapter I

¹Herbert Fingarette, Self-Deception (London: Routledge & Kegan Paul, 1969); Gary Watson, "Free Agency," Journal of Philosophy, LXXII:8 (April 24, 1975), 205-220.

²Herbert Fingarette, "Self-Deception and the 'Splitting of the Ego'," ed. Richard Wollheim, Sigmund Freud (New York: The Viking Press, 1971); Freud, 1940e.

³Fingarette, ibid.; Harry Frankfurt, "Freedom of the Will and the Concept of a Person," Journal of Philosophy, LXVIII:1 (January 14, 1971); Frankfurt, "Identification and Externality," ed. Amelie Okensberg Rorty, The Identities of Persons (Berkeley: University of California Press, 1976), 239-251; Watson, ibid.

⁴William James, The Principles of Psychology, Vol. I (New York: Dover Publications, 1950); George Herbert Mead, Mind, Self and Society, ed. Charles W. Morris (Chicago: University of Chicago Press, 1962).

Chapter II

SELF-DECEPTION

Introduction

The present view of self-deception is based upon personal experience and observation, analysis of opinions sought in conversation with colleagues and friends, and current philosophical literature on self-deception.¹ There seems to be a phenomenon, self-deception, which is described here. The present description seems to be suggested by the data. It may be that this is illusory and that self-deception is nothing more than error or rapidly alternating conflicting commitments. But the present view of self-deception is conceivable as well as being suggested by the phenomenon that will be described in the following discussion. Thus, this thesis concerns certain problems about the self vis à vis self-deception as described here, if there is such a thing.

The present discussion is a sketch of what will be explained more fully in the later sections of this chapter. When persons self-deceive, they purposefully, sincerely, and in a peculiar sense, partially erroneously, deny their beliefs or feelings, because of interests they have in not thinking they believe or feel what they

believe or feel. These interests grow out of persons' desires not to believe or feel what they believe or feel and their inability to eliminate the undesired beliefs or feelings. Self-deceptive denial is purposeful, not accidental. It is motivated by persons' interests, which are their reasons (and not just causes) for their denial. Such purpose rules out thinking of self-deception as being mere error. It is not an unwitting or innocent mistake. It is part of the projects of persons. Self-deception is not merely lying either, as the element of sincerity means that self-deceivers tell themselves the same things that they tell others-- that they do not believe or feel something that they do believe or feel.

(What is denied self-deceptively will be referred to throughout as beliefs or feelings. But what can be denied self-deceptively can include other facts about the persons involved, such as desires, attitudes, intentions.)

Another view about self-deception that will be defended in this chapter is that the denial it involves is simultaneously an avowal of beliefs or feelings that conflict with what is denied. For example, when the secretary denies that she hates her boss, she is claiming that she does not hate him. This denial-avowal is sincere, and, as argued below, because of this sincerity it is not simply false. The denial-avowal is either an expression of beliefs or feelings that the persons

had before self-deceptively denying what they deny (as in the case of those who both love and hate and this leads to self-deceptive denial) or it is an expression of beliefs or feelings that have been created by the denial and that are perpetuated by it. Self-deception, then, involves either the creation of conflict or the perpetuation of conflict.

An example of the first-mentioned kind of situation that can give rise to self-deception is that persons may both love and hate someone, while they also disapprove of having one of these feelings. They may wish they did not hate as well as love. Not wanting to hate and being unable to eliminate this feeling can lead to self-deception. In self-deception persons will sincerely deny that they hate and they will sincerely claim to love. In such cases, they continue to hate and to love. Thus, self-deception can perpetuate conflict.

In the simpler kind of situation that can give rise to self-deception, an example would be simply hating and not wanting to hate. If persons are unable to resolve this conflict they might self-deceptively deny that they hate and claim that they do not hate or even that they love. The avowal that one does not hate creates a new feeling.

In this chapter an attempt will be made to show that the result of both kinds of situations that give

rise to self-deception (when there is initial conflict of beliefs or feelings and self-deception perpetuates it; and when there is merely initial conflict between beliefs or feelings and a desire not to have them, and self-deception creates new beliefs or feelings that conflict with the denied ones) is that when persons are in self-deception, they have conflicting beliefs or feelings.

In some cases the self-deceptive denial-avowal (which conflicts with what is denied) may report beliefs or feelings that persons have only in a shallow sense. There is a sense in which persons really think of themselves as believing or feeling what they claim-- for example, that they do not hate--, and they exhibit some types of behavior that are relatively mechanical or that resemble acting. The sincerity and habitual aspect of the avowal makes it true of them, but only in a very shallow sense, as they merely "go through the motions" of believing or feeling it. In such cases, evidence that persons really have the denied beliefs or feelings is stronger. Thus, they may claim to love and may do some apparently considerate acts for the "loved" one, but these actions do not ring true. In addition, such actions might be far outweighed by aggressive behavior.

The passive-aggressive person is a good illustration. The passive aggressive person has the capacity to seem gentle, innocent, and even loving, doing things

for the benefit of others that are really destructive. The man who claims that it is out of a sense of honesty and respect that he tells his lover that he has met a new lover might be burdening his original lover with information that will cause unnecessary and avoidable pain. In the name of concern and respect he might be brutally honest. In such cases, the avowal that his intentions are constructive might be true in only a very shallow and nearly meaningless sense. One might argue that his honesty about his infidelity has some elements of true benefit and concern for the first lover-- she may now be able to plan her future and to gear her expectations more realistically. But if such prudence on her part is possible without her having to be told painful details, and particularly if she prefers not to have such knowledge and if her lover knows this, the evidence seems to be that he is expressing hostility more than genuine concern. Thus, this man's claim to be doing something that is considerate and not hostile might be a shallowly true claim.

The view presented so far of the status of the self-deceptive avowal is reminiscent of Fingarette's² remark that there is a shallow kind of truth to self-deceptive avowals. But he minimizes this point. He does not emphasize that self-deception involves having conflicting beliefs or feelings, perhaps because he is concerned with placing a higher value on a single consis-

tent set of beliefs or feelings. In his view, then, if persons claim to believe or feel something that conflicts with what he calls their "governing values," their claims, no matter how sincere, are basically false. The sense in which these claims might be true is, for him, not only shallow but irrelevant. In the position taken by this thesis, however, the moral judgments Fingarette makes about the status of individuals' commitments in self-deception are eliminated. The present view emphasizes that self-deceivers have conflicting beliefs or feelings.

The present view also emphasizes that the self-deceptive denial-avowal could also be true of persons in a deeper sense. In such cases, the denial creates a new belief or feeling (unless it is perpetuating a previous one) in a deeper sense. Thus, the person who hates can create loving feelings by avowing them sincerely and vigourously. In such cases the denied hostility may continue to operate, but it does not discount the truth of the avowed love. Feelings as extreme as love and hate often appear together. Another possibility is that newly avowed beliefs or feelings will transform persons until the self-deceptively denied ones no longer exist (in which case self-deception dissolves), or at least until both hold equally. Thus, as Rorty³ notes, self-deception might be an "adaptive" or "integrative strategy." This would also be consistent with

Fingarette's⁴ view that self-deception is motivated by integrity, although he does not approve of self-deception as a means to the end of self-transformation. This difference between Rorty and Fingarette is accompanied by the parallel difference that Rorty⁵ emphasizes that self-deception involves conflicting beliefs or feelings and Fingarette does not.

A final possibility is that the self-deceptive denial-avowal that is true of persons in the deeper sense does not transform them in that it does not eliminate the denied belief or feeling. Rather, it creates or perpetuates an additional, and conflicting, belief or feeling. The result is that there are conflicting deep beliefs or feelings. The distinction between the deep and shallow senses of having beliefs or feelings is discussed at greater length later in this chapter.

Thus, in summary, it will be argued in this chapter that self-deception can create or perpetuate conflict between deep and shallow beliefs or feelings and between deep feelings or beliefs. Self-deception can result in deeply true avowals. One's self-deceptive avowals no longer resemble role-playing or "going through the motions," as Fingarette suggests.⁶ That self-deception can create such conflict in persons is one of the more interesting features of self-deception from the point of view of the nature of the self. The sincerity and apparent double-mindedness of self-deceivers can make

us wonder about persons' capacities for and attitudes towards inner conflict. This is the aspect of self-deception for which we are seeking an explanatory theory of the self in this dissertation.⁷

To summarize, what will be argued in the following pages is that the source of self-deception is discomfort about having certain beliefs or feelings. The self-deceptive denial-avowal creates or perpetuates a claim that persons have conflicting beliefs or feelings. This avowal is true to one degree or another. Persons actually have the avowed beliefs or feelings. The depth of their having the beliefs or feelings may vary, but habitual sincere denial means that they really do feel or believe what they claim. Thus, self-deception can create conflicting beliefs or feelings in the self, and not always just in the weak sense of creating shallow beliefs or feelings. But even when it is just in the weak sense, the significance of this in the present view differs from Fingarette's view.

In the case of initial discomfort involving conflicting beliefs or feelings, the self-deceptive denial-avowal perpetuates conflict. The avowal in this case is true (again, the depth can vary), but, as in all cases of self-deception, it is incomplete because it misrepresents the person as being single-minded, as not also having the denied beliefs or feelings.

The underlying condition for all of the conditions of self-deception is that self-deception is by and of the same self. Otherwise, it would not be self-deception. Self-deception of the sort that has just been described involves the following central features:

- 1) Sincere denial and avowal: what persons sincerely and not fully erroneously claim to believe or feel involves or is a denial of something else that they believe or feel.
- 2) Interest: persons have interests in denying that they believe or feel what they do in fact believe or feel (thus, they also may have an interest in denying the evidence for what they believe or facts that would justify their feeling; this interest is derivative).
- 3) Purpose: self-deception (what is described in conditions 1) and 2)) is done by persons and it is not mere error or wishful thinking.

Sincere Denial and Avowal: The Subject-Matter of Self-Deception

The view presented here is that in all cases of self-deception what one is deceived about is something about oneself (namely, whether or not one has a certain belief or feeling), not something about something or someone else. This is one of the conditions that distinguishes self-deception from error. This view is also held by other writers on self-deception. Rorty says that a bona fide case of self-deception--i.e., a case that is clearly distinguishable "from being

mistaken about one's beliefs, or from being conflicted about them. . . involves deception of the self by the self, about the self, for the sake of the self."⁸ If there is any such thing as self-deception, it must be by and of the same self. Otherwise, it would be deception of others or lying.

That self-deception is about the self is connected with its being for the sake of the self. What is at stake is some interest persons have in not having the denied beliefs or feelings.⁹ Belief due to an interest in what is believed being the case is not necessarily indicative of self-deception, but when the interest is in having or not having a certain belief (and when one already has a conflicting belief), it is self-deception.¹⁰ In the case of belief due to interest, a mother may not want to believe that her son is dead because she does not want him to be dead. She does not want him to be dead out of an interest--but it is not necessarily an interest in having the belief that he is alive. But in self-deception, she believes he is dead and she deceives herself about her belief. Thus, her interest in his not being dead, or her reaction to believing he is dead (painful attitudes caused by thinking he is dead), leads to her interest in not thinking he is dead. This latter interest motivates self-deception. It may be partially for the sake of others that she wishes he were not dead, but it's fully for her own sake--to

save herself pain--that she wishes not to think of him as dead. Thus, the interest persons have in having-- or in not having--certain beliefs leads them to self-deceptively deny their beliefs. Persons also might have an interest in transforming themselves. Thus, if they hate their boss but wish they liked him, their interest in liking him might lead to self-deception, whereby they create the desired feeling, possibly with the aim of phasing out the undesired one.

Not all writers agree that self-deception is always about the self. Thus, Gardiner is critical of the view, which he traces back to Butler, and he gives examples in which people deceive themselves about "how things stand in the world."¹¹ (Butler's view is also rather narrow in that he sees self-deception as always involving one's presenting oneself to oneself in a better light than is true. In my view, what is required is that one presents oneself to oneself in a different but not necessarily better way than the way one believes oneself to be.) But the cases to which Gardiner refers are not significantly not-about-the-self. The characteristic result of self-deception is that it creates and protects one's view of oneself as believing or feeling what one avows without conflict. We end up thinking of ourselves as if we were single-mindedly denying what we deny. We do not want to admit to ourselves that we also believe or feel what we deny.

This is the reason we might deny facts about the world that we seem to be in a position to acknowledge. Such denial is not part of self-deception unless there is also denial about oneself. In ordinary wishful thinking we erroneously deny evidence for certain beliefs because our vision is clouded by our desire that something be the case. But wishful thinkers' denials are not based on denials that they have certain beliefs that they do in fact have. Self-deception involves having a belief or feeling that one wants to deny, whereas wishful thinking just involves wanting to have a certain belief or feeling. Thus, in self-deception we ultimately deny that we believe something, not just that the belief is true. So-called examples of "deceiving oneself" about external facts--say about someone else's affections or abilities--are ultimately examples of self-deception because they are about oneself. If I self-deceivingly think someone loves me this means that I am trying to present myself to myself as believing or feeling that I am loved. The denial it involves that is essential to self-deception is a denial of my belief or feeling that I am not loved. Thus, I am deceiving myself about my belief or feeling that I am not loved. My avowed belief that I am loved is self-deceptive because it purports that this is my only belief, that I am not in conflict about it. (This point will be explained further, below in the discussion

of self-deceptive beliefs and feelings). Denying the external fact and accompanying evidence that I am not loved is a consequence of my deceiving myself about myself. This is what is meant by the present claim that this and all other cases of self-deception are about the self: they are about the self in the sense that they are about whether or not one has a certain belief or feeling.

The point that self-deception is about the self is easier to see in examples in which persons are deceived about what they feel, rather than what they believe. The belief case can be confusing because it usually includes person's denying both beliefs about themselves and about external circumstances. Consider the following supposed example of self-deception given by Rorty:

Mrs. Faith Beloyal, having been presented with evidence that her son, Willie, has given 9-year-old Miranda an apple with a razor blade stuck in it, stands firmly behind her son, her hand on his shoulder, and says: 'Willie is a good boy; he would not willingly or knowingly harm anyone.' She suggests that perhaps when Willie gave Miranda the apple he wasn't aware of the fact that his roommate, Archie Angel was in the practice of storing his used razor blades in apples, and then carelessly leaving them on the windowsill. Or she says that Willie prepared the razored apple as part of his assignment for a course, 'Form and Symbol in Modern Art', and that he asked Miranda to deliver the apple to his instructor, believing she understood the errand and was capable of running it. Before attributing self-deception to Mrs. Faith Beloyal, we would want to know whether her attempts to discredit the evidence were reasonable, and even if they were, whether she raised them for their reasonableness or whether she'd have used any doubts she would have laid her mind on.¹²

Rorty claims that if a person were to give wildly imaginative counter-explanations for evidence that something is true about the world--in this case, that Willie is guilty--we might think the person is being unreasonable. But, Rorty thinks, we would not suspect self-deception unless we had evidence that the person would "have used any doubts she could have laid her mind on."

The trouble with this view and example is that what such persons are denying is not clearly presented as being about themselves, and evidence that they would use any doubts they might think of could be evidence of stubbornness or deep conviction, and not of self-deception. For self-deception we need something more--we need evidence that Mrs. Beloyal denies the evidence for her son's guilt because she denies having the belief that he is guilty. If there is evidence that she believes what she denies in the core denial--that she believes that he is guilty--there is evidence for self-deception. But the fact that she has evidence that he is guilty is not enough to establish self-deception. If Mrs. Beloyal refuses to take apples from Willie or to let anyone else near his apples, this, in conjunction with her unreasonable denials of evidence for his guilt, would suggest that she is in self-deception. Her denials about her son's guilt would then be the result of her denial that she believes he is guilty.

This example is meant to illustrate the ambiguity that might arise when self-deception is about beliefs rather than about feelings. Denials of beliefs about the world might accompany such cases of self-deception or be reason to suspect self-deception. Another problem is that it might be difficult to determine whether a particular case of denial that persons have certain beliefs is self-deception. We need evidence that they really believe what they deny. Even in the case of Mrs. Beloyal's not allowing anyone to take apples from Willie, it is conceivable that there is some reason other than that she thinks he is guilty. But it might also be the most plausible interpretation to assume that this is evidence that she thinks he is guilty. The point of this discussion is to point out that cases of self-deception about beliefs can be difficult to identify.

In the example of being deceived about one's feelings it is easier to avoid ambiguity. Take the case of being deceived about whether one loves or has affection for someone. For example, if a person is in an unfair situation, in which one would expect her to be angry and to rebel, and instead she justifies the conditions to which she is subject, makes excuses for her oppressors, and describes positive features that do not exist, she may be engaging in self-deception. Take the example of a woman who has been mistreated by her employer. She will probably try to strengthen her denial of her

hostility by denying external facts--evidence that there is reason to be hostile (his unfair treatment of her, etc.)--but the essential self-deceptive denial is about herself. The important point is not her denial of any evidence that justifies the view that anger or hate is an appropriate response to him. This in itself does not indicate self-deception. What indicates self-deception is evidence that she feels what she denies--such evidence would be strong if, accompanying these avowals of hers is some very hostile behavior towards him. What she is really denying is that she feels angry. The other denials are made because of this essential denial. In the same way, when self-deception is about a belief, it is about a belief we have about ourselves.

Sincerity

It is important to emphasize that self-deceivers' denials and avowals are sincere. If they were not, it would be impossible to distinguish self-deception from lying or deception of others.¹³ Self-deceivers really think they believe or feel what they claim to believe or feel (and they really do believe or feel it-- this claim will be substantiated below), and they really think they do not believe or feel what they deny. Thus, the woman who denies hating her boss really thinks she does not hate him. She may mean by this that she has no

feelings at all about him, or that she likes him. But still, she also hates him, as her other behavior reveals. If Mrs. Faith Beloyal is a self-deceiver, she sincerely thinks that she believes that Willie is innocent, and this is consistent with her denial that she thinks he is guilty and that he is guilty. But, as her behavior reveals, when she refuses to take apples from him, she also might believe he is guilty. Self-deceivers are not simply liars even though self-deceivers also do not believe or feel what they claim to believe.¹⁴

This feature, which has been taken to be one of the paradoxes of self-deception, is referred to by Fingarette as "insincere sincerity."¹⁵ Fingarette points out that a paradox of self-deception is that self-deceivers appear to be sincere from the perspective of their denial-avowal, but other behavior is evidence for their believing or feeling what they deny and from this perspective they seem insincere. Fingarette explains that one can be sincere and irresponsible. Sincere statements do not require behavior consistent with them to be sincere. So persons may sincerely deny that they feel hate and act as if they feel hate. Fingarette thinks this kind of sincerity is shallow-- but it still is sincerity; and that to the extent that this is taken to be "sincere insincerity" it is taken in everyday terms to be a kind of self-deception. But, he says, it is different from the more complex kind of occurrence that we are taking to be self-deception.

Shallow sincerity occurs when strong emotions exaggerate persons' commitments. This can happen, for example, due to the influence of drugs. The more complex kind of occurrence--self-deception--is purposeful. And deep insincerity is lying. Fingarette sets out three conditions to clarify the distinctions between shallow sincerity, lying and self-deception:

(1) that there is no intentional difference between what individuals tell themselves and what they tell others

(2) that what individuals tell themselves correctly reflects the situation

(3) that individuals come to make these expressions purposefully, even if they are inaccurate, in which case the individuals would be purposefully wrong.

If conditions (2) and (3) hold but not (1), lies are the result. If condition (1) alone holds, but not (2) and (3), we have shallow sincerity. Finally, if we have conditions (1) and (3), but not (2), we have self-deception. The paradox of self-deception is that the more we have (3) but not (2), the more we see self-deceivers as insincere. The more we emphasize (1), the more we see them as sincere. So, the more we look at the denial as being purposeful and erroneous, the more persons seem to be liars. And insofar as we look at the denial as purposeful and unintentionally erroneous we see them as self-deceivers.

Conflict between What is Denied
And What is Avowed

The view of this thesis is that the self-deceptive denial-avowal creates conflicting beliefs or feelings, or it perpetuates them. Although self-deceivers are mistaken in claiming not to have the denied beliefs or feelings, there is a sense in which they are also not mistaken: because they also have conflicting beliefs or feelings and this is what they are saying when they deny that they have the original beliefs or feelings. Thus, when persons self-deceptively deny that they hate their boss, they sincerely claim not to hate him. The present interpretation is that this sincere claim is also true at least to some degree. Thus, in the preceding section on sincerity, where it was said that condition (2) (that what self-deceivers tell themselves correctly reflects the situation) does not hold in self-deception, what was meant was that what self-deceivers tell themselves is incorrect because it is incomplete.

The view with which this interpretation takes issue is the view that what is claimed by self-deceivers is claimed falsely and is untrue of them. In this view, self-deception involves denying one's true beliefs or feelings and claiming to have beliefs or feelings that one does not have. But the interpretation preferred here is that the situation is more complex. There is a sense in which the claim is false, but it is more accurate

to say that it is incomplete, by which it is meant that self-deceivers have the beliefs or feeling that they claim, but that they also have those that they deny.

This interpretation does justice to the experience of being in self-deception, in which one seems to oneself to have the beliefs or feelings that one claims. Otherwise, self-deception would be no different from simple lying.¹⁶ It also does justice to the observer's view of self-deceivers. From the observer's perspective self-deceivers are puzzling because they seem to have conflicting beliefs or feelings. They give evidence of sincere ascription to conflicting beliefs or feelings. Their behavior and history are testimonies to their having the beliefs or feelings they deny and in this respect the sincere denial seems false. From this perspective, sincerity does not lead us to think that what is claimed must be true. But at the same time, from another perspective, this very same sincerity does tempt us to think that self-deceivers do not have the denied beliefs or feelings (i.e., that they have the beliefs or feelings they claim). One way to eliminate this paradox is to posit that self-deceivers have both beliefs or feelings.

The obvious objection at this point is that the sincerity of a claim is not enough to establish its truth. Indeed, from one perspective the sincerity of self-deceivers' claims that they do not hold the denied beliefs or feelings is overridden by evidence that they have them.

Such evidence is taken to be decisive. Thus, it would seem to follow that such a claim--that one does not have the self-deceptively-denied beliefs or feelings--is simply mistaken. However, the present argument is that a sincere claim about one's own beliefs or feelings cannot be fully mistaken, although it can be incomplete. A sincere claim about something other than one's own beliefs or feelings can be erroneous. This is what error is: sincere but faulty thinking. But sincere error about one's own beliefs or feelings can be self-deception because it is purposeful. When the reason that persons are incorrect in claiming to have certain beliefs or feelings is that they have conflicting ones too (not just that they have conflicting ones instead), that they have an interest in the denial-avowal, and that this interest is motivating the "error," we have self-deception. When the error is about something other than one's own beliefs or feelings, what one claims is simply false. But when the error is about oneself, what one claims is incomplete, and there is still some truth to one's claims not to have the denied beliefs or feelings.

One reason a sincere claim about one's own beliefs or feelings cannot be fully erroneous is that it is possible that the sincerity of such claims means that one thinks of oneself as having these beliefs or feelings and that such thoughts will cause one to cultivate having them. Part of having a belief or feeling can be

sincerely thinking one has it. If one really wants to have a certain belief or feeling, a good way to cultivate it is to claim it and to start acting as if one has the belief or feeling. The degree to which one has the belief or feeling can vary, just as the depth of self-deception and the amount one gives elaborate defenses against acknowledging that one also has the denied belief or feeling can vary. This interpretation is substantiated by the psychoanalytic interpretation of ambivalence, in which persons often experience themselves as having alternating beliefs or feelings. (Ambivalence also includes experiencing oneself with conflicting beliefs or feelings at once). The psychoanalytic view is that all of the beliefs or feelings belong to persons, none of them are claimed falsely, although persons' experiences or claims are incomplete if they do not acknowledge both beliefs or feelings of any pair of conflicting ones.¹⁷

Another objection that can be raised against the present argument is that it rules out the possibility of genuine error due to ignorance about oneself. Could it not be that one is just mistaken about one's own beliefs or feelings? Error about one's own beliefs or feelings is possible but when it is purposeful it is self-deception. When persons have interests in such "error" and those interests motivate the "error," we have self-deception, and it follows that persons have the avowed beliefs or feelings. In cases of true (not self-deceptive)

error, persons do not really have the beliefs or feelings even though they sincerely think they do. The difference is that sincerely thinking one has certain beliefs or feelings will lead one to act like one does. But if one is so ignorant as to be mistaken about having the beliefs or feelings, one must not know much about them and what it's like to have them, and one will therefore not know how to act or feel like one has these beliefs or feelings. But self-deceivers, whose claims to have the beliefs or feelings are not based on ignorance, but are purposeful, know what kind of behavior the beliefs or feelings entail, how to express such beliefs or feelings, and how to experience them. In this case, sincerity means persons have the beliefs or feelings. In the case of genuine error, it does not really mean this, or it means it in an indirect, differently incomplete way. A sufficient condition for having a belief or feeling is present in a very empty way. But in self-deception one's sincere avowal can have some content, due to the element of purpose.

It follows that the self-deceptive denial-avowal can create a conflicting belief or feeling: it is possible that the avowal of self-deception is not simply a false avowal. We shall say, then, that self-deception involves conflict of beliefs or feelings.

Deep and Shallow Senses of Having
Beliefs or Feelings

The present view is that: beliefs and feelings that belong to persons in a shallow sense ought not to be regarded as falsely attributable to persons or as not truly representing them; that such beliefs or feelings may come to belong to persons in the deep sense; and that what self-deceivers claim to believe or feel can be true of them in either the shallow or deep sense. These claims support the present view of self-deception, according to which self-deception involves persons having conflicting beliefs or feelings. This view contrasts with views that see self-deception as involving falsely claimed beliefs or feelings that do not represent such persons. The most important difference between these two views is that in the present view many cases of self-deception that are more than momentary involve claimed beliefs or feelings that are more than shallow, whereas the opposing view sees what is claimed by self-deceivers as being shallow at most and not representative of them. And in these views, when persons have shallow beliefs or feelings they are just acting. But in the present view, shallow beliefs or feelings are more than that and self-deception can involve conflicting beliefs or feelings that are either shallow or deep. In the following discussion the distinction between the deep and shallow senses of having a belief or feeling will be discussed.

Persons have beliefs or feelings in a deep sense if they give evidence of understanding what it is to have such beliefs or feelings and if they can display behavior of a wide range that indicates having such beliefs or feelings. The more such evidence for their having these beliefs or feelings seems sincere and does not seem to be play-acting, the greater likelihood that the beliefs or feelings are deep. And evidence of having conflicting beliefs or feelings is not necessarily evidence that the original beliefs or feelings are not had in a deep sense. In fact, it might be evidence to the contrary, and this is one way that self-deception can be distinguished from wishful thinking. Wishful thinking does not involve conflict and the wishful thinker's claim to believe or feel something is mistaken. But self-deceivers are "mistaken" in what they claim to believe or feel because they claim those beliefs or feelings only, to the exclusion of conflicting beliefs or feelings that they also have. To deny this is to confuse shallow belief or feeling with false claims to believe or feel.

Beliefs or feelings are had in a shallow sense when there is little of the whole range of appropriate evidence available that persons believe or feel the belief or feeling in question. In such cases persons may seem to be only play acting, but they may also seem to be sincere in making their claim. But they display

a lack of the full range of either understanding what it is to have that feeling or belief, and/or behavior that would indicate such a belief or feeling. Behavior that seems perfunctory or easily distracted by conflicting demands, or both inappropriate and counterbalanced by conflicting behavior (as in the case of the man with two lovers discussed earlier) can be evidence of shallow beliefs or feelings.

This distinction is difficult to make because it is tempting to rely on conflict as evidence for shallowness and lack of conflict for deepness, and it is tempting to confuse shallow beliefs or feelings with false beliefs or feelings. But making conflict the indicator that a belief or feeling is shallow is not always permissible because what we are going to claim is that one can have either deep or shallow beliefs or feelings and still be in conflict about them. Conflict cannot be relied upon to make this distinction. We want to deny the claim that necessarily the more persons are in conflict the more shallow are their beliefs or feelings. For the same reason, we cannot make the deep-shallow distinction in terms of sincerity, where sincerity is construed in terms of absence of conflict. Perhaps the best we can do in making this distinction is to cite the range of behavior and extent to which this behavior shows understanding of what it means to have the beliefs or feelings. Referring to range and extent is tricky,

because these concepts bring to mind the concepts deep and shallow, and this suggests that our explanation of the distinction is circular. Given the difficulty of making this distinction, the strongest claim we can make is that by referring to the range of behavior and extent of understanding we are spelling out what is meant by deep and shallow senses of having a belief or feeling.

Given the difficulties of making this distinction, the distinction will not be very precise. But it is basically meant to mark a difference of status between certain of one's beliefs and feelings. Deep beliefs or feelings are better established and more permanent. They have more of a history in persons. Shallow beliefs or feelings are either new or only partially formed and established. They have not become habits and it may take a special effort to call them into play. Thus, self-deceivers start out with beliefs or feelings that are deep and a desire to eliminate them that also may be deep. The self-deceptive denial is not deep, though. It's new, it takes effort, it can be seen as a contrivance. It's a first step towards developing deep beliefs or feelings, a first step in self-transformation or self-elaboration. As self-transformation, it takes the form of creating new beliefs or feelings that will replace the old ones. As self-elaboration, it takes the form of creating additional beliefs or feelings.

It may be thought that in the case of belief it does not make sense to say that the self-deceptive belief is deep in the sense of being grounded in persons' histories. In the case of the mother who believes her son is dead and wants to deny that she believes it, the belief that her son is dead is not necessarily part of her history, in the usual sense of the concept of history. But what is meant when referring to persons' histories in defining the deep sense of belief is that these are beliefs that we should reasonably expect them to have, given certain facts about them--their intelligence, levels and capacities of awareness, rationality and objectivity, and so on.

Looking at some examples might help to clarify the distinction between deep and shallow beliefs or feelings. First look at the distinction between being in error about whether or not one has a certain belief or feeling and having the belief or feeling in the shallow sense. Take the case of wishful thinkers who mistakenly think that they believe or feel what they want to believe or feel, but who do not know what it is like to have that belief or feeling. Their lack of knowledge on this score is evident from the inappropriate behavior that they claim to be expressions of these beliefs or feelings. The reason we know they are mistaken is that although their claims are sincere, they give evidence that they do not know what it means to believe or feel what they

claim. Suppose such a person does not want to be jealous. She knows what being jealous is like, but she does not know what not being jealous is like. Yet her desire not to be jealous leads her to think that she is not jealous. But she gives no evidence of knowing what not being jealous is like. The only evidence that she's not jealous is that she claims not to be. But this is not enough.

In some ways such a woman is similar to persons who claim not to be jealous and who are non-jealous only in a shallow sense. Non-jealous persons in the shallow sense do some non-jealous things and make attempts to be non-jealous. But they may show a partial lack of understanding of what it is to be non-jealous. They may also show a partial commitment, whatever their level of understanding (even if they do know what it is to be non-jealous) to jealous behavior. Such conflict is not sufficient for having a shallow belief or feeling. There can be conflict between deep beliefs or feelings too. In such cases persons behave jealously and non-jealously, but insofar as they behave jealously, they do so in a wider range of situations and with a greater variety of behavior than they are capable of in terms of non-jealous behavior. So they are non-jealous in a shallow sense. And insofar as they behave non-jealously, they do so in a wider range of situations and with greater variety of behavior than the woman who mistakenly thinks she is non-jealous.

Believing or feeling in a deep sense, then, is, as just stated, exhibiting a wide range of behaviors and a more thorough understanding of what is required of persons who have such beliefs or feelings. Belief or feeling in the deep sense requires more than just an attempt to be non-jealous. The person in the shallow state exhibits such attempts. And, of course, belief or feeling in the deep sense is more than acting in ways that one might mistakenly think are non-jealous.

Being mistaken about whether one has a belief or feeling, and having a belief or feeling in a deep or shallow sense, are all characterized by sincerity. But being merely mistaken means exhibiting no behavior that is indicative of being in the state beyond one's sincere avowal of it. Being in the state in the shallow sense means having only partial understanding and/or exhibiting only partial behavior. The deep sense requires more full understanding and more appropriate behavior. Whether the understanding and behavior are appropriate for the deep or shallow senses often depends upon the manner of their expression.

The distinction between deep and shallow beliefs or feelings brings to mind Aristotle's morally weak and morally strong man and the *sophron*.¹⁸ Take the case of persons who control their jealousy and act non-jealously. Are such persons, whom Aristotle would call morally strong, non-jealous in a shallow or deep sense?

One might think the morally strong person is non-jealous in a shallow sense because she is less deeply non-jealous than the sophron for whom there is no conflict. But on my definition, the morally strong man could be non-jealous in a deep sense. Continuing conflict would not rule out having beliefs or feelings in the deep sense.

In the present view there is another category --being in the state in a shallow sense--which is between the morally weak and the morally strong categories. On the road to becoming morally strong, one can have beliefs or feelings in a shallow sense. (The morally strong person can have shallow beliefs or feelings too). This development from shallow to deep states may not be a process of developing consistency, as the integrated theory claims (this is explained in the next chapter). A possible implication of the integrated theory is that beliefs or feelings are more representative of persons the more they are consistent with their other beliefs or feelings.

We start out wanting to say that the beliefs or feelings self-deceivers deny are deep. They fully have these feelings or beliefs. Then we have sincere counter avowals. Self-deceivers seem to have conflicting beliefs or feelings. Does this detract from the depth of the denied state? The secretary acts angrily, but claims not to be angry too. Insofar as she acts as if she is not

angry, it seems as if her anger is shallow and vice versa. This shows that both feelings can be deep, even though they conflict. Each one can seem shallow from the point of view of the other. In addition, whether a belief or feeling is deep or shallow can shift depending upon whether one chooses to encourage and cultivate it, to transform or eliminate it, or to replace it or to complement it with conflicting beliefs or feelings.

In conclusion, the extent of persons' nonmistaken commitment to beliefs or feelings through appropriate, various, and innovative behavior, indicates whether the beliefs or feelings are deep or shallow. Conflict is not necessarily a mark of this distinction, although at times it is suggestive of it. But we must look beyond conflict to the other signs of commitment, or lack of it.

The Concept of Conflict

We have made the claim that self-deception involves persons having conflicting beliefs or feelings. What follows is a discussion of what is meant by conflict.

It is difficult to be clear about what is meant by conflict of beliefs or feelings at this point. What is meant by conflict depends upon a theory of the self which would indicate persons' attitudes towards conflict. But the problem of this thesis is to find a theory of the self that explains self-deception. Thus, just

what is the meaning and nature of conflict in self-deception is not entirely explicable at this point. There are two possible attitudes towards conflict that are considered in this thesis. One view is that persons can accept conflict, create or cultivate it, and that there is no inherent impulse involved in it to eliminate it. This is the view of the fragmentary theory of the self, which will be discussed in Chapter IV. The other view, that of the integrated theory, which is described in Chapter III, is that one thing that conflict means is that there is an impulse to eliminate it. It's a situation of tension and persons are not content until it is eliminated. From this perspective, what the fragmentary theory calls conflict is impossible.

The best we can say is that conflicting beliefs or feelings are those beliefs or feelings that the integrated view would have persons trying to make consistent. And that they are those beliefs or feelings that the fragmentary theory would say are different attitudes about the same situation. We might be tempted to say that these different attitudes might manifest mutually exclusive behavior. But this is not an adequate way to characterize conflict, since much behavior is mutually exclusive: at the same time one can behave only one way.

A better tack might be to say that some conflict involves felt tension between beliefs or feelings, and that it is an open question, depending upon which theory

of the self is true, as to whether such tension leads to the impulse to eliminate it. Sometimes conflict involves wanting to do different things. Sometimes it is manifested more behaviorially in the form of patterns of which persons may not be aware.

In general, then, a symptom of conflict can be tension of various kinds. Persons can have conflicting impulses, in which case there is tension because they cannot do more than one thing at the same time. Or persons can have conflicting impulses, feelings or beliefs that create behavioral tensions. Such tensions might be less clearly felt at certain times. Such conflicts might be manifested in parapraxes, as pointed out by Freud, and by ambivalent behavior.

We now turn to further remarks about how the conflict involved in self-deception differs from erroneous conflict.

Conflict Due to Error Distinguished
from Self-Deceptive Conflict

Although self-deception is an instance of holding conflicting beliefs or feelings, it is important to note that it differs from other cases of conflicting beliefs or feelings, whether they are due to error or not. One difference is that when persons hold conflicting beliefs they may sincerely claim to believe both of them, both in cases where they both know and in cases where they do not know that they are incompatible.

If one's holding of conflicting beliefs or feelings is due to error, one might claim both beliefs or feelings, mistakenly thinking that they are compatible. Thus, if persons believe that the ice is safe and that it is about to melt they may be holding incompatible beliefs mistakenly because they do not know that the beliefs are incompatible. This holds for beliefs about one's own beliefs or feelings too. If persons believe both that they think the ice is safe and that they think no one should go near it, they may be mistakenly holding conflicting beliefs about themselves, because of an error about whether these beliefs are compatible. In addition, persons may have conflicting beliefs and acknowledge them both and that they conflict. These claims hold for conflicting feelings too. Loving someone can involve wanting the best for them while also wanting to consume, enslave, or control them. One may be unaware that these feelings conflict, or one may acknowledge that they do and that one has them. Neither the condition due to error about whether beliefs or feelings are compatible, (even when such beliefs or feelings are about oneself), nor acknowledged conflict are possible descriptions of self-deception.

In self-deception, persons must not claim the denied beliefs or feelings. If they did, self-deception would be unnecessary and impossible. Part of self-deceiving is concealing from oneself the fact that one believes

or feels, in addition to the avowed beliefs or feelings, exactly that which one denies. Unlike persons who hold conflicting beliefs due to ignorance, and unlike persons who acknowledge having conflicting beliefs or feelings, self-deceivers end up not taking responsibility for all of their beliefs or feelings. Self-deceivers know that such beliefs or feelings conflict and that it is not possible to claim both of them and to claim to be consistent at the same time. The self-deception consists of persons' maintaining both beliefs or feelings without resolving the conflict between them, while they sincerely deny that they have one of them. The result is their not acknowledging that they are in conflict. Self-deceivers' avowals are presented as if the conflict has been resolved, as if self-deceivers have only one belief or feeling. A woman who is angry at her boss might feel both nonhate and hate for him, or hate and a desire not to hate, but if she avows having both feelings of any pair of conflicting feelings, she has no need to self-deceive. Her self-deception enables her to express hate without thinking of herself as doing so. But if she claims both feelings, she must take responsibility for them and she cannot be in self-deception. Thus, self-deception is different from error or acknowledgment of conflict in that it cannot involve avowal of conflict, or of the denied belief or feeling, even though it must be based upon some kind of initial rec-

ognition of what is denied and that there is conflict.

Self-deceivers know that they have the beliefs or feelings they deny. They have to know this in order to deny them.²⁰ As James points out in a very different context (when he writes about the selective awareness of persons experiencing the phenomena of "blind spots"), we might by analogy describe persons in self-deception as giving evidence of carefully blocking out of present awareness just certain things, and this seems to require that they initially be aware of them. James explains in the case of blind spots that persons are not blind to the kind of thing they are looking at, but to this particular one, and that this requires that they initially recognize it:

Paradoxical as it may seem. . . , he must distinguish it with great accuracy from others like it, in order to remain blind to it when the others are brought near. He disconnects it, as a preliminary to not seeing it at all. We have then, to deal in these cases neither with a blindness of the eye itself, nor with a mere failure to notice, but with something much more complex; namely, an active count-int out and positive expulsion of certain objects.²¹

In self-deception persons have interests in not having the denied beliefs or feelings and this is what motivates their sincerely denying them. Thus, they must also have been able to take them into account. But there is no sense in which persons who are unaware due to error that they have conflicting beliefs or feelings, really know they have such beliefs or feelings.

The following objection might be put forth: that there are cases in which persons who hold conflicting beliefs or feelings due to error claim to believe or feel only one of them and know that believing both of them is inconsistent, and that in these cases the mistaken persons are indistinguishable from self-deceivers. But self-deception cannot be taken to be mere error. In the case of error, persons do not know that they really believe or feel what they deny. But in self-deception persons must know that they believe or feel what they deny, even though they sincerely deny that they have these beliefs or feelings. In error there is sincere denial, and persons do not know in any sense that they really have the beliefs or feelings. But self-deceivers' knowledge is their reason for denying the beliefs or feelings.

One way that the difference between self-deception and error might manifest itself would be that persons who mistakenly claim only one of their beliefs or feelings might acknowledge the other ones when demonstration makes it clear that they hold them. Upon being confronted with the appropriate evidence that they believe or feel what they deny, it is possible that they will recognize their beliefs or feelings. For example, the man who denies that he thinks the ice is unsafe and who also takes great pains to avoid it and to keep others from it, should admit, when having his attention drawn

to his behavior, that it reveals that he believes the ice is unsafe. If he acknowledges evidence that he believes what he denies, it could be that he is not in self-deception, but simply in error. If this man is deceiving himself, such evidence will make him admit that he thinks the ice is unsafe only if it is sufficient to break the spell of being in self-deception.

This point about mistaken persons being open to acknowledging evidence for what they believe or feel is not a sufficient way of distinguishing self-deception from error. This is due to the facts that (1) if persons acknowledge the evidence that this is their belief or feeling, it does not mean they were not in self-deception. Maybe this is the end of the self-deception. Thus, acknowledgment of conflict or of what one denies at a certain point is not a sure mark of having been in error. It could happen at the end of self-deception too. And, (2) there are conceivable non-self-deceptive cases in which persons do not acknowledge the evidence that they believe or feel what they believe or feel. Persons might hold a belief that they do not believe that poor people are inferior. This belief could be so deeply entrenched that such persons would provide what seemed to them plausible counter-explanations for any evidence to the effect that they do not actually hold this belief (and that, instead, they believe the poor to be inferior). It may be that in such a case we would think it likely

that such persons are deceiving themselves, but we would not yet have cited all the conditions that make it so, for we have not distinguished this case from being in stubborn error about one's beliefs.

The present claim is that it is conceptually possible that the following conditions hold: persons believe or feel what they deny and avow, but only believe that they believe or feel what they avow; and they know that the denied and avowed beliefs or feelings are incompatible. But such conditions could be evidence for self-deception too. It may be that all cases of such "error" about oneself are cases of self-deception, or that it is a matter of degree. But it is conceivable that there is such a distinction. Whether this is a case of self-deception or error depends upon persons' motivation for resisting the evidence of what they believe. If the motivation has to do with an interest they have in presenting themselves to themselves as not believing or feeling what they deny, it is self-deception, not error.

Thus, the difference between self-deceivers and persons who are in error about their beliefs or feelings is not always that self-deceivers will not acknowledge evidence that they believe or feel what they deny, while mistaken persons will. As we have seen, in the case of belief, persons who hold conflicting beliefs because of error, like self-deceivers, will not always acknowledge

evidence that they believe what they deny. When the self-deception is about what persons feel, not what they believe, it is easier to see the contrast between self-deception and error vis à vis whether persons will acknowledge evidence that they believe what they deny. For if persons are mistakenly unaware of what they feel they should acknowledge it when shown the evidence. Feelings are more closely connected with interests, whereas beliefs can have many different kinds of sources. So evidence that persons believe something may not always be what it seems. There might be some other way of reading the situation. The mistaken belief might be due to stubbornness rather than self-deception. But the behavioral expression of feeling will verify more strongly that persons feel what they deny, and if persons will not acknowledge evidence about their feelings this leads us to suspect an explanation other than error--self-deception.

Thus, persons' denials of the evidence that they feel what they deny can be grounds for saying that persons are purposefully denying what they feel. Purposeful denial of what one knows to be true is lying, but this is not simply lying because the denial is also sincere. In some sense, persons really think that they do not have this feeling.

In the case of belief, it may be more difficult to determine whether persons' denials are due to self-deception or error, but such cases of self-deception about beliefs exist. The difficulty may be seen in a case like that of Mrs. Faith Beloyal. There we did not have decisive evidence that her bizarre denials that she thought her son to be guilty were self-deceptive and that this indicated that she also believed that she thought of him as guilty. Such determined attempts to resist evidence countering her claim that she thinks he is innocent could be due to stubbornness or ignorance. We pointed out that we needed something other than such denials as evidence that Mrs. Beloyal believed that she thought her son to be guilty. Thus, we said, if she refuses to take apples from Willie (they might have razor blades in them), we would have a stronger case for attributing self-deception. In this case she would be not just denying unreasonably, she would also be acting like she thinks she is guilty. But in the case of self-deception about one's beliefs, such evidence still is not enough to decisively determine whether a particular case is self-deception rather than error. Although it is unlikely, Mrs. Beloyal's refusal to take apples from Willie could conceivably have some other explanation than purposeful error based on interest (i.e., than self-deception). It might not be an indication that she knows she believes Willie is guilty.

It could be merely that she erroneously and perhaps stubbornly believes that she thinks Willie is innocent. But the fact that self-deception is sometimes difficult to distinguish from error does not mean it does not exist. In a case like this we would look for evidence that persons deny what they deny because of interests.

The most plausible explanation for denying that one feels something when one has evidence for one's having such feelings is still that one is purposefully denying what one feels, for it seems likely that one knows in some sense that this is what one feels. But since there is the possibility, however remote, of error being the explanation, it may be that in some cases we cannot tell until after the self-deception is over if the person has self-deceived. For example, when Mrs. Beloyal finally realizes that she did think Willie guilty, what her interest was in denying this, and that her interest motivated her denial, she realizes that she was deceiving herself.

Another good example of this difficulty of determining whether a denial is an indication of error or of self-deception until the person stops denying is provided by Freud in his account of his missing a train. Freud misses a train in Cologne because "it was nowhere to be found." He was travelling to visit his brother in England and was under pressure to come directly. But he wanted to stop in Cologne to see the

museum. Missing the train gave him the opportunity to do so. The next day, after he had visited the museum he had long been wanting to see, and as he continued his journey, he writes: "a clear memory emerged of my having seen a large notice in the station at Cologne There, waiting for me, had been the train in which I ought to have continued my journey."²² On the previous day, he failed to notice that he perceived the train. This enabled him to believe that his stop-over in Cologne was an accident. This was one method of dealing with his conflict about staying over in Cologne while not eliminating it. Out of respect for his not wanting to displease his expectant brother he sincerely denied awareness of the train. That way he did not have to think of himself as deciding to miss the train. This conflict about what to do (to continue the trip or to stay overnight in Cologne) yielded another conflict--about seeing the train. Not wanting to continue the trip was his motivation for not wanting to see the train. The conflict about what to do resulted in his being self-deceived about whether he saw the train. He sincerely claimed not to see it, denied it was there, and yet to an outsider, it must have looked as if there was evidence that Freud saw the train. Also there would have to be evidence that he did not see it--that his denials were sincere and non-mistaken. Perhaps Freud was looking right at the train or at the notice for it.

The conductor and other bystanders inform him that this is the proper train. The train is properly marked and in full view. Yet Freud thinks it is the wrong train. If we know about his interest in staying over in Cologne and his conflict about that we might have a hunch that he is in self-deception. Or we might hypothesize that something has gone wrong with his senses. There would be no way to decisively ascertain that this is self-deception, that his interest motivated his sincerely claiming not to see the train, and that he also really saw it. But surely after the fact, the next day when Freud realized that he had failed to notice his perceiving the train, it becomes plausible that this is a case of self-deception.

We see that we need some other way of distinguishing self-deception from error because sometimes the type of evidence that we get for persons' believing something (such as their unreasonable denial of any evidence for believing it) is not enough to indicate that they believe it. Strictly speaking, the mother's denial that her refusal to take apples from her son is evidence that she believes in his guilt, is not necessarily evidence that she really believes him guilty. It could be simply that she mistakenly believes he is innocent. Likewise, that Freud denies seeing the train (plus apparent evidence that he sees it) does not mean he sees it. It could be that he does not see it and

that he is simply mistaken when he claims that it is not there. And so it could be that his behavior that seems to confirm his recognition of the train has some other explanation.

What we have been pointing out is that although in self-deception persons really believe or feel what they deny, and although certain features of the particular way they deny it give evidence that they really believe or feel it, such evidence is not enough for us to conclude that this is self-deception. The evidence would be stronger if we could also claim that they are motivated to make such a denial. When simple error seems like an unlikely cause of a persons' denials of evidence about their feelings, we need some other explanation. The difference between error about one's beliefs or feelings and self-deception lies in the kind of motivation each has.²³

Interest

As has been discussed, the motivation for the self-deceptive denial-avowal consists of interests persons have in not having the denied beliefs or feelings. Thus, the beliefs or feelings they claim to have are based upon these interests. Persons have the avowed beliefs or feelings because they want to have them. Self-deception differs from simple wishful thinking, however, because self-deception involves having conflicting beliefs or feelings. Wishful thinkers may believe or feel some-

thing because of desire, but the desire is not connected to not wanting to have the beliefs or feelings they already have. This is another point of distinction between self-deception and error. Erroneous belief based on interest differs from self-deception in that self-deception about belief involves conflicting beliefs and an interest in not having one of them.

Several points are important concerning the role of interest in self-deception. One is that it is because of persons' interests in not having certain beliefs or feelings that persons deny them. This is connected with the fact that persons must know they have these beliefs or feelings. This is different from persons who hold contrary beliefs erroneously (they know such beliefs are contrary but are only aware of holding one of them).

Another point is that the initial situation that gives rise to self-deception is one in which persons have beliefs or feelings that they do not want to have. Thus, they have interests in not having those beliefs or feelings, and in the name of those interests, they try to eliminate them. If they fail and do not choose to accept the resulting conflict in themselves, they self-deceive. Thus, self-deception grows out of situations in which persons try to eliminate conflict, fail, and will not accept this. The original interest,

then, that gives rise to self-deception, is an interest in resolving a certain conflict.

Finally, as discussed above, interest, in the present view, is one source of belief. Therefore, self-deceivers' claims to believe on the basis of interest can involve genuine beliefs. They are not ruled out because they are not held on the basis of evidence.²⁴

Purpose

Self-deception is purposeful. It has its source in persons' interests. In this sense, self-deception is something persons do, not something that happens to them. The cause of self-deception is not something external to them, such as actions done by other persons, or something that might be within them but not part of them, like a drug or a hypnotic suggestion. Self-deception issues from a genuine, full-fledged, representative part of persons. But calling it a part of persons is meant to emphasize that it is not external to them, and is not meant to indicate that there is only a part of persons that self-deceives or that is in self-deception. The whole person is in self-deception, the whole person self-deceives. This is what is meant by the claim that self-deception is by and of the same self. Self-deception would not be something a self does if it were not by and of the same self.

That self-deception is something persons do, distinguishes it from error. The network of interests

behind self-deception are operative in self-deception. Such interests may exist in mistaken persons, but they are not operative in such persons' error in the same way. While interests play a causal role in wishful thinking, they are not the wishful thinkers' reasons for wishful thinking. But the interests behind self-deception are persons' reasons for self-deceiving.

Conclusion

There seem to be two major paradoxes about self-deception. One is that self-deception involves purposefully being unintentionally inaccurate about oneself. This paradox involves problems of knowing and not knowing and relates to questions about the unconscious. Much of the philosophical literature on self-deception focusses on this problem. But this is not the paradox of concern in this thesis. Rather, this thesis is concerned with the second problem--that self-deception involves conflicting beliefs or feelings and the self's attitude towards and capacity for dealing with such conflict. Self-deception grows out of situations in which persons are trying to eliminate conflict. But it involves creating or perpetuating conflict. The issue here is what kind of self can have the capacity to include conflict in itself in the way required by self-deception.

Footnotes to Chapter II

¹Of the views found in contemporary philosophical literature, the present view of self-deception has most in common with the views of Herbert Fingarette, 1969, and Amelie O. Rorty, "Belief and Self-Deception," Inquiry, XV (1972), 395-410. In addition, the present view of self-deception is influenced by descriptions of phenomena such as parapraes, defense, ego and super-ego formation found in the writings of Freud that are cited in the bibliography of the present work. An excellent survey of the development of modern philosophical thought on the subject of self-deception can be found in Daniel Shapiro, "Self-Deception" (unpublished Ph.D. dissertation, Dept. of Philosophy, The City University of New York, 1975).

²Fingarette, 90-91.

³Rorty, 1972; also, "Adaptivity and Self-Knowledge," Inquiry XVIII:1 (1975) Appendix, 19-22. See also D.W. Hamlyn, "Self-Deception," Proceedings of the Aristotelian Society, Supp. V. XLV (1971) 59.

⁴Fingarette, 139-150.

⁵Rorty, 1972, 393, 395, 399; 1975, 19.

⁶Fingarette, 90-91.

⁷Compare Rorty, 1972, 399, who says that what must be explained about self-deception is not only that it is apparently a move to eliminate dissonance, but that it preserves it.

⁸Rorty, 1972, 393.

⁹Rorty, 1972, 398-400, points out that the claim that self-deceivers deceive themselves about themselves is not trivial. She claims that "cases of self-deception do have the same type of explanation," and that some restrictions must be placed on "the sort of interest served by self-deception." She suggests that what is at stake in self-deception, in general, are self-deceivers' conceptions of what sorts of persons they are. To detect self-deception in other persons it is necessary to understand their self-concept, what the beliefs or feelings mean to them, what their attitudes are towards having them.

Footnotes to Chapter II, cont.

¹⁰Cf. B.A.O. Williams, "Deciding to Believe," in B.A.O. Williams, Problems of the Self (Cambridge: Cambridge University Press, 1973), 136-151. Williams, 149, makes a distinction between truth-centered motives for having beliefs and non-truth-centered motives. In the former case we want what we want to believe to be true, and that is why we want to believe it. In the latter case, we want to have the belief for some other reason (perhaps Rorty would suggest "self-image"). Williams thinks that when we have non-truth-centered motives we have to self-deceive, because there is something wrong with a "conscious project to make myself believe what I want to believe." Belief is fundamentally a truth-centered concept. The notion of willing to believe is odd because we think of belief as being truth-oriented. Thus, self-deception seems to be required for such beliefs. Compare this thought about the deviant sorts of beliefs existing against the background of truth-oriented beliefs with Rorty's comment that self-deceivers must hold normative criteria for rational belief in order to believe self-deceptively (and thus, irrationally). If persons do not hold such criteria, they may be in error but not in self-deception. Thus, she says that while self-deception is intentional, it is not deliberate: "self-deception must be seen in bas-relief to be identified at all" 1975, 22.

Finally, Fingarette, 1969, 30, also asserts that self-deception is about whether one has a belief, not about whether the belief is true. Thus, he says that self-deception cannot be described simply as belief in the face of evidence. What is important is the strength of the person's belief, not the truth of it.

¹¹Patrick Gardiner, "Error, Faith and Self-Deception," Proceedings of the Aristotelian Society, LXVIII, new series (1969-70), 223-224.

¹²Rorty, 1972, 393.

¹³Most philosophers see the sincerity of the self-deceptive claim as basic. It is for them the source of seeing self-deceivers as "double-minded," as doing inconsistent things at once, or as having conflicting beliefs or feelings. What would be intelligible if it were divided between two persons is seen as paradoxical and in need of special explanation in one person. See Ronald B. deSousa, "Self-Deception," Inquiry, XIII, (1970); H.O. Mounce, "Self-Deception," Proceedings

Footnotes to Chapter II, cont.

of the Aristotelian Society, supp. v.XLV (1971); David Pears, "Freud, Sartre and Self-Deception," in Richard Wollheim, 97-112; Fingarette, Ch.II; Rorty, 1972.

¹⁴Some philosophers have discussed the pros and cons of talking about self-deception as a form of lying or other-deception. Fingarette rejects this approach because it emphasizes contrary beliefs, and he wishes to shift the emphasis to a notion of avowal and away from the cognitive notion of belief. Shapiro shows that seeing self-deception as simply a form of lying to oneself sheds no light on the paradox of contrary beliefs that self-deception generates, or it unjustifiably reduces self-deception to error. He suggests that lying and other deception are concepts that are just as complicated as self-deception, and that they ought not to be used to elucidate self-deception. Thus, when self-deception is herein referred to as lying to oneself, the concept of lying to oneself must not be understood as being explanatory of the problem.

That self-deception is lying to oneself is also ruled out because such a view presupposes a transparency model of the mind and omits the screening that goes on in self-deception. Thus, Pears is critical of the view he attributes to Sartre, that self-deception is lying to oneself.

¹⁵Fingarette, 50-57.

¹⁶Although, lying might lead to self-deception if the theory of cognitive dissonance (Leon Festinger, "Cognitive Dissonance," Scientific American, 207, 4 (October, 1962), 93-102), is right. One variation of what the theory of cognitive dissonance points out might be that when persons try to reduce psychological dissonance, they do not change their beliefs or feelings, rather, they deny them self-deceptively. Thus, in order to reduce the dissonance created by lying in situations in which there is not much external reinforcement for lying, persons might deceive themselves into thinking they are not lying and that they believe or feel what they claim. There is a close connection between lying and self-deception, but they are different phenomena.

¹⁷Freud, 1940e, writes that the defensive person who is undergoing a trauma "replies to the conflict with two contrary reactions, both of which are valid and effective." Cf. Wollheim, 237. The view that sees persons as contradictory vies with the view that there is a true self or a simple direct way to be true to oneself, for prominence in contemporary thought. Cf. Lionel Trilling, Sincerity and Authenticity (Cambridge: Harvard University Press, 1971), 52-56.

Footnotes to Chapter II, cont.

¹⁸Aristotle, *Nicomachean Ethics*, trans. Martin Ostwald, (New York: Library of Liberal Arts, 1962), Bk.III.

¹⁹Freud, 1901.

²⁰Fingarette, 1969; Hamlyn, 58; Rorty, 1972.

²¹James, 212.

²²Freud, 1901, 227-8.

²³See David Wood, "Honesty," ed. Alan Montefiore, *Philosophy and Personal Relations* (Montreal: McGill-Queens University Press, 1973), 191-223, for discussion of the difficulty of determining what persons' engagements in the world are and whether persons are in self-deception. Gilbert Harmon, "Knowledge, Reasons and Causes," *Journal of Philosophy*, LXVII: 21 (Nov.1970), makes a point that can be interpreted as supportive of this when he claims that belief is the result of a reasoning process and the reasoning process is not necessarily conscious. What persons say are the justifications for their believing what they believe (and not just for their belief) is not necessarily the real justification.

²⁴The ideas in this section are similar to some of Rorty's claims. She emphasizes that many of our beliefs are irrational, not just the "difficult fringes of belief--akrasia, self-deception, ambivalence, and their kith and kin," (1975,19). Thus, she argues against Szabados (Bela Szabados, "Rorty on Belief and Self-Deception," *Inquiry*, XVII (1974), 464-473), who claims that having conflicting beliefs is incomprehensible. For example, she accuses him of taking "the normative criterion for rational belief to provide a descriptive model for believing" (1975, 19). That self-deceptive beliefs conflict is not, for Rorty, a distinguishing mark of them, since she believes that many of our beliefs are irrational in this way. What distinguishes self-deceptive beliefs from erroneous ones, and from those based upon wishful thinking, is that self-deceptive beliefs are based upon a certain kind of interest. Rorty says (1972) the interest of self-deceivers is in having certain self-images. This point is similar to the present claim that what interests self-deceivers is having or not having certain beliefs or feelings, not primarily, having those beliefs or feelings be true or appropriate.

Hamlyn also mentions the irrationality of much of our mental life and Fingarette's analysis of self-deception emphasizes that it involves affirmations or denials of identities, and that it is not mere wishful thinking.

Chapter III

THE INTEGRATED THEORY

Introduction

Certain views about persons tend to emphasize integrative aspects of their motivation. By integrative aspects I have in mind those features of persons in virtue of which they strive to eliminate conflict in themselves. Views that emphasize integrative features are associated with a possible theory about the self that is referred to in the present work as the integrated theory.¹ The claim of the integrated theory is that persons always strive to eliminate conflict. The integrated theory may not be entailed by or implied by the views that bring it to mind, and it is not a theory that is being attributed to a specific philosopher. But as a view that comes to mind in association with certain other points of view, it deserves study. The task of the present chapter is to see what the integrated theory claims.

First let us look at some of the views that emphasize integrative aspects. All have in common the view that defines persons in terms of a process by which they aim at eliminating conflict. In the present work, three contemporary philosophers have been chosen for

illustration of this view--Fingarette, Frankfurt, and Watson.

For Fingarette the self is constituted by a process called avowal, by which he seems to mean consistent avowal of one's engagements.² By engagement Fingarette means

to characterize, in most general terms, what someone does or what he undergoes as a human subject; it is how an individual finds and/or takes the world, including himself. It is a matter of the activities he engages in, the projects he takes on, the way the world presents itself to him to be seen, heard, felt, enjoyed, feared, or otherwise 'experienced' by him. It is logically necessary that it should be typical of our description of an individual's engagement in the world that the description be cast in terms of such categories as aims, reasons, motives, attitudes and feelings, or understanding and 'perception' of the world and himself. What a person does not somehow take account of is not part of his engagement in the world, though it may have effects on the course of his engagements.

By avowal Fingarette does not mean identification, for, identification, he says, presupposes a self. The significance of this distinction between identification and avowal is puzzling, as there does not seem to be any empirical difference between identification and avowal, and the same questions that might arise about identification if it were not in the context of an already avowed self, arise for avowal. Perhaps these are terms meant to describe different phases of a process. Perhaps when one is already constituted as a self, the result of avowal of one's engagements, one is capable of identifying with things external to oneself, such

identification being dictated by one's avowals which contain one's "governing values."⁴ In any case, what both avowal and identification denote for Fingarette can be interpreted to include a striving for consistency.⁵

Another philosopher, Frankfurt, also thinks that identification is a process of conflict-resolution.⁶ We identify with our desires when we make a conflict-resolving decision that we want them to be effective. What we identify with in this way, Frankfurt says is internal to us as persons, as opposed to being ours as human beings but external to us as persons. For example, he explains that one might consider refraining from one's impulse to compliment a friend because of jealousy, but if one decides to compliment the friend in spite of the jealous impulse, the jealous impulse is external.

A third philosopher who suggests the integrated theory is Watson.⁷ Watson's approach is like Frankfurt's in that he is looking for a way to distinguish different statuses of the impulses in conflict within persons. In particular, Watson is trying to understand how persons can act "in spite of themselves," and he has in mind weakness of the will and other cases of unfree action. If self-deception can be seen as an instance of acting in spite of oneself, his view is applicable to that too. Watson differs from Frankfurt in that he rejects the criterion of identification and higher-

order decision. He is critical of the approach that tries to distinguish persons in terms of decision or attitude, and he favors distinguishing them in terms of rational judgment. But both views have in common an emphasis on consistency and conflict-resolution. Watson favors dividing the person into motivational systems, along the lines of the platonic division of the soul into appetitive and rational parts. He assigns a higher status to the rational part.⁸ Thus, Watson says we have two motivational systems--blind desire and judgment (or valuing). When our desiring system is effective in spite of our conflicting values, he says, we are acting in spite of ourselves. The implication is suggested that the real person is the one who makes rational "cool" judgments, not the one who simply desires or whose reasons for desiring are based on desire alone.

On all of these views there is a tendency to identify a part of the human being as the true self. Some of one's states--desires or beliefs (or engagements for Fingarette) are part of one as a person, some are not. For Fingarette this means a distinction between the individual and the person. For Watson it means a coherent set of values, a single point of view, in distinction from desires. And for Frankfurt it means an internal attitude based on decisiveness, where decisiveness implicitly involves conflict-resolution, as opposed to what is external in us because we do not

want it to be effective.

As stated earlier, views like those just described, insofar as they emphasize integrative impulses, bring to mind the integrated theory. The fundamental claim of the integrated theory is that persons constantly strive to eliminate conflict. The integrated theory is a motivational theory about persons. It differs from the views just mentioned in that they have as their purpose an emphasis on the question of what is ideal behavior for persons. Those views are concerned with setting up criteria for good action. They are not trying to answer a question about responsibility in terms of the question of agency. Rather, they are more interested in the normative question: what is the essence of being a good human being. In this sense of "responsibility," persons might be liable (responsible) for their actions even though these are not good (responsible in the other sense) actions.

The integrated theory, a motivational theory, is suggested by these other kinds of views because of their emphasis on consistency. And insofar as these other views pick out only certain features of human beings' behavior for inclusion in the concept of a person, there is another similarity. For in denying that persons can have any other attitude towards conflict than wanting to eliminate it, the integrated theory suggests the implication that conflicting impulses in persons

do not represent them. Persons are only represented by consistent impulses. If there is a conflict between Tom's wanting to act in Mary's best interest and his wanting to hurt her, it is either the case that neither of these impulses represents Tom, or only one of them is his true impulse. The integrated theory rules out the possibility that both impulses truly represent him as a person. He is in conflict about what he really wants to do and something he mistakenly thinks he wants to do, and he does not know which is which. Or he wants to do some third thing that excludes both of the original alternatives. What he wants above all, in all of these cases, is to be single-minded, to be consistent.

In saying that persons always try to eliminate conflict, the implication about policies is that persons cannot make policies in spite of conflict (they must resolve conflict to make policies), and that persons are not capable of accepting conflict in themselves. If persons fail to eliminate conflict they will continue to try to do so. Thus, in conflict persons are always driven towards consistent integration.

The concept of eliminating conflict in oneself may be ambiguous. It could mean elimination of conflict in one's decisions, judgments, and avowals. But Frankfurt, Watson, and Fingarette think that decisions, judgments, and avowals are, by definition, conflict-free. This suggests the possible view that once policies

are made, conflict is eliminated in the person. Such a view might rule out the possibilities that persons make policies in spite of continuing conflict, that policies might include conflict, or that different policies might conflict. These are possible implications of the integrated theory on the issue of policy-making. Such implications are suggested by the view that our better selves, what Frankfurt, Watson, and Fingarette call "persons," are unconflicted. This normative view suggests the empirical view that our real selves are unconflicted. The empirical thesis is that persons are not motivated to act by conflicting impulses. A contrasting view, which is discussed in Chapter IV (the fragmentary theory), is that unconflicted policies do not rule out, necessarily, conflict in persons, and that the remaining conflict (not just the policy) represents persons too.

Frankfurt has a different reading for such cases. He discusses an example in which a person has conflicting impulses and makes an arbitrary decision to follow one of them. The conflicting impulse continues, but in his view it is external to the person-- unless, he adds, the decision itself is external.⁹ One implication of this might be that when a decision does not try to put conflict to rest, the decision does not belong to a person; in this case, the decision would be external. Or, alternatively, that if the decision does belong to the person, the remaining conflict does not. These two interpretations

suggest the integrated theory, even though Frankfurt discusses them for a different reason. They suggest the idea that there is always discontent when conflict remains in a person and always an attempt to eliminate it.

Since human beings sometimes fail to make the kind of decision Frankfurt and the integrated theory see as being characteristic of persons, the integrated theory needs some way of explaining putative counter-examples. At this point it can persist with an empirical claim that anything human beings do can be properly interpreted and described as being guided by consistency impulses. Fingarette says that when integrative impulses fail, persons regress or dissolve.¹⁰ And Frankfurt claims that inability to be decisive destroys persons.¹¹ But they also might claim that integrative impulses, whether they fail or not, are behind everything we do. They might claim that the individual, the regressed or dissolved person, is struggling to regain consistency. In another interpretation, the normative theory that suggests the integrated theory says it is recommending the concept 'person' be reserved for those situations in which human beings are guided by impulses to achieve consistent integration. The latter seems to be what Frankfurt has in mind when he claims that in the inquiry into what persons are, what we are looking for is a concept that captures what is most important to us.¹²

This view of reserving the concept 'person' for an ideal of human existence is echoed in Fingarette's developmental view of the self. He claims that we develop from being individuals with unintegrated engagements to being persons--i.e., to having consistently avowed engagements.¹³ Being a person is an ideal that we aim at and that is the ultimate moral and healthy state. And for Watson, insofar as we are persons we are represented by our best and most coherent values.¹⁴

There is a link between the integrated theory and the normative view that suggests it, and it results in there being two aspects to the integrated theory. The two aspects of the claim made by the integrated theory are (1) That persons always strive to eliminate conflict: all so-called counter-examples--apparent self-acceptance of conflict, arbitrary decision without eliminating conflict, a surveying and ranking type of policy-making, conflicting policies (all of these phenomena will be discussed at greater length in the following discussion and in Chapter IV)--can be properly redescribed as being guided by or in the context of integrative impulses. (2) Secondly, if such alternatives do occur, the claim of the integrated theory is that when integrative impulses do not dominate, what is denoted by the name "person" dissolves or regresses. This claim is derived from the normative view that the concept of a person should be reserved for those times when human beings fit the

description given by the integrated theory. In this latter view there are two phases to the existence of human beings vis à vis motivation, and human beings can slip back and forth between them. Presumably the integrative phase is preferred for the concept of a person, either because it is seen as being the way human beings most frequently act, or because it is seen as being, not necessarily most frequent, but better. Perhaps the empirical aspect of this claim is that human beings are no longer agents when they regress to the non-person stage, and thus, they are no longer responsible.

In summary, then, the integrated theory either claims that human beings always strive to eliminate conflict or that when they do not, they are not responsible in the sense that they are not agents.

Another question that the integrated theory seems to bear upon is whether unity is possible in a self that is not always striving for consistency. The question of unity in the self, and the questions of decision-making, judgment, and identification are linked. Decision-making, judgments, and identification require unity because persons require unity. The integrated theory suggests that unity requires constant consistency impulses and that the alternative is a multiple-self view. If conflict can thrive and be tolerated by persons, if decisions can embody conflict, and persons can identify with conflicting things and have conflicting values,

and there is no overall impulse to become consistent, the integrated theory takes this to indicate the impossibility of having policies, of making decisions or judgments. Such fragmentation indicates a split into different selves.

In the following pages a fuller view of the integrated theory of the self is given by way of a description of Fingarette's view. His view suggests a clear picture of what the self looks like on the integrated theory. We shall also describe some aspects of his view of self-deception, because it suggests one of the possible interpretations of self-deception that an integrated theory could have. Next, we'll discuss the concept of a person and the handling of self-deception by the integrated theory in terms of the versions of the integrated theory of Watson and Frankfurt. Here the virtues and some of the problems of seeing self-deception from the point of view of the integrated theory will be pointed out.

Fingarette: Persons Defined in Terms
of Consistent Avowal

Fingarette¹⁵ makes a distinction between the individual and the person. A person is an achievement of the individual. At first the individual is made up of various systems. The goal of achieving consistent unity among these systems is the goal of becoming a person. One possible interpretation of Fingarette

is that he assumes that to have the goal of integrating the atomic fragments of an individual is to aim at consistency. As mentioned earlier, he calls the systems that constitute the individual "forms of engagement." (Engagement is defined in the previous discussion and below.) They are ways of taking into account one's involvements in the world. The process by which individuals try to synthesize their engagements into a unified self is called "avowal." As mentioned above, avowal is a kind of inner act similar to but broader than psychological identification; avowal constitutes the self, while identification presupposes it.¹⁶ "Avowal is necessarily manifested by integration of the avowed engagement into the system of the personal self."¹⁷

The resulting self is a kind of community:

In the community of the self, as I see it, however the members are the various originally independent forms of engagement, the rudimentary but unified complexes of reasons-motive-feeling-aims-means and moralistic reaction. When these atomic forms acquire citizenship in the community of a self, they are significantly transformed by virtue of the mutual interaction and interdependence, yet they do not entirely lose their distinctive character or quasi-autonomy. Ideally, the uncivilized and unsocialized child becomes a civilized member of the community, and yet he retains his individuality; and, analogously, the primitive and independent forms of engagement are ideally integrated with each other within a self, and they remain specific even though subtly transformed, even though serving as a medium in which the one person is constantly expressing himself.¹⁸

To become a person, for Fingarette, is to avow one's engagements. Avowal constitutes the self as it trans-

forms engagements into an integrated unity. That transformation of engagements is necessary for integration, could be interpreted to mean that avowal means consistent avowal. We must endorse our engagements. But there are limits to what we can avow legitimately. There are requirements that must be met for proper avowal. The procedure of avowing involves transforming the engagements, as is necessary, so that they fit together with one's other avowed engagements. The goal is an "ideal integration." That Fingarette means this ideal integration to be consistent integration is suggested both by his talk of the resulting "governing values," and by his remarks about self-deception, which he sees as growing out of an attempt to avow engagements that conflict with one's governing values.

Fingarette¹⁹ explains that transformation of engagements is necessary in avowal because some engagements are there to be discovered-- they are part of our past and to properly constitute ourselves (to correctly claim who we are) we must acknowledge them. Other engagements are created and who we may choose to be in this sense may conflict with what our history deems us to be. Fingarette also marks this distinction between discovering and deciding or creating our identities in Self in Transformation,²⁰ and he urges there, and in Self-Deception, that the self must integrate what is given and what is creatively chosen and that the choices

we make in defining ourselves must include both sorts of elements even when they are initially in conflict. Avowal includes eliminating conflict--the engagements we avow must be transformed as much as is necessary ("subtly transformed"²¹) to create a consistent system. It could be, of course, that avowal is simply a process of making one's engagements coherent as all being part of a single person. But since Fingarette thinks avowal involves transformation of engagements, this view suggests that the coherency aimed at is one that eliminates conflict.

Fingarette's View of Self-Deception

For Fingarette, the situation that gives rise to self-deception is that individuals are in situations in which they are attracted to engagements that are inconsistent with persons' governing values.²² Fingarette's distinction between individuals and persons is that individuals are constituted by many unintegrated engagements, but persons transform these engagements into a consistent whole via a consistent set of governing values. As an individual one may identify with different engagements, and these may conflict or reflect inconsistent avowal. A person may have many different engagements but they cannot be in conflict. Thus, the situation that can give rise to what he calls the person's putting "himself in self-deception" is one in

which what he calls the individual

is strongly inclined to a form of engagement which is strongly inconsistent with the person's governing principles (the person's avowed aims, ideals, values, cultivated tastes, moral principles).²³

Fingarette says that the person has three options when this situation arises. The first option is to give up the inconsistent engagement. This is the option normally chosen by the adult, according to Fingarette, and he adds that

to say that a person has put himself in self-deception, however, is to say that the person could not bring about a total abandonment of the engagement.²⁴

Thus, we see that for Fingarette one way that self-deception can arise is when there is a deadlock conflict between the individual's engagement that is attractive to the person and the person's governing values. In such cases, he says, the person cannot abandon the dissonant engagement.

The second option is to pursue the engagement, "the person avowing it as his."²⁵ Fingarette calls this option self-acceptance and he claims that it is a spiritual crisis and that it requires great courage:

it would be to confess one's finitude and, if the conflict is predominantly moral, one's sin; it would be to lay oneself open, vulnerable, as a radically divided nature, and to hope for grace of some healing movement which it is not at the moment entirely within one's own personal powers to effect or even to foresee.²⁶

Fingarette adds that "from this perspective, self-deception is always a spiritual failure"²⁷ and yet the kind of perfection involved in such self-acceptance is so difficult that it can be disastrous. Here Fingarette seems

to appeal to self-acceptance as a kind of ideal, while also warning that it may be a disastrous, and thus unhealthy, solution. Here it is important to note that Fingarette seems to be advocating two different states as ideal. He first seemed to advocate consistent integration as the ideal, but now he seems to be saying that when there is irresolvable conflict, self-acceptance is the ideal. There is no contradiction here, but it is important to emphasize that the ultimate ideal for him is integration. This contrasts with a view that might regard self-acceptance as a possibility independent of integrative impulses. Hence, the ideal would be honesty, not consistency. Fingarette's idea of self-acceptance seems to involve the hope for integration, and it occurs in the context of failed integrative impulses. Self-acceptance is a kind of tragedy for the self, while on a theory that does not emphasize consistency, it need not be so.

It must also be noted at this point that Fingarette in this passage has allowed for the possibility that in self-acceptance persons could avow engagements that conflict with their governing values. This is confusing because we have been interpreting him as defining "avowal" as "consistent avowal," in which case it would be impossible for persons to avow dissonant engagements (although it would be possible for individuals to claim them). Furthermore, Fingarette goes on to claim that self-

deception occurs because persons are unable to avow dissonant engagements--because they conflict with governing values--, implying that such avowal is not possible by persons. All of this can be seen to indicate the confusion generated by the distinction between the notions 'person' and 'individual' and the corresponding notions of 'consistent avowal' and plain 'avowal'. Fingarette's notion of 'avowal' seems to be ambiguous, sometimes meaning "consistent avowal," sometimes not. The important point for present purposes is that in the integrated theory self-acceptance as described by Fingarette would not be possible-- persons could not avow engagements that conflict with their governing values. But self-acceptance is possible in the integrated theory if it is accompanied by attempts to resolve the conflict, and this is what Fingarette says too.

The third option, when there is a "stalemate between inclinations which the individual will not give up, and the refusal by the person to avow these inclinations as his," is that "the individual does engage himself in the way to which he is inclined, but the person refuses to acknowledge the engagement as his."²⁸ This is self-deception, in which the person is "at war with himself." The self-deceiver

attempts to save his integrity at a price which amounts to surrendering, however indirectly, the very integrity he cherishes.²⁹

Fingarette points out that psychoanalysis uses morally neutral language for this phenomenon--"inner psychic conflict"-- but he goes on to justify describing self-deception in moral terms as "spiritual cowardice."³⁰ The cowardice he refers to is the failure of integrity-- by which he means the failure to become consistent and not just the dishonesty or inaccuracy involved in the claims of the self-deceiver. It is just this point of contrast between the neutral view of self-deception as psychic conflict and the view of it as spiritual cowardice that marks the integrated theory. For the integrated theory there are good and bad sides at war.

Fingarette cites two elements in the "movement to self-deception"-- they are spiritual cowardice and inner warfare. He claims that these two elements presuppose "a person with a certain integrity,"³¹ and he adds that "not infrequently what is threatened is some aspect of integrity rooted in moral concern. So, it is Fingarette's view that persons self-deceive because they care about being integrated, they cannot do it, and they cannot accept this: "the movement to self-deception is rooted in a concern for integrity."³² But he is not just saying that self-deception grows out of a frustrated desire to make inconsistent engagements consistent. He is making the additional claim that concern for integration is the "authentic" concern of persons. He explains that this is why self-deceivers are not mere

cheats--we have compassion for them because their "authentic inner dignity" motivates their "self-betrayal."³³ There are several other comments that Fingarette makes to enlarge upon his emphasis on integrity in self-deception. One is that this integrity is presupposed, counted on and aimed at in "insight psychotherapy":

It is this authenticity of moral concern rooted in personal integrity which constitutes the basis for optimism using insight psychotherapy.³⁴

Fingarette can be interpreted to mean that psychotherapy aims at persons' assumed concern for the consistency of their engagements with their governing values.

The Integrated Theory and Self-Deception

The foregoing-described view of the self and self-deception is an example of the integrated theory because it emphasizes consistent integration. Even though self-deception is seen as destroying consistency, it is seen as being the result of an attempt at consistency. Thus, when something like self-deception occurs, in this view, persons regress or dissolve into individuals. Translated into the language of Watson, we could say that in self-deception persons are acting in spite of themselves. And in Frankfurt's language we could say that certain beliefs or feelings involved in self-deception are external to persons.

An idea that is linked with the integrated theory's view that persons are always striving for

consistency and that this applies to self-deception too, is that there is one point of view that represents persons' true values and that it is in virtue of this point of view that persons are trying to become consistent. This suggests that in self-deception there are integrative impulses that dominate and that these impulses are connected with a certain set of values, beliefs or feelings, or attitudes. Hence, Watson says that in cases of weak will, persons are acting in spite of themselves. The self in spite of which one acts is the true self. It is represented by the consistency impulses to which the integrated theory refers. Let us look at this notion of acting in spite of oneself as it might apply to self-deception. The claim is that in self-deception persons maintain beliefs or feelings that they do not want to have (the ones they self-deceptively deny) in spite of their desire not to have them and their resulting sincere claim that they do not have them.

But another idea of Watson's plays a role here. He claims that the impulse in accordance with which we act in weak will is ours--its source is our desiring motivational system. This view of the self, he tells us, enables us to see how persons can be obstructed by their own will. Thus, Watson is making two points: in weak will we act in spite of ourselves (i.e., in spite of our better judgment) and we are obstructed from doing what we most want to vis à vis our better judgment by our

own will. Such an explanation aims at preserving the unity of the person in situations that appear to split persons into factions or selves. Watson is saying that persons can act in spite of themselves and be obstructed by their own wills because persons consist of different motivational systems. Applied to self-deception, this idea of being obstructed by one's own will can be used to support the claim that self-deception is by and of the same self.

But if we go along with Watson's view we cannot maintain without qualification that self-deception is a case of acting in spite of oneself. Since on Watson's view the beliefs or feelings that one self-deceptively denies (and thus maintains) are part of oneself, self-deception is partially acting in accordance with oneself. This is due to the fact that to self-deceptively deny a belief or feeling is to maintain it, and to maintain it is to endorse it. One might take issue with this claim, and point out that persons can accept that they have certain beliefs or feelings that they are unsuccessful in eliminating without endorsing these beliefs or feelings, and that to say that such acceptance is endorsement is too broad a notion of endorsement. But we could respond to this view by emphasizing that to the extent that we do not want to have beliefs or feelings that we have, we do not endorse them. But at the same time, to the extent that we have them, we do en-

dorse them. Self-deception could be described as acting at the same time in spite of and in accordance with oneself.

A view such as the integrated theory, that sees all phases of self-deception as being guided by impulses of consistency, suggests that in self-deception there is one attitude that represents persons more truly than other attitudes and this suggests the possibility of a split in the self such that self-deception is seen as involving a true self and a false self, a deceiver and a deceived self, that are not the same entity. The integrated theory attempts to preserve the unity of persons when it emphasizes consistency. But in taking sides, which this emphasis on consistency seems to require, it appears to lose the unity that is required to have both a single self and self-deception.

One of the conditions that Fingarette cites for self-deception--that persons are attracted to dissonant engagements that belong to individuals--is a possible attempt at solving the unity problem. A similar strategy might be at play in Watson's distinction between the judging and desiring parts of the self. The attempt is made to see the fragments as different phases of the same entity. Watson, for example, says that these different motivational systems ought not to be seen as different selves. But there is an additional claim that the self is identified with the judging sys-

tem more deeply, and a possible implication of this claim that what does not belong to persons in a deep sense, does not belong to them at all. The idea of ranking systems, judgment, etc. brings confusion to this issue. We need a nonarbitrary criterion for picking out the impulses in human beings that belong to persons. Any impulses that are chosen seem to be chosen on arbitrary grounds. Therefore, the choice of consistency impulses is arbitrary. Furthermore, the arguments that support the distinction between internal and external are weak as we shall see below in the discussion of Frankfurt's treatment of this problem.

A view like Watson's does not do justice to the extent that persons can be in conflict, nor to the nature of conflict. If conflict is irresolvable, we have a state of ambivalence. This kind of conflict is behind self-deception. Otherwise, persons would have been able to resolve the conflict and to avoid considering self-deception. Self-deception also leaves the possibility that persons purposefully do not want to resolve the conflict, and that they want to be in self-deception. Self-deception could be useful in cases in which persons want to preserve conflicting states. Another point about conflict is that it is not always simply between a rational and an irrational part. A man leaves a business deal in which he has made a huge profit. He is proud of himself and in his euphoria and

self-confidence he drives carelessly, ending in a self-destructive accident. He is both confident and self-deprecating, proud and guilty. That the self-deprecating response ends in having the greater impact does not mean that is is the stronger part of the person, or if it does, it does not mean that the self-confidence was false and that he was really only guilty. It does not mean the other state is not a true part of the person. All of this relates to the question, in relation to self-deception, of the status of self-deceivers' claims to have certain beliefs or feelings. In Chapter II, we said that a sincere claim to have beliefs or feelings is grounds that persons have them. The same argument applies here--preferences can be based on desire. The fact that impulses exist in persons makes such impulses their preferences, at least to some degree. There may be other stronger preferences, but this does not rule out the "irrational" ones from representing persons.³⁵ But there is an attempt made by the integrated theory to justify the view that rational impulses are more representative of persons. The theorists who have been cited as suggesting the integrated theory all define rationality in terms of consistency, but some, like Fingarette and Watson, also include better values in the notion of rational. If they grant, as Watson does, that all the impulses are part of persons, then how will they justify that persons should be identified only with impulses

for consistency? And if they do not grant that all impulses are part of persons, how can they explain self-referential occurrences of conflict like weak will or self-deception as being by, of, about and for the same self.

Frankfurt: Identification--The External and the Internal

Frankfurt makes a distinction between persons and biological individuals, using, not only the criterion of policy-making, but also the criterion of identification, which, he claims marks the internal from the external in persons.³⁶ If we can determine with what values persons really identify, the assumption is that we can determine which poles of the conflict represent their policies. Frankfurt's notion of identification, spelled out in terms of the internal/external distinction, can be seen as an attempt to clarify how persons' policies can be distinguished. But as we shall see, such an attempt does not show how the persons' policies must be consistent. Another problem is that the criteria for determining what is internal, or what really belongs to persons, are difficult to specify, and Frankfurt even acknowledges this. He rejects the approach that identifies real persons, pointing out that what persons regard as their better selves is not necessarily coincident with whom they are or that with which they identify.³⁷ As we shall see in the following discussion, he also

points out the inadequacy of falling back upon persons' attitudes to mark the distinction between internal and external, for attitudes themselves may be either internal or external.³⁸ He finally says that the criterion is decision, and that decision cannot be external. But he is unable to say what is the criterion for decision.³⁹

First, let us look at what Frankfurt means by the internal/external distinction. Frankfurt asserts that we can think of things within us as falling into the two categories of external and internal. He gives the following kind of illustration: The motion of my body may have an external or an internal source. If I fall into someone when the bus lurches there is a sense in which I must answer in the affirmative when asked if it was I who bumped into the other person. It is a motion that belongs to me, according to Frankfurt, but it is not really my action in the sense of being internal to me, i.e., of being that with which I identify. It is part of my history but not part of me as a person. Frankfurt is trying to establish that there are events that are "neither actions nor mere happenings."⁴⁰ There is a difference between the motion of a spider's legs caused by an electric shock and that when it is idly walking. The former is truly passive, but the latter is not really passive or active. Many of our actions are like the idle walking of the spider--this is something we do "mindlessly," it's not really internal.

The same dichotomy that has been applied to our bodies can be applied to our minds, Frankfurt thinks. We may bear witness to certain thoughts and feelings that, although they are within us, are not internal, for we do not identify with them. I may want to compliment a friend and at the same time experience a desire to withhold the compliment because of jealousy. If I choose to act on the former impulse, the latter is external to me--I do not approve of it if I do not choose to act on it. I am merely a witness to this impulse and possible action. This view is similar to Watson's distinction between the two motivational systems of desire and judgment. Desire, we might say, is external to persons as persons, it is part of what Fingarette would call "individuals." (But Watson is making the opposite point, that this is how we can see dissonant impulses as being part of persons.)

As mentioned above, in discussing a possible criterion for identification, Frankfurt sees an ultimate problem in identifying what is internal in a person by the person's attitudes. The person's attitude itself could be either external or internal:

Suppose that a person has mixed feelings about one of his passions: he is aware of having an inclination to approve of the passion and also of having an inclination to disapprove of it. Suppose he resolves this conflict by decisively adopting an attitude of disapproval toward the passion. He may find nonetheless that his inclination to

approve of the passion persists, though it is now external to him and not properly to be attributed to him as his own.⁴¹

But a person's attitude can determine the internality or externality of a passion "only if the attitude in question is itself genuinely attributable to him"⁴²-- that is, if it is the attitude with which the person is decisively identified. But this line of argument, Frankfurt points out, leads to an infinite regress. Conflicting attitudes require higher-level attitudes to cut through the conflict. And if there is genuine conflict, as Rorty notes, it will be duplicated on higher levels.⁴³ Although he acknowledges the difficulty of this view in respect to the infinite regress argument, Frankfurt maintains the position that the criterion for what desires a person has with which he/she identifies internally has something to do with his/her attitudes.⁴⁴ We need, ultimately, an attitude that must be internal, a "decisive" final policy, Frankfurt says, that really represents the person.

Attitudes or identifications mark what belongs to persons. But if one thinks that persons' identifications or attitudes must be consistent, then to appeal to attitude or identification as the final criterion for personhood is circular. It assumes there is a single unified and integrated attitude and that is just what the integrated theory must establish. It rules out by stipulation the possibility that persons

might have conflicting attitudes.

Frankfurt's distinction between internal and external is useful but his criterion for the distinction is arbitrary. That there is such a distinction ought to be granted by any theory of the self. But what the distinction is based upon need not be consistent integration, or persons' real attitudes, or identifications, or approval. If decisions cannot conflict, there must be something--attitudes, values, avowals--in virtue of which this cannot happen. So we are thrown back to the problem that attitudes, etc., are not always internal. We need a criterion for their being internal. This turns out to be something like decision, and we are back into the circular form of reasoning which again leads us to seek a criterion for decision in terms of attitude.

Frankfurt: Decisiveness

Frankfurt believes that the mark of personhood is the ability to have preferences about which of our desires will motivate us to act. He makes a distinction between desires and higher-level desires. Desires are mere forces, but higher-level desires, which are about desires, are decisions. Persons can have desires but not want them to be ones with which they identify. This means that they do not want them to be decisive, they do not want them to motivate them to act. On this view, as Rorty⁴⁵ points out, persons do not coincide with biological

human beings, for there are human beings who, Frankfurt⁴⁶ says, along with all nonhuman animals, do not care about which desires motivate them. He calls such beings wantons.

Frankfurt thinks that persons are identified with what is on the higher level. But unlike Watson and Fingarette, Frankfurt does not require that what is on the higher level represents ideal persons, persons' better values, or moral values:

In speaking of the evaluation of his own desires and motives as being characteristic of a person, I do not mean to suggest that a person's second-order volitions necessarily manifest a moral stance on his part toward his first-order desires. It may not be from the point of view of morality that the person evaluates his first-order desires. Moreover, a person may be capricious and irresponsible in forming his second-order volitions and give no serious consideration to what is at stake. Second-order volitions express evaluations only in the sense that they are preferences. There is no essential restriction on the kind of basis, if any, upon which they are formed.⁴⁷

That persons might be identified with values that are not their better ones suggests that the integrated theory is not bound to a claim that persons represent an ideal. Consistency defines persons but not necessarily consistency of better values. But there is another sense in which the integrated theory cannot avoid being a theory based on an ideal. It picks out one thing that persons do and that thing is something we prefer them to do. It then defines them in terms of it. And the thing it picks out is referred to by Frankfurt himself as being

that which is most important to us. So it turns out that consistency itself is an ideal and it is built into the integrated theory.

Frankfurt's view is that persons are represented essentially by higher level desires. We might call these policies or decisions. This view suggests the integrated theory because it suggests that policies are consistent and that in any conflict there is one consistent set of attitudes that represents persons. But just as there were problems with Watson's view that some of our actions are in spite of ourselves, because they are in spite of our cool and rational judgments, and with Frankfurt's distinction between the internal and external, so are there problems with this view about policies. This view about policies and its problems can in fact shed some light on what is wrong with these other views. The difficulty of ranking our desires into an absolute hierarchy is that, as Rorty points out, our policies shift and our desires get ranked differently depending upon our attitudes--whether we want to strengthen or weaken our impulses--and these attitudes can shift too. A desire that belongs to us in a shallow way at one time can become deeply internal. Furthermore, policies can conflict--we can have several rankings at work within us at the same time.

Another problem that Rorty points out is that it is not always possible to distinguish a desire from a

policy. But the integrated theory assumes such a distinction. If all desires are accompanied by a higher level desire ("let this desire be enacted"), Rorty points out, desires and policies would be isomorphic. Such a result would not help us to understand how desires get ranked when there is conflict. And if only some desires are connected to higher level ones we have the problem of infinite regress and of determining the criteria for the link between certain desires and policies.⁴⁸

Finally, the idea of higher level policies does not help us to understand the problem of conflict in the self because, as Rorty points out, policies will either reduplicate the conflict that exists on the lower level, if they are going to be accurate conveyors of the impulses on that level; or they will be so general or ideal as to misrepresent persons' impulses. In the latter event, the policy resolution that is characterized by consistency is arbitrarily assumed:

The decision which picks out 'the best' desires, and makes those dominant, already presupposes the reordering of personality that it is designed to achieve. If a person is faced by a conflict between his desires, or even between his desires and his policies, he is precisely in the position of not knowing what rationality requires of him.⁴⁹

In referring to rationality here, Rorty is assuming that decision-making and rationality are the same. She is trying to suggest that there might be a meaningful kind of rationality that does not include consistency. In both cases that Rorty points out, conflict remains

and it seems inaccurate to define persons in terms of attempts at unconflicted policies.

Rorty thinks that what is intended in the concept of a person is the concept of responsibility, by which she means agency. A person is a responsible human being. Thus, the criteria for personhood should be those conditions in virtue of which individuals are responsible. On this reading the integrated theory is claiming that persons are responsible only insofar as they are striving for consistency. But responsibility also can be ascribed to individuals when they do not seem to be striving for consistency, as in self-deception or weak will. As Rorty points out, we strive for unity even in dissociated states, and creating conflict may be one manoeuvre towards unity. But

the rational integration of desires is not a person's sole task, nor even necessarily his dominant one. It may be the one that is socially useful in (re)identifying responsible individuals. But there is no a priori reason to treat those characteristics as more essential to a person than the voice that bellows in pain, or persistent habits of refusing integration.⁵⁰

Rorty criticizes Frankfurt's type of view (which she calls "PRAT"--the view that persons are rational agents) on this issue. She argues that disintegrating or dissociated persons are still persons, at least because they still "go through manoeuvres towards unification that are recognizably theirs," and because strategies of integration are not necessarily persons' only

strategies.⁵¹ When Rorty writes of "manoeuvres towards unification" we need not read consistent integration into this. She could be referring to policy-making. The point is that policy-making could be essential to personhood, without this meaning that the integrated theory of the self is true. Persons may be persons by virtue of having strategies ("unification" means that each person is one person), but not necessarily only integrative strategies. Rorty argues that the concept of a person must be able to accommodate the different kinds of conflict of which persons are capable:

Frankfurt suggests that a decisive identification of a conflicted person with one (set of) his first order desires terminates the regress of second order policies. But we need a time qualifier here: the decisive identification may be made in one context, to be later revoked by a different policy.⁵²

Here Rorty is pointing out that policies might be different at different times. She goes on to talk about synchronic policy conflict:

What's worse, a person may commit himself to several, potentially conflicting policies, half-realizing, sometimes self-deceptively, that some of his identifications must undermine one another. Second order conflicts may remain latent or recessive; though only later, they may have the effect of defeating 'decisive' identifications, so that a person's policies often have perfectly predictable consequences that counteract the major lines of identification, though they may suit latent or minor ones. We are familiar enough with patterns of predictable failures and other sorts of self-defeating behavior on the part of highly capable persons. Yet none of these complexities should force us to abandon the concept of a person; unless there were the presumption that such struggles were those of one person, we could not even think of this as the phenomena of conflict.

The difficulty for PRAT lies in specifying the extent to which a person may be conflicted and sun-dered, yet for other reasons still be regarded as identifiably the same person, and not merely the same biological individual.⁵³

Rorty's remarks support the view presented here that the integrated theory does not do justice to the nature of conflict.

Self-Deception

On the one hand, the integrated theory seems right for self-deception because it is because we want to integrate, to eliminate a conflict, that we self-deceive. If being in conflict were acceptable to us, we would turn to self-acceptance when we were unable to eliminate conflict. Thus, the integrated theory seems to provide an explanation for the motivation for self-deception. But on the other hand, self-deception creates and maintains conflict in persons, and this seems incompatible with overriding integrative impulses. The fact that we self-deceive can suggest that we do not care about eliminating conflict--that not wanting to have the beliefs or feelings we end up denying in self-deception is not the same as not wanting to be unconflicted. The fact that self-deception can be useful and adaptive for purposes of self-transformation may indicate that we can strive to be inconsistent (contra the integrated theory) even if the ultimate result is consistency. And perhaps the ultimate goal need not be consistency. It

is even possible that persons prefer self-deception to conflict-resolution--as a way of having conflicting beliefs or feelings that are both desirable to them. Persons might see it as advantageous to have conflicting beliefs or feelings--to love and to hate, for example.

Much of what we do is an example of letting ourselves be overcome by what we do not think is best, by impulses we do not want to be effective. Such impulses play an important role in the dynamics of persons.

The integrated theory provides something that is essential for self-deception--an impulse to resolve conflict and a vision of how to understand conflict in a single self. It gives an explanation for dynamics that require one part of persons to try to overrule other parts, for the apparent difference in status of these parts. What persons want to deny self-deceptively seems to have a different status from their wanting to eliminate it, and from their sincere denial-avowal: I hate my boss but I do not want to. I sincerely claim to like him. Which is the real me? The one who hates, the one who does not want to hate, or the one who likes him? The integrated theory shows these different feelings to have different statuses within persons. Yet there is a nagging question about the arbitrariness of the criterion it proposes and a persistent inkling that the full meaning of conflict is not taken into account by the integrated theory.

Even if there is a ranking of desires and some are deeper, one can still want to be in conflict. Wanting to be in conflict and maintaining conflicting desires do not imply that both desires have the same status. So the integrated theory can be right about the hierarchy of impulses, and wrong about consistency impulses defining persons. Even though the desire not to have a desire, belief or feeling is, in a sense, a policy impulse--the motivating and decisive impulse of self-deception--it does not follow that other impulses are external to persons. The difficulties the integrated theory has in establishing the criterion for persons can be seen in the example of self-deception, where picking one side as representing persons seems arbitrary.

For these reasons, it is useful to look at an alternate theory of the self. A theory that can deal with the difficulties the integrated theory is unable to solve satisfactorily might give an easier account of self-deception. Thus, in the next chapter, we turn to the fragmentary theory.

Footnotes to Chapter III

¹The tradition of such views can be traced back through Kant and Descartes to Plato. Also, see Kierkegaard, for whom purity of heart is to will one thing. Views that emphasize certain interpretations of rationality as the defining characteristics of persons suggest the integrated view. Such views often define persons in terms of their rational choices or their ideal selves. Rorty refers to a kindred view as "PRAT"--the view that persons are rational agents, in Amelie O. Rorty, "Persons, Policies and Bodies," International Philosophical Quarterly, XII (March, 1973), 63-80.

²Fingarette, 1969, 71-72, 75, 87.

³Ibid., 40-41.

⁴Ibid., 138.

⁵A review of Fingarette's book on self-deception that presents useful criticism of his views of self-deception and that points out problems it raises about the self is found in de Sousa, 1970.

⁶Frankfurt, 1976, 239-251.

⁷Gary Watson, 205-220.

⁸Cf. Terry Penner, "Thought and Desire in Plato," ed. Gregory Vlastos, Plato--A Collection of Critical Essays II: Ethics, Politics, Philosophy of Art and Religion (Garden City, New York: Doubleday Anchor, 1971), 96-118.

⁹Frankfurt, 1976, 248-250.

¹⁰Fingarette, 1969, 87-8.

¹¹Frankfurt, 1971, 16.

¹²Ibid., 6.

¹³Fingarette, 1969, 149: Persons develop via "the integration of what is avowed into that achieved synthesis which is the personal self;" also, 84-88.

¹⁴Watson, 215-16.

¹⁵Fingarette, 1969, Ch. IV.

¹⁶Ibid., 71-72.

Footnotes to Chapter III, cont.

¹⁷ Ibid., 149.

¹⁸ Ibid., 85.

¹⁹ Ibid., 138.

²⁰ Herbert Fingarette, The Self in Transformation,
(New York: Harper Torchbooks, 1965).

²¹ Fingarette, 1969, 85.

²² Ibid., 138-142.

²³ Ibid., 138.

²⁴ Ibid.

²⁵ Ibid.

²⁶ Ibid.

²⁷ Ibid., 139.

²⁸ Ibid.

²⁹ Ibid.

³⁰ Ibid., 141.

³¹ Ibid., 139-140.

³² Ibid.

³³ Ibid.

³⁴ Ibid., 140.

³⁵ Rorty, 1972, 404-406; 1973a, 67-9, 73; 1975, 6, 21.

³⁶ Frankfurt, 1976.

³⁷ Ibid., 245.

³⁸ Ibid., 246.

³⁹ Ibid., 251.

⁴⁰ Ibid., 239.

⁴¹ Ibid., 247.

Footnotes to Chapter III, cont.

⁴²Ibid., 248.

⁴³Rorty, 1973a.

⁴⁴Frankfurt, 1976, 246.

⁴⁵Rorty, 1973a.

⁴⁶Frankfurt, 1971, 11.

⁴⁷Ibid., 13, n.6.

⁴⁸Rorty, 1973a.

⁴⁹Ibid., 67.

⁵⁰Ibid., 69.

⁵¹Ibid., 67-68.

⁵²Ibid., 67-68.

⁵³Ibid.

Chapter IV

THE FRAGMENTARY THEORY

Introduction

We saw in Chapter III that the integrated theory claims that there is a constant impulse to become consistent in persons. The fragmentary theory denies this claim. It acknowledges that persons may have impulses for consistency, but it downplays such a possibility and it emphasizes that there are many impulses, none of which are necessarily constant. The fragmentary theory can be inspired by recognition of the plurality and variety of influences that contribute to the formation of persons and to the consequent plurality and variety that results in each individual. Such views of the formation of the self are held by Mead, Freud, and James.¹ Freud also emphasizes inner psychological contributions to the mind's capacity and predelection for plurality of attitudes. His contributions to our understanding of ambivalence, ego and super-ego formation, and motivation yield a picture of persons as being basically irrational and not always driven by consistency impulses. Such theories reveal an enormous capacity in human beings for being in conflict.

James's findings, also of a psychological nature,

suggest that we see ourselves as different selves both in different social situations, and in relationship to ourselves. James seems to suggest that this is quite natural and he does not mention a constant impulse to become consistent.

Finally, Mead concentrates on the role society has in forming persons, and sees the social origin of the self as being indicative of plurality. He also does not mention a consistency impulse.

Such are the kinds of thoughts that lead to a fragmentary theory. The fragmentary theory can now be defined as the view that persons are pluralistic and inconsistent, that they lack a constant impulse to become consistent, that they can accept conflict in themselves, as well as being able to create it, seek it out, and maintain it. A desire to be consistently integrated is just one of other possible motivations of persons, according to the fragmentary theory.

The fragmentary theory is not a multiple self view. But it might be accused of being such because philosophers such as James and Mead refer to our different selves. It also might be accused of being such from the point of view of the claim that it is not possible to be a unified self or to make policies without an impulse for consistency. The fragmentary theory must meet such challenges. It must show that it need not be taken to be a multiple self view and that it can

explain unity and policy-making in the self without appealing to a consistency impulse. These are issues that will be taken up in this chapter.²

The idea that persons are fragmentary is suggested by certain interpretations of Fingarette and Watson. As we saw in Chapter III, these philosophers also suggested the integrated view. But insofar as they point out plurality in the self they bring to mind the fragmentary theory. Fingarette writes about the "community of the self" which is made up of "the rudimentary but unified complexes of reasons-motive-feeling-aim-means-and moralistic reaction."³ He calls these complexes "forms of engagement," or projects. As discussed in Chapter III, his view differs from the fragmentary theory, in that he maintains that these "atomic forms" become transformed when they become part of the "community of the self":

They are significantly transformed by virtue of the mutual interaction and independence, yet they do not entirely lose their distinctive character or quasi-autonomy.⁴

Becoming part of the community of the self for Fingarette means becoming part of an integrated self. The engagements are thus not only united into a single self, but they are united by virtue of a single system or set of policies. This marks persons from individuals. Acting from such "engagements independently as autonomous projects, without integration into the complex unity of a personal

self,"⁵ is regressive. Fingarette thinks that self-deception is such an action. The fragmentary theory differs in its view that it is possible for persons to act independently of any particular principles.

Another view that recognizes plurality in the self is Watson's. His view is that persons may contain different motivational systems--the desiring system and the valuing system. It is this view that enables him to see akrasia as an expression of persons, an instance of one's being "obstructed by one's own will."⁶ He is noting that persons may be motivated by conflicting impulses. But then, as explained in Chapter III, he goes on to imply the opposite--that only the valuing part truly represents persons. Views like this seem to suggest--contrary to Watson's avowed view that persons can be conflicted--that a motivational system (the desiring part in this case) can belong to persons but it does not always represent them. It seems that what issues from the desiring motivational system, for Watson, belongs to persons, but does not really represent them. Similarly, Fingarette disallows engagements that are not consistently integrated with central ones from being part of persons.

Another source of the notion of a fragmentary self comes from remarks about these problems made by Rorty. Like Mead, Rorty emphasizes that our notion of what persons are originates in social practices.⁷ Very

often, it is a notion of what we would like persons to be--unified and integrated--because the ascription of responsibility is easier in this case. Rorty criticizes a view that claims that persons are rational agents (she calls it "PRAT") and one of her criticisms is that conflicts in the self are reproduced on a policy-making level.⁸ In this criticism she suggests the absence of a constant consistency impulse, and her view that disintegrated persons are still persons⁹ suggests that unity or personhood are not the same as and do not require consistent integration:

But a disintegrating person is still, at least in some measure, a person: and dissolved persons, even those who have conflicting interpretations of their own memories, whose desires conflict sharply with their policies, nevertheless go through manoeuvres towards unification that are recognizably theirs.¹⁰

Her views that unresolved conflict is compatible with personhood, that irrational grounds for belief are common, and that our ideal self (an integrated one) is only one of the systems with which we identify, also contribute to the origin and appeal of the notion of a fragmentary self.

The fragmentary theory is suggested partially because of problems raised by the integrated theory, both in itself and as it applies to self-deception. We saw that the integrated theory had the difficulty of giving a nonpluralistic definition of persons. Concepts such as value, avowal, choice, identification, aim, have been suggested as criteria for personhood, but it

seems that what is representative of persons is the result of all of these factors, and that they do not necessarily coincide, or yield consistency. What is really representative of persons cannot be narrowed down always to a monolithic point of view. For one thing, the determination of what represents persons at any one time depends upon one's perspective. Are we looking for values that are validated by their behavior? If so, are we looking for typical behavior or the rarer variety? Shall we determine persons' dominant values by what is avowed, by identifications, by choices? The different approaches available yield different and often conflicting answers. Should we pay attention only to those things to which one commits oneself in "a cool moment," as Watson does, or to what is the result of higher-order decisions, as Frankfurt suggests? This kind of strategy will yield one kind of answer to the question of what persons are but it seems to rule out arbitrarily other eligible candidates. Finally, there is the possibility that from a single perspective what represents persons will be seen as multiple and inconsistent. This may be what happens in self-deception, where there seems to be perpetuation and/or creation of conflict.

The integrated theory is faced with two alternatives: it can either claim that apparently conflicting beliefs and feelings are really consistent or that

they really conflict but are subserviant to a policy of consistency; or it can claim that not all of the beliefs or feelings are really parts of the self. The first alternative is difficult because it requires denying obvious cases of irresolvable or purposefully maintained conflict. This strategy seems to assign dominance to consistency impulses arbitrarily. The second alternative is equally arbitrary. We find this kind of move in Fingarette's claim that when consistency breaks down, so does the self;¹¹ and in Frankfurt's relegation of whatever is not part of the consistent system to the "external."¹²

A parallel set of alternatives arises for the question of self-deception. The integrated theory must claim that self-deception is either an attempt to preserve or achieve consistency or that it is a violation of a pre-established consistency. Fingarette's view is the latter, that self-deception is a violation of an integrated self.¹³ He believes that self-deception is only possible in a self that is integrated. But he also sees self-deception as the former view, that self-deception is an attempt to preserve or achieve consistency; it is because of the integrity of the self--because of a dominating concern for integrity, which only an integrated self can have--that we self-deceive. Self-deception is here seen as a self-defeating attempt at integration. The inconsistency in Fingarette's view

of self-deception is that he implies that it is both done by the self and by an engagement external to the self. He inconsistently assigns the project of self-deception to both a fragment of the self and, thus, to an engagement that is not part of the person, and to the integrated person. This is the flaw of the integrated view. It assigns responsibility to either a narrow part of a person--what coincides with consistency impulses--or to a system that is not even part of the person. In the case of self-deception, since each alternative seems to require us to omit essential parts of the self from our explanation of the dynamics of self-deception, it is plausible to turn to the fragmentary theory, which may allow that self-deception is by and of the same self and that it grows out of, preserves, and expresses inconsistency.

The appeal of the fragmentary theory is that it is compatible with many of our experiences and theories about ourselves and others. We notice that in different circumstances and with different people we are different in the way that we think of ourselves, in the roles that we play, and in the way we experience the world. If we put all these different attitudes together we will have an inconsistent whole. An integrated self seems to be nothing more than a goal or an ideal. The fleeting moments when we seem to ourselves to be integrated often turn out to have been illusions. If it turns out that

integration is one of the goals of persons, this might be due to the fact that persons are fragmentary. Theories such as Fingarette's, that account for the nature of persons in terms of a developmental theory, might actually imply the fragmentary theory, for it does not seem that integration is the only impulse persons have.¹⁴ Fingarette seems committed to an implication that is puzzling in view of his developmental theory of the self. The implication (i.e., having consistency as one's constant impulse) is a precondition for being a person and thus for having consistency as one's goal. The difficulty of this view is this: how can a developing self--one that develops from an individual to a person, as he says--have a dominant goal of integration. It seems that it must already be integrated (have this goal) to do this; but then it would already meet the criteria for personhood. Fingarette gets around this by saying that the person is already integrated, not the individual. What persons build themselves out of are the engagements of themselves as individuals. But if persons develop from individuals there must be some initial dominating-integrating impulse. Perhaps such an impulse exists in individuals--but then integration is already established and there is no need to develop into persons; persons exist from the beginning--but then there is no justification for the distinction between individuals and persons. To posit integration as the guiding impulse

in a developing self is to rule out a previous stage where there is no such impulse, and it makes it difficult to understand cases of conflict-creation or maintenance that seem to be counterexamples. If the developing self were already integrated, in the sense that consistency were its goal, integration would not be so difficult.

There are obvious examples of deliberately avoiding integration, as when we are morally inconsistent for opportunistic or self-indulgent reasons, or due to weakness. People often do not want to be integrated, and this may not be due always to moral weakness or confusion, but even then it is no less a part of the self.¹⁵ It is conceivable that people cultivate fragmentation in order to keep alive parts of themselves that they regard as valuable or useful, that give them pleasure, and that are sources of enrichment. Cultivation of inner fragmentation may be a way of adapting to society and the many roles it imposes upon persons. Conflict situations are conducive to increased fragmentation. If the self were integrated such situations would make it difficult for persons to survive as persons. And if people have the beliefs that they are integrated they might experience less self-acceptance in such situations. Such attitudes could be, not only unrealistic, but destructive. One might be perpetually in the position of thinking of oneself as having fallen apart, as not a person, as out of touch with one's

true self. And one might struggle in vain to make oneself consistent, thinking that this is in keeping with one's true nature and with mental health. If the self is fragmentary, perhaps self-acceptance can lead to greater harmony in the self, more realistic and constructive goals, more honesty vis à vis oneself, and an informed attitude towards self-deception that might include using it constructively. One objection to the fragmentary theory might be that it makes it more difficult for others to know us, and, thus, makes us less predictable. But even this would not necessarily follow. If it does, in that sense it may be impractical. But, nevertheless, it may be still what the self is like.

We often discover that we are not integrated. An example is the sense of alienation we sometimes feel from part of ourselves. We may look upon certain of our own avowals, identifications, choices, aims, ideals, with a feeling of remoteness or estrangement, as if they belonged to someone else, because they are inconsistent with other attitudes that we also have, and with which we might feel a stronger identification at the moment. These other points of view represent us too, even if we identify more strongly with others. One wonders how one could have been so silly yesterday, or written those strange words ten years ago, and in the present one may feel disconnected from the sound

of one's own voice. Proust writes of these kinds of feelings of self-estrangement:

It is not because other people are dead that our affection for them grows faint, it is because we ourselves are dying. . . . My new self, while it grew up in the shadow of the old, had often heard the other speak of Albertine; through that other self . . . it thought that it knew her, it found her attractive. . . . but this was merely an affection at second hand.¹⁶

Feeling similar disclaiming tendencies about present elements of oneself may be one of the things that occurs in self-deception. It should be noted that when we disclaim parts of ourselves as in the examples cited above we do not mean that we literally think these beliefs or feelings are not ours. It is because we realize they are ours that we remark upon the fact that it seems strange that they seem remote, as if they belonged to different persons. The point is that the fragmentary view allows persons to act on the basis of certain beliefs or feelings at a certain time, and still to be represented by all of their beliefs or feelings at that time, no matter how remote they may seem or inconsistent they might be. Examples of the claim that we see ourselves as having different and possibly conflicting beliefs and feelings need not all be in terms of self-alienation. As we shall see in James's and Mead's views, we may see ourselves as a conglomeration of different and conflicting roles with correspondingly different and conflicting beliefs and feelings.

It may be that we cultivate fragmentation, sometimes, perhaps, as a practical matter, while giving lip-service to integration. In this way we preserve ourselves as we are, and sometimes, as we want to remain. Fragmentation is natural. Integration is artificial. In this sense, self-deception may be what Rorty calls an "adaptive strategy" but not with integration as a goal.¹⁷ Rorty has a pragmatic interpretation of self-deception.¹⁸ She suggests that self-deception may be an attempt to preserve our conflicting beliefs or feelings while still presenting an unified facade to the world. She also suggests that self-deception might be a method of self-transformation. Thus, self-deception can be seen as involving recognition of fragmentation and a process that perpetuates it, in situations that call for at least the appearance of integration. The difficulty we have in achieving integration may be the result of the fact that it is not possible. Integration--or the illusion of integration--may seem necessary to assume for practical social reasons but it may be an impossible achievement. Self-deception may be an attempt to deal with the conflicts that exist between the forces of integration and fragmentation.

In the next section of this chapter, two views of the self are presented that are suggestive of the fragmentary theory and of how its major problems can be handled. The views presented are those of George

Herbert Mead and William James. Both Mead and James have views that suggest to me what might be meant by the notion of a fragmentary self. Mead emphasizes how the origin of the self contributes to its fragmentary nature, but he does not give an adequate account of how the different fragments of persons can be seen as all part of single persons. At this point, it is useful to turn to James, who also suggests a rich picture of what a fragmented self is like, but who, even more importantly, has a theory that suggests to me a starting point for a possible explanation for the problem of unity in such a self. Thus, James's view suggests a way to understand how a fragmentary self can make policies that reflect purpose. These views are not explicitly employed by James to support the kinds of claims I am making, but they are suggestive of such claims. It is necessary to explain how a fragmentary self can make policies so that we can meet the objection that a fragmentary self might be a multiple self. We must show how a self can be complex, pluralistic, inconsistent, and yet, not multiple.

The idea of a fragmentary self is meant to accommodate the things for which the integrated view is inadequate. There is no official version of it, so that we find that various kinds of theories can be instances of it. Basically, though, it represents a rejection of a constant impulse for consistency in persons.

Finally, although the fragmentary theory does not subscribe to the view that there are literally many selves or persons per individual organism, sometimes I shall refer to the fragments as selves because that is how James and Mead refer to them. But strictly speaking, "self" and "person" are synonymous in my usage; therefore, we may talk about the selves of persons, as long as we keep in mind that by "selves" we mean roles or attitudes and not whole selves or persons. The problem of what kind of entities the fragments are is discussed at greater length later in this chapter.

George Herbert Mead and the Fragmentary Theory

No discussion of the fragmentary theory of the self would be complete without acknowledging its indebtedness to the views of Mead. Rorty, who emphasizes the fragmentary, pluralistic nature of the self, cites him as an influence too.¹⁹ It is useful to study Mead for our purposes, because in giving a picture of the self as consisting of different "selves," and in showing how these "selves" originate and how they are structured, he presents a picture of persons that suggests the likelihood of conflict and he does not posit a consistency impulse. Mead has a vision of what persons are like that includes plurality and conflict, but from which the idea that persons are characterized by a constant impulse for consistency is conspicuously absent. He also

tries to give an explanation for the unity of the self. As we shall see, his answer to this problem is incomplete and inadequate. Pointing out how Mead meets or fails to meet these problems indicates what the fragmentary theory must account for in order to be distinguished from a multiple self view.

Mead builds his notion of the self in the following way. He first distinguishes between the organism and the self.¹⁹ This distinction is reminiscent of Fingarette's distinction between the individual and the person. Both are distinguishing personhood from a pre-personhood stage. They share a developmental view of the self. But for Fingarette the distinction is based upon the presence or absence of a single integrating mechanism, what we have been calling a constant impulse for consistency. As we shall see, since Mead's self is created by social interaction, from which it incorporates systems, it is unlikely that it would contain only one system, but if a person had very narrowly focussed social interactions with an unusually well-systematized segment of society, the possibility that such might be the result is conceivable.

Mead claims that selves are products of society and that they thus cannot exist unless society exists. An organism has experiences in the sense of having bare sensation, but these experiences are not organized and the organism cannot reflect upon them. Social inter-

action, however, causes the organism to "import" the values of society and the roles of its members. This gives the organism the ability to become an object to itself, ie., to reflect about itself, and this is the origin of the self. Mead thinks that the essence of selfhood is self-consciousness. Thus, he thinks that the most important question in understanding the self is "how can an individual get outside himself (experientially) in such a way as to become an object to himself?"²¹ The answer is found in "the process of social conduct or activity in which the given person or individual is implicated." We interact with others, and we internalize the process of interaction. The result is that we may, so to speak, interact with ourselves. The process of communication sets up "others" within us. Internal dialogue, or introspection, is thereby made possible and this is the origin of the self. Thus, "the self, as that which can be an object to itself, is essentially a social structure, and it arises in social experience."²² With the internalization of others, the two aspects of the self--the "I" and the "me"--come into existence. The internal dialogue is between these two aspects. The "me" is constituted by the internalized others and the "I" is the individual's spontaneous response to the "me."²³

Another difference between Mead's self and Fingarette's person is that for Fingarette the person

is not a social structure, but mainly an internally created or chosen one. Thus, we see that another difference between Fingarette and Mead is that for Mead systematization in the organism is accompanied by self-awareness, and both are essential for selfhood. But for Fingarette, self-awareness can exist prior to the special kind of self-awareness that marks personhood. The special kind of self-awareness is characterized by an impulse for monolithic systematization; for Fingarette, a person is a single system. For Mead, systematization in the self does not refer to a single consistent system, but to many systems, depending upon the many roles that are internalized. To the extent that Mead's notion of systematization refers to the organization of the self as a whole, it tends to mean unity, but not consistently integrated unity.

Fingarette recognized two processes that construct the self--discovery and choice. These two processes correspond to two kinds of elements in the self--those that are determinants, things persons find within themselves; and those the persons create or that are the result of cultivation, choice, or transformation. Both Mead and Fingarette think that the selves (the individual and its engagements, for Fingarette) consist of given, passive determinants, and novel, spontaneous, active forces. It is the latter kind of element that Fingarette identifies with real or achieved persons. The other kind of element

is seen as the raw material upon which persons, guided constantly by a consistency impulse, work to create themselves.

Mead's view differs in that for him persons are just as much constituted by what he calls the "me," i.e., by the given elements they find within themselves, as by their responses to the "me." For Mead, these two elements, the "I" and the "me," always appear together. Another difference is that the "I," the part of the self that responds to the "me," is not represented by a consistency impulse. The "I" systematizes what it finds in the "me," but not necessarily into a single system. So we see that Mead also believes that the inner self contributes to what the self is, as we shall see when we discuss his notion of the "I" more fully.

At this point Mead seems to have provided a picture of a self that has two features: it is constituted by social roles and other elements incorporated from society; and it has the ability to reflect about itself. As will become apparent, there are some problems in interpreting what Mead had in mind about how the "I" functions and this will be discussed later.

Returning to Mead's exposition, we see that from the fact that the self is a product of society, it follows that it is split up into a number of "selves." Mead explains:

We carry on a whole series of different relationships to different people. We are one thing to one man and another thing to another. There are parts of the self which exist only for the self in relationship to itself. We divide ourselves up in all sorts of different selves with reference to our acquaintances. We discuss politics with one and religion with another. There are all sorts of different selves answering to all sorts of different social reactions. It is the social process itself that is responsible for the appearance of the self; it is not there as a self apart from this type of experience.²⁴

We see that there are many possibilities for plurality and the development of conflict within one individual. Mead goes on to explain that the kind of unity within a self depends upon that in society:

The unity and structure of the complete self reflects the unity and structure of the social process as a whole; and each of the elementary selves of which it is composed reflects the unity and structure of one of the various aspects of that process in which the individual is implicated.²⁵

There may be many elementary selves, reflecting aspects of society, and forming the complete self, which is "a reflection of the complete social process."²⁶ Since society is not organized like a consistently integrated self, as described by the integrated theory, it follows that on Mead's view the unity of the self is not an integrated unity.

Mead thinks that two different kinds of incorporation take place. One not only identifies with attitudes of others towards oneself, but also with "the attitudes of the organized social group to which"²⁷ one belongs. Mead's term for the organized social group insofar as it affects the individual is "the gen-

eralized other."²⁸ These two ways of being connected to and influenced by society--through individuals and through groups--are distinguished by Mead via a parallel distinction he makes between play and games. Play and games are two ways that individuals interact with society; they are two processes by which we internalize roles and values. Play, indulged in by children, is the activity of the early stages of the development of the self. It involves random and aimless imitating of the roles of certain individuals in their environment. To become a complete person, one must develop to the point of participating in games. Games involve incorporating systems, the structure of society and its values, as well as roles and attitudes of others. Games require more selection on the part of the individual and an organizing capacity, so that the individual roles one has earlier incorporated are now seen as part of the generalized other. Games are the ultimate mechanism by which the individual organizes itself into a person.²⁹

In the first stage of development, then, the individual's self is constituted simply by an organization of the particular attitudes of other individuals towards himself and toward one another in the specific social acts in which he participated in them.³⁰

Such a self would think of itself as someone about whom others think well or ill, to whom others respond in certain ways, and so on. One's self-image would be constituted by one's impressions of the opinions others have

of one and would be tied to specific situations. One would see oneself as a being who interacts in certain ways with certain individuals and who is thought of in certain ways by these individuals. But these self-perceptions are not organized into a system or a pattern in this early stage of the self. Individuals do not think of themselves as consisting of different roles and they have no concepts of how these different roles affect one another. At this point our existence consists of

a simple succession of one role after another. . . .
The child is one thing at one time and another at another, and what he is at one moment does not determine what he is at another. . . . [The child] is not organized into a whole.³¹

By contrast, the ultimate stage in the development of the self is a unity because it places the individual roles it incorporates in the context of the social organization as a whole. This requires that the individual organizes and generalizes the individual roles it incorporates. The result is a self that is a reflection of society. The individual, in essence, puts the atomic roles it has imitated into a systematic organization. This gives it a definite personality and it is in virtue of this systematization that it, as a unified self, reflects society.

If we look at this view of the self as a fragmentary theory we see that fully developed selves experience themselves from the points of view of different roles and attitudes, as well as from the points of view of

social institutions and values. Thus, I am as if I were one person when I think of myself as and organize my actions from the perspective of a student, and I am as if I were quite another person when I think of my student role in its relationship to my other roles and to the group. In the former case there is the possibility of experiencing myself in an isolated way purely in terms of how certain others respond to me and interact with me as a student. In other cases, I might put this role into the context of the social system as a whole and think of my role as a student as part of my systematic self which is made up of many roles. I may, that is, think of myself as an organized complex whole, or I may experience myself as isolated parts of that complex whole. In the latter case, I experience myself as a simple whole. Having the perspective of each of these different points of view is possible and it is possible to operate as the same self with each kind of perspective, even if they conflict. This is what is meant by the fragmentary theory of the self.

Mead provides for the unity of the fragmentary self in his notion of games. We import roles and systems and may experience ourselves from the points of view of any of them. Each one may seem like a different self. But they are all unified into one self by an overall organization which reflects the organization of society. Thus, each person is a unified whole constructed out of

each individual's attempt to systematize what it internalizes. The two factors that contribute to unity, then, are the systematic nature of society, and the organizing and reflective capacity of the individual. Mead refers to the "ongoing processes of the individual" which try to organize what is internalized into a unified self on the model of the organization found in society. This organizing capacity need not be seen as a consistency impulse. We incorporate many systems but they are not all part of one overall system, save that called society. And society is not a system that is guided by a drive to eliminate inconsistent systems. According to Mead, society is seen as systematic when its various structures and institutions are seen in terms of how they affect one another. The organizing capacity of the self that he speaks of may be interpreted as an attempt to make all internalized elements parts of the self in the sense of having them come to be understood in terms of how they affect one another. They are seen as a whole in that sense (just as society is seen as a whole in that sense), in contrast to the child in whom "what he is at one moment does not determine what he is at another."³² This kind of organization or "systematization" need not be thought of as a consistent system. It is simply that individuals may see themselves as wholes; they may have a vision of how all of their "selves" affect one another or exist as part of a whole. They may also identify

themselves successively with different roles, thus seeing themselves from the points of view of different "selves" at different times.

We just referred to the two factors that contribute to the systematization of the self--the internalization of social systems and the "ongoing processes" of the individual that perform the function of systematizing what is internalized. The entity that performs these processes is what Mead refers to as the "I." The "me" is the part of the self that has been internalized from society. It thus represents what is conventional within the individual. The "I" is the part of the self that reacts to the "me." It contributes what is spontaneous and novel in the individual. Both the "I" and the "me" come into existence simultaneously. They are the two aspects of the self. Mead says they never appear without one another. The internal dialogue that marks the existence of the self is between the "I" and the "me." When the organism gains self-consciousness, what it is aware of is the "me," and what is aware is the "I." The "I" is never an object--it is the part of the self that reflects. It is thus a part of the self that we can never fully experience. We experience it only in retrospect, through memory, after it has become a "me": "The 'I' of this moment is present in the 'me' of the next moment."³³ It gets into the person's "experience only after he has carried out the act. Then he is aware of

it."³⁴ The nature of the "I" is the reason we never fully experience ourselves. We never know what the contribution of the "I" is until we act. It is not a given, like the "me." While the "me" is wholly calculable, the "I" is not. The result is that there is always something predictable and an unpredictable element in the persons. We discover our values and ideals when we act. The "I" contributes what is unique in every individual.³⁵

Mead talks as if the "I" is the part of each "self" that systematizes. But since Mead thinks that we not only systematize each "self" but the self as a whole, the "I" must also be the entity that performs this procedure. This is the point at which problems arise for Mead and for a fragmentary theory. We want to save such a theory from falling apart into a multiple self view. To do this we need a way of seeing a fragmentary self as being a single self. Mead's "I" systematizes the self, but we have been emphasizing that there are many different systems in the self and that it follows that the "I" is not governed by a single system of values. We want to preserve this plurality of values when the "I" takes an overview, surveys all "selves" and systematizes them in the sense of seeing how they affect one another. Thus, there may be different overviews the "I" can take, different ways it can see the self as a whole. But the crucial thing that this means is that the "I" that surveys can systematize them in a way that does not involve making

them consistent. But to preserve unity, the idea that this is a single self, we want the "I" to also be simple, nonpluralistic, not able to manifest itself via conflicting attitudes. The point is that Mead does not make it clear why the self that sees all the "selves" as being systematized in a particular way is the same self as one that sees them as being systematized differently or as one that sees itself from the perspective of any of the selves individually.

We can see the source of this problem in Mead's view as being due to (1) a failure to provide an explanation for the unity of such a self and (2) building too many functions into the "I," so that it is an ambiguous concept. The "I" is seen as the systematizing part of every "self" but it is also seen as the agent of systematization of the entire self. The latter role could be interpreted in two different ways. It could mean that the "I" is the subject of the "selves"; and it could also mean that the "I" is an agent of unity in the self as a whole. But it is difficult to see how it could be both. How could the self be unified by one of its "selves" and not be eminently identified with that one? The other "selves" would seem to take on a secondary status, to be subsumed under the higher-level or surveying "self." Consider the possibility that such a "self" is not represented by a consistency impulse, like an integrated self. But if it is not an integrated self, and there

are different "selves," and the present surveying "self" is no more representative of the person than each other "self," we still want to know in virtue of what the "selves" all belong to a single self and why they do not fall apart into different persons. Further, the surveying "self"--if it is a role--must also be an integrated system. So, if the surveying and unifying "I" is a "self," in this sense, this means that all the "selves" of the person are consistently integrated and we have an integrated theory. So we need a way, on the fragmentary theory, to allow for the fact that persons can survey their "selves" and see them as a system in the sense that Mead posits, i.e., insofar as they affect one another, but not as an integrated system. Answering this question will also show how a fragmentary self can make policies when it is in conflict. Mead implies that he thinks this can be done, for he does not rule it out, but he does not show how.

The other alternative, that the "I" is an agent of unity in the self, could be what Mead means when he speaks of the "I" as the manifestation of the "ongoing impulses" of the person that systematizes what is incorporated from society. He seems to be saying that the different "me's" all belong to one person because of this systematizing "I." The problem we have been pointing out is that there seems to be a difference between this "I" and the one that manifests itself in

each "self" or role as an integrating force of that "self" or role. Mead fails to spell out this difference. To the extent that this is not spelled out, the fragmentary theory is left with a problem. Mead is not explicit about these two different kinds of "I's." Nor does he say that there is an explicit distinction between the kinds of systematizing that each one does. I am assuming that because he talks about the self as being made up of different "selves" that are incorporated systems, and that are further made into systems by the "I," that each self is a consistently-integrated system, what he calls a social role. The idea that we identify ourselves with different social roles suggests that we think of ourselves in each role as being consistent. Mead's claim goes beyond how we think of ourselves, though. The idea of a social role is the idea of a consistent system. A role requires a certain kind of behavior and reflects or expresses a single system of values. The "I," in its response to each role, contributes its values and attempts, according to Mead, to resolve any conflict between it and this particular "me" or social role. The "I" seeks harmony with society--internalized or not--and will attempt this by transforming the self or society if necessary. Thus, it follows both because of the nature of social roles and of the "I," that each "I"- "me" system (each "self" or person in a role) is an integrated system.

The important point is that in this context the "I" as a systematizer is governed by a consistency impulse. But when Mead talks about how the "I" systematizes the person as a whole he says it does a number of different things, none of which imply a constant consistency impulse, although some come close. Mead might be taken to be holding the view that the "I" is guided by a consistency impulse because he does assign it the function of transforming the self, of achieving harmony, and of representing ideal values. But he ultimately seems to be committed to a view that such impulses in the self are not necessarily dominating ones. Also, when he speaks of transforming the self with the aim of harmony, he means harmony with society, not internal consistency.

Mead's view does not succumb to the criticism that it is really an integrated theory. There is no constant consistency impulse in his notion of the self. It is also clear that he does not think of the self as a multiple self. Somehow all the different "selves" are systematized into a single self. They are a system in the sense that they are seen by the self in terms of how they affect one another. Such a notion of systematization does not entail consistent integration, but it is, as it stands, vulnerable to the criticism that it is a multiple self view. Although Mead claims to be talking about a unified, single self, he does not show

how this can be. This means that we cannot see how such a self could make policies. The problem vis à vis self-deception would be that self-deception requires a single self, and the fragmentary theory must address itself to this problem. We shall see in the next section that James's view of the self is suggestive of some concepts that could be useful in solving this problem.

William James and the Fragmentary Theory

James's theory of the self is like Mead's in that James talks about the different "selves" that make up a person. But James can be interpreted as having a view that shows how such a fragmented self can be unified. His distinctions between the "I" and the "me," and particularly, between the "I" and the spiritual "self," respond to this problem in a way that Mead's work does not. Mead explains that he is not interested in

the metaphysical question of how a person can be both "I" and "me." The question is: "where in conduct does the "I" come in as over against the "me"?³⁶

This difference in orientation marks a major difference between Mead and James. James's view is suggestive of a way that we might think of a fragmentary self that would allow it to make policies, to be one self, and thus, to be capable of self-deceiving. So we turn now to an examination of James's view, first to see what it contributes to the idea of plurality in persons, and then to see how it suggests a way of dealing with the

problem of unity.

When referring to James's different "selves," quotation marks will be used to distinguish such entities from the self, which is composed of these "selves." James, however, does not use this convention. Therefore, in passages quoted from James, it will be noted that he uses no grammatical convention to distinguish these two different senses of the concept 'self'. But in the present discussion, such a convention is instituted.

James: The Material and Social "Selves"

An example of the view that persons have experiences and act from varying points of view can be found in James's theory of the self. James talks about the empirical "self," which he calls the "Me" and the judging "self," the "I."³⁷ The "I" is the principle of unity in the self, and we shall discuss it later. The empirical "self" is constituted by material, social and spiritual "selves" and each one of these takes many forms. Thus, the possibilities for fragmentation are many.

The empirical "self" is the self as we experience it. Materially, we experience ourselves through such things as our bodies, clothes, families, home and labor.³⁸ Of course, the self is not literally, say, a part of one's body, a material object, one's labor, home or family. But the self can view itself and the world and it can operate from the points of view of such

material things. It is the self's ability to do this, I believe, which forms the basis of James's claim that the self consists of many "selves" and that one of our kinds of "selves" is a material "self." This means that from the point of view of our material "selves," we can act, and have sets of beliefs and feelings. This, I believe, is what James means when talking about the different "selves." There are three major kinds--material, social, and spiritual. And each person may have many "selves" of each of these types.

James explains that we have certain impulses and behaviors in respect to our material "selves."³⁹ When any of our family members die

a part of our very selves is gone. . . .If they do anything wrong, it is our shame. If they are insulted, our anger flashes forth as readily as if we stood in their place.⁴⁰

When we think and act from the points of view of family members, we are doing so as material "selves." If such loyalties should conflict several of our material "selves" would be in conflict. Each one by itself is, however, a systematic unit.

Another kind of material "self" centers around our affectionate feelings for our home. Persons might be sensitive to criticism about their homes, James says. We might also imagine that this is the "self" that seeks to make the home comfortable and attractive. Another kind of material "self" is exhibited in persons' impulses to watch over their bodies and to deck them with clothing.

Sometimes we identify ourselves so intimately with our clothing that we would, James says, rather have beautiful clothes than a beautiful body. That persons might be capable of having a system of values with this orientation shows the extent to which fragmentation might occur. To the extent that we systematically focus upon bodily matters, either by focussing on clothes or on the body itself, we are functioning from the point of view of a material "self." Finally, James notes that we have an impulse to collect property and that the collections we make "become, with different degrees of intimacy, parts of our empirical selves."⁴¹ This commonly occurs in respect to the wealth that is derived from our labor. But in some persons, material "selves" that are manifested by acting from the points of view of their possessions are more prominent than in others-- as in the case of misers.

Loss of any of the material "selves" results in our experiencing

a sense of the shrinkage of our personality, a partial conversion of ourselves to nothingness.⁴²

This is reminiscent of the Marxist philosophy that when the labor of persons is exploited they suffer a loss or alienation from themselves. And Rorty refers to societies in which "a chief may be treated as the embodiment of a tribe."⁴³ In such a situation the loss of members of the tribe would be experienced by the chief as a partial

loss of the self.

A related idea is found in Proust, as quoted by Parfit in an article in which he tries to establish that we are different "selves" at different times.

Proust writes:

Our dread of a future in which we must forego the sight of faces, the sound of voices, that we love, friends from whom we derive today our keenest joys, this dread, far from being dissipated, is intensified, if to the grief of such a privation we reflect that there will be added what seems to us now in anticipation an even more cruel grief: not to feel it as a grief at all-- to remain indifferent: for if that should occur, our self would then have changed. It would be in a real sense the death of ourself, a death followed, it is true, by a resurrection, but in a different self, the life, the love of which are beyond the reach of those elements of the existing self that are doomed to die. . . .⁴⁴

Proust's passage expresses the loss of self we feel when we lose a loved one. We mourn, not just the person lost, but the loss or death of our own self. For the self that was attached to the lost person will cease to exist as the result of this death. This is a more complicated notion than what James is trying to express. For James, we feel things to be a part of us, and so we feel a "conversion of ourselves to nothingness" when we lose them. For Proust, the things we lose are not seen as a part of ourselves, but in losing them the self that loved them ceases to exist and it is this loss that we mourn. In both cases we lose a part of ourselves, a set of beliefs or feelings ceases to exist or becomes more remote. In slightly different but sufficiently similar ways both

James and Proust, indicating that when we identify with material things, this often includes other people, and pointing out that we thus feel a loss of self when these things are lost, are illustrating the fragmentation of the self. James emphasizes that the material things with which we identify are experienced as parts of the self. The self seems to shrink when they are lost and is affected when they are altered. Proust also notes the idea of the shrinkage of the self when he points out that when loved ones die the part of us that loved them eventually dies too. When we mourn the loss of a loved one, we simultaneously mourn the death of one of our selves-- in so doing, we experience a sense of shrinkage of ourselves.

James's kind of shrinkage is an indication of fragmentation. An additional indication, suggested in the above-mentioned article by Parfit, is that we may feel distance from some of our "selves." Parfit makes this point in relation to past "selves." But the point might also be made in reference to present ones. If I am seeing the world and acting from the point of view of one of my material "selves" at the moment, my other conflicting material "selves" might seem quite alien to me, and might not be recognized by others. For example, Proust suggests that my mourning "self" will one day seem quite alien to me. The possibility of this kind of distance and of James's kind of shrinkage are indications

of fragmentation.

The second kind of empirical "self" that James mentions is the social "self." The social "self" is the most complex, and James says, the most interesting.⁴⁵ The social "self" is first of all "the recognition. . . (one) gets from his mates."⁴⁶ This leaves open the possibility of having as many social "selves" as there are others in one's life: "A man has as many social selves as there are individuals who recognize him and carry an image of him in their mind."⁴⁷ This can be refined, James claims to the statement that one "has as many different social selves as there are distinct groups of persons about whose opinion he cares." This point is reminiscent of Mead's claim that we internalize both the attitudes of other individuals and of groups.⁴⁸ Both Mead and James emphasize that different people, groups, and social situations bring out different "selves." Contact with society also brings out the self's active participation in maintaining its fragmentation. People show different sides of themselves to different groups, and they actually cultivate the separate roles. Goffman writes about how we establish what he calls "audience segregation," by which "the individual ensures that those before whom he plays one of his parts will not be the same individuals before whom he plays a different part in another setting."⁴⁹ James says that the influence of attitudes of others on us and our corresponding res-

ponse leads to

what practically is a division of the man into several selves; and this may be a discordant splitting, as where one is afraid to let one set of his acquaintances know him as he is elsewhere; or it may be a perfectly harmonious division of labor, as where one tender to his children is stern to the soldiers or prisoners under his command.⁵⁰

Here James is pointing out our capacity for fragmentation and how various social roles can bring this out. Social expectations and roles can result in the following kinds of expression:

"As a man I pity you, but as an official I must show you no mercy; as a politician I regard him as an ally, but as a moralist I loathe him. . . ." ⁵¹

Like the material "self," the social "self" can produce a sense of loss or a feeling of "nothingness." James's most interesting example of this is in his note about the social "self" one has vis à vis the person with whom one is in love:

The most peculiar social self which one is apt to have is in the mind of the person one is in love with. The good or bad fortunes of this self cause the most intense elation and dejection--unreasonable enough as measured by every other standard than that of the organic feeling of the individual. To his own consciousness he is not, so long as this particular social self fails to get recognition, and when it is recognized his contentment passes all bounds. ⁵²

In a similar spirit R.D. Laing points out that our self-images fluctuate with our images in the minds of others, especially loved ones. ⁵³ Such examples show that the recognition others give us affects our social "selves" because it affects our beliefs about ourselves. If

our self-images are affected by the opinions of others, our behavior in these roles is also likely to be affected. Thus, one might carry out one's role as a lover, daughter, or mother, differently, depending upon one's self-image. And the self-image is very sensitive to the reactions of others. We want to be recognized by others in specific ways--we want certain roles recognized. Thus, we want the person we are in love with to relate to us as a lover. To the extent that this role is not recognized by our lover we feel annihilated--one "is not," as James says.

James's ideas about the empirical "self"-- the "me"-- are very much like Mead's view that the self is composed of social roles. They both emphasize that social interaction constitutes the self. The significance of this for the fragmentary theory is that it underscores the notion that the self is fragmented. The view that the self depends upon social interaction and internalization of social elements suggests a fragmentary theory because society is pluralistic. Of course, we need more than this for a fragmentary view. We need a self that does not have a constant impulse for consistency. But what we have so far from James and Mead, in terms of how society contributes to what the self is, gives us part of what we need. As we shall see later, James gives us the rest, because he does not posit a constant consistency impulse in the self, and as we have seen, nor

does Mead.

James: The Spiritual "Self"

When James talks about the spiritual "self," he says he means the empirical "self" as thinker, the subjective "self." There are two ways we may consider ourselves as thinkers: the abstract strategy involves dividing the spiritual "self" into "psychic faculties or dispositions such as the ability to argue and discriminate, moral sensibility, conscience, and will," identifying ourselves with each in turn.⁵⁴ This strategy is abstract because such faculties never occur in the self in isolation from each other or from the other "selves," but they may be thought about that way. The other method for considering our subjective "self" is as our entire concrete stream of consciousness or as the present section of that stream. James says that the spiritual "self" thought about as the present section of the stream is often confused with the principle of the unity of consciousness (what he later calls the "I"). In fact James himself does not seem to be consistently clear about the distinction. Perhaps this is due to the fact that he sees the spiritual "self" in the form of the present segment of the stream of consciousness as the foundation for the "I." We should keep this relationship and distinction in mind as we discuss the spiritual "self." It will be-

come clearer later when we discuss the "I."

The spiritual "self" is what everyone would agree in calling the "active element in all consciousness."⁵⁵ It goes to meet or receives one's feelings or the content of one's thoughts--it welcomes, rejects or assents to perceptions, and it is the seat of interest effort, and attention. It is different from the other elements of the "mental life" in being the most permanent element. All the other elements belong to it. Thus, the spiritual "self" in this abstract sense is the "self of selves."⁵⁶ James says that everyone singles out "some central principle" that has these general characteristics, as the spiritual "self." But when it comes to characterizing it further, there is much disagreement and little clarity.

It is at this point that it seems that people confuse the spiritual "self" with the "I," and as mentioned earlier, it is not always clear how James distinguished them either. Thus, he goes on to cite characteristics of the spiritual "self" that make it sound like the principle of unity of consciousness. He says that the general characteristics of the spiritual "self" are that it is the most enduring and intimate part of the self. The "I" is later described this way. In addition, James speaks of the spiritual "self" as being both the nucleus of the self and the central self. One might mistake this nucleus or central self for the "I," for

the "I" is later described as the owner of all the other "selves." Further, this spiritual "self" is never partially experienced; it is never split up. James says that when it is experienced in a moment of consciousness, it is fully present in that moment, no less than in a lifetime of such moments. So this spiritual "self" does not have a partial existence. When it is experienced it is experienced as a whole. But all the other "selves" can be split up and partially experienced. Experiencing one of our material "selves" does not include, necessarily, experiencing others. They also can fade in and out of existence or prominence. But it seems that anytime a spiritual "self" is experienced, it is experienced in all its forms. And its presence is more stable and permanent than the other empirical "selves." The "I" is also defined as a whole, simple entity; in this respect it is contrasted with the plurality of the "me."

Finally, in speaking of the intimacy of the spiritual "self," one of the things James means is that this "self" is most vividly experienced as bodily movements ("adjustments") in the head, along with a faint and obscure feeling of something more.⁵⁷ According to James, the principle of unity is also based upon these cephalic movements and other bodily experiences, and these form the true center of the self.

We have seen that it is difficult to distinguish what James means by the spiritual "self" from the "I."

They both are assigned the same characteristics in spite of the fact that they are supposed to be different in the following two crucial ways: (1) The "I" is supposed to be a subjective "self," while the spiritual "self" is part of the "me," the objective "self"; and (2) the "I" is a permanent principle of the self and the "me" is not a principle, but a "self." The upshot of both of these features is that the "I" is not a "self," while the "me's" are. Thus, the question is whether the spiritual "self" is a "self." James does mention, in explaining the spiritual "self," that everyone picks out a central principle in themselves, thus, suggesting that the spiritual "self" might be a principle, but he also explains that he does not mean by this the principle of the unity of the self. It seems that he does not mean the same thing by "principle" in each case either, for he insists that the spiritual "self" is a "self" or a kind of "self."

These two distinguishing marks are problematical. Since the spiritual "self" is a part of the "me," it must be an objective "self." Yet James calls it a subjective "self" and says that the spiritual "self" is the "self" as thinker. And although the spiritual "self" is said to be distinct from the unifying principle of consciousness, the "I," it is assigned the same characteristics of permanence, intimacy, centrality, and wholeness.

Perhaps the following interpretation makes sense

of these problems. When James says that the spiritual "self" is the "self" as thinker, perhaps he means that it is the "self" we experience or know when we know ourselves as a thinker. When we think about this "self," we think about its being a thinker. In this sense, it is "subjective"--it is thought about as being a subject. But since it is the object of our thought in this case, it remains a "phenomenal" objective "self." So, we might say that for James, there is a "subjective" "self" that is part of the objective "self," and a true or more basic subjective "self" that cannot be experienced, the "I." This will become clearer when we discuss the "I" and the "me" and how the spiritual "self" forms the foundation for the "I."

To give substance to the view of the spiritual "self" as one of the objective "selves," consider the possible things that James may mean by a spiritual "self" as thinker. We may think of ourselves as thinker in the sense that we reflect upon ourselves as thinking. I may now reflect upon the fact that I am thinking about James. I may further reflect upon the fact that I am thinking about my thinking. Thinking about my thinking or that I am thinking are different from thinking about the content of my thoughts--say, just thinking about James. It is as a spiritual "self" that I may think about myself as a thinker. I may also be thinking about things that are part of my spiritual "self," such as

my choices or sensibilities, my likes or dislikes. These parts of the spiritual "self" may be thought of as "selves" too. The spiritual "self" thinks about these "spiritual" things. In addition, I may think about my psychic abilities and functions--my ability to discriminate, will, conscience, etc. Thus, when we experience our spiritual "selves," what may be meant by spiritual "self" is (1) the "self" as it reflects upon the fact that it is engaging in certain psychic activities; or upon the fact that it is reflecting upon these engagements (as when one notices that one is noticing that one is analyzing something); or (2) the "self" that is constituted by or that reflects upon such thoughts as those that express choices, moral sensibilities, likes and dislikes, interests, etc.; (3) the "self" as it reflects upon its ability to think.

It is difficult to see what James may have meant by insisting that the spiritual "self," in all these senses, is characterized by permanence, intimacy, wholeness, and centrality. If it shares all these features with the "I," how is it different from it? Perhaps the answer that James intended was that since this spiritual "self" forms the basis for the "I," it comes closer to having these features than the other "selves." But it would seem that the way in which a part of the "me" can be most permanent, intimate, central and whole differs from the way these features are true of an "I," for the

"I" is a permanent principle of the self, and the spiritual "self" is not said to be a principle. If we make the assumption that all our "selves" have spiritual elements (reflective faculties, sensibilities, etc.), perhaps it is possible to see these features as permanent, intimate central and whole. Perhaps there is a spiritual "self" present in every other manifestation of a "self." This would make James's claim that the spiritual "self" is the most intimate, permanent, central and whole "self" more plausible. But it also makes the spiritual "self" sound like a principle. We cannot expect the objective "selves" to have these characteristics. They are in flux, "the same object being sometimes treated as a part of "me," at other times as simply mine, and then again as if I had nothing to do with it at all."⁵⁸ But it is possible that James means that the objective "selves" share certain kinds of elements that do have these characteristics, and that in this sense the spiritual "self" is like a principle. Perhaps such features are partly responsible for the unity of the self.

The view I am suggesting is that the spiritual "self" is a kind of experienced "self," like the material and social "selves." It is one of the kinds of "selves" that we experience, and we experience it as having the characteristics that the "I" has. This should not be surprising since, as will be explained in the next section, the "I" is based on the spiritual "self." These

comments still do not explain how the spiritual "self" can be a subject and can share certain characteristics with the "I." Nor does it adequately distinguish the spiritual "self" from the principle of unity of consciousness. Perhaps we have, however, shed some light upon what James meant by "spiritual self." This is the part of ourselves that we think of as having these characteristics and we can think of ourselves from these "spiritual" points of view. Perhaps this is the most that we can say about James's spiritual "self."

A final word must be said about the relationship between James's spiritual "self" and the integrated self. The fact that James talks of the spiritual "self" as the "self of selves" might raise the question of whether he is talking about a consistent system or a consistency impulse and of whether he thus has an integrated theory of the self. But James does not include a consistency impulse in the spiritual "self." The spiritual "self" is not trying to strive necessarily for consistency. It is our thinking part, or the part of us that thinks of ourselves as thinker. The spiritual "self" is central by its nature--as the thinking part of us it appears within or with the other "selves." As one of my social "selves" I may also be a thinking "self." But this does not rule out the existence of conflicting "selves" that are not under the aegis of a consistency impulse. Thinking is not connected, necessarily, with consistency

impulses, although it is a central feature of the self. And the spiritual "selves" may take different forms, just as there are many social and material "selves." The different spiritual "selves" may represent different principles or values, and not just one, like the consistency impulse. In this sense, spiritual "selves" may conflict too.

This is a point that requires clarification, since James has stated that the spiritual "self" is indivisible, like the "I." The way we can see how this can be is to sort out the different meanings James attributes to the spiritual "self" and the different phenomena and functions he was trying to explain. As an experienced "self," the spiritual "self" may manifest itself differently. The spiritual part of one's different social roles results in different impulses. In this sense, spiritual "selves" may conflict. But the spiritual "self" is a unity too and cannot be split up or experienced partially. I believe this is James's way of saying that we are all of our spiritual "selves" and that in being all of them there is a sense in which they are all the same thing. This kind of idea leads to James's "I." As his spiritual "self" acquires this sort of unifying function it has more and more in common with the "I." Thus, the more we want to focus on the spiritual "self" as the same subject in each "self," the more it becomes necessary to turn to an examination of James's notion of the "I."

James: The "I" and the "Me"Introductory Note

At this point it is appropriate to point out some differences between James's purposes, the things he emphasizes, and his views, and the interpretation proposed in the present work. For James did not use his ideas the way they are used here. For James, the different "selves" of persons are such in virtue of what he called "self-feeling." He held that persons must be aware. The question of the nature of the "I" for James was a question about the unity of consciousness. Persons, on his view, consist of streams of consciousness. James wanted the "I" to unite the different moments of the stream of consciousness, but he did not want it to be thought of as a substantial entity. Thus, as we shall see below, he posited that unity of the self comes from the fact that at every new moment, although there is a new "I" (or subject), each new "I" appropriates the contents of the stream from the moment before. The point of these remarks is the following: James did not think of the different "selves" as "selves" about which one may or may not be aware; and he did not posit the "I" to unite these different "selves." In fact, the only sense in which he thought there really were fragments in persons was in abnormal cases, illustrated by such things as hysteria and hypnosis. In these cases, separate "selves" exist-- they are separate consciousnesses. And

James analyzed any instances of the unconscious in the same way. These other consciousnesses were actually other selves for James--other subjects that did not appropriate the past moment of consciousness from the stream of the primary self. Although they belong to the person, they are not part of the self in the same way as the material, social and spiritual "selves."

These views differ considerably from the fragmentary view of the self, which allows that one need not be aware of one's attitudes. The fragmentary theory is not just a view of the abnormal self. But James's view is useful for the fragmentary view because his notion of the different "selves" of persons suggests the possibility of plurality in the self, or different attitudes and impulses. That there could be such plurality in persons is suggestive of the view that conflict is likely and that the fragments (James's "selves") may not always be experienced. James' notion of the "I" is also useful because such a mechanism helps us to envision a self that is both fragmentary and unitary. Keeping in mind, then, that James did not mean what is meant by the fragmentary theory, and that his views are being taken as suggestive of it, we turn now to a further look at what he did mean and how the fragmentary theory can make use of his ideas.

The "I" and the "Me"

James divides persons into "selves," the types of "selves" being material social and spiritual. There are numerous "selves" of each type that make up or belong to each person. We may now ask why such pluralities should be considered single persons. As an answer to a problem like this, James's distinction between the "I" and the "me" is suggestive. The "me" is the empirical "self" that we have discussed so far. It is made up of the different "selves," it is experienced, and it is objective. The most central part of the "me" is the part we experience most intimately, intensely, and directly--the bodily sensations associated with thinking. This makes one kind of spiritual "self" the core of the "me." The other "selves" are claimed by persons depending upon their "warmth," as either belonging to them or being part of them. In essence, then, the "me" consists of the different empirical "selves" that we apprehend to be ours or a part of us because of their warmth.

The "I" is the part of us that does the claiming, or "appropriating," as James says. It is the present thought of the present section of the stream of consciousness, so it was formerly a "me" and will end up as such. When it passes away the next thought will "inherit its title." What this means is that it will inherit the function of appropriation as well as all that the previous thought appropriated and inherited. The "I" is a subject, a thinker,

but it originates as a thought or objective "self" and ends up as one. Since there is continuity within what is appropriated and inherited, the "I" is the principle of unity of consciousness or the principle of personal identity. This continuity is the source of the "I's" permanence. Its unity is empirically discovered--we only experience ourselves as one thinker at a time, even though as a single subject we identify with (or "appropriate") many different "selves" both at different and at the same times. And as the "I" we feel a warmth in respect to former "I's" or past "selves" as subject.

In the beginning of his consideration of the principle of personal identity (the "I") James says that for this problem we must begin with "the sense of a sameness perceived by thought and predicated of things thought-about."⁵⁹ It is not simply the recognition that a subjective synthesis is necessary for thought that prompts us to think of personal identity,⁶⁰ but rather, this sense of sameness in the objects of our thought. The reason we think that the present and past "selves" are the same is not just that we are aware that our thoughts belong to "one thinker and not to another,"⁶¹ but also that we recognize the ones that belong to us as having a warmth and intimacy. All our thoughts come to us with warmth and intimacy. This is true of the bodily part of the self, as "we feel the whole cubic mass of our body," which "gives us an unceasing sense of

personal existence."⁶² In a similar manner, we feel the "nucleus of the spiritual self" as faint "physiological adjustments" or "the pure activity of our thought taking place as such."⁶³ Our more remote spiritual, material, and social "selves" are also felt with warmth. James explains that

the character of 'warmth', then, in the present self, reduces itself to either of two things,-- something in the feeling which we have of the thought itself, as thinking, or else the feeling of the body's actual existence at the moment,-- or finally to both.⁶⁴

We can separate "selves" that can be felt like this from thoughts that cannot. Our thoughts about someone else's thoughts are cold.⁶⁵ In the case of our own thoughts, it is as if we put our brand on them. This feeling of warmth is associated with all of our "selves" and "makes them into a whole, which we treat as a unit, no matter how much in other ways the parts may differ inter se."⁶⁶ In addition to warmth, continuity and similarity add to our sense of personal identity. Our "selves" are the same on the same grounds of continuity and similarity that are used in our judgments of sameness among other things: "it is a conclusion grounded either on the resemblance in a fundamental respect, or on the continuity before the mind, of the phenomena compared."⁶⁷

The sense of warmth, as well as continuity and similarity, make up our sense of personal identity. But it would be a mistake to see this identity "as a sort

of metaphysical or absolute Unity in which all differences are overwhelmed."⁶⁸ James warns against thinking that the different "selves" are alike in more respects than they are. When it is said that the feeling of warmth gives the "selves" unity, what is meant is that they are the "same in kind," but:

if from one point of view they are one self, from others they are as truly not one but many selves.⁷⁰

James later explains that as the "I" they are one self but as the "me" they are many. Thus, he is in agreement with those he calls the associationists insofar as he believes that "the self is an aggregate of which each part, as to its being, is a separate fact,"⁷¹ and that it is empirical. But he believes that the associationists have ignored some subtle facts about consciousness. Common sense tells us, James claims, that the unity of the "selves" is not an appearance, but

that it involves a real belonging to a real Owner, to a pure spiritual entity of some kind. . . . There must be a real proprietor in the case of selves, or else their actual accretion into a 'personal consciousness' would never have taken place.⁷²

This is not the usual empiricist or associationist view, James says. Their view has the parts of the self somehow "integrating" together or "fusing into a stream."⁷³ James thinks such fusing is incomprehensible without a "medium," and that the medium that is required is "the real, present onlooking, remembering, 'judging thought' or identifying 'section' of the stream."⁷⁹

As soon as the Thought comes into existence it finds past facts its own. This indicates that Thought does not create the unity of the self, but that the unity cannot be without thought. This fact seems to require that "the Thought have a substantial identity with a former owner," and "not a mere continuity or resemblance" as in the previous account; "a real unity" seems to be required.⁷⁵

But James claims that the unity we are looking for is nothing more than the one we experience, and what we experience is a present segment of consciousness that finds other "selves" warm. Those who disagree might be tempted to postulate what James calls a dominant Thought, an Arch-Ego, like the 'soul' of Metaphysics, or the "Transcendental Ego" of Kantian philosophy. But James's account is that each Thought inherits the title of the former one. Each Thought "dies and is replaced by another;" the other finds its predecessor warm:

Each thought is thus born an owner, and dies owned, transmitting whatever it realized as its Self to its later proprietor.

It is this trick which the nascent thought has of immediately taking up the expiring thought and 'adopting' it, which is the foundation of the appropriation of most of the remoter constituents of the self.⁷⁶

Thus, each section of the stream of consciousness stands as the "representative of the entire past stream."⁷⁷

This is the only possible verifiable account of personal identity: one thought appropriates another. It may be

pointed out that the act of appropriation is obscure. How can Thought appropriate itself, it might be asked-- appropriation implies an agent. The answer that James provides is that the Thought "is a vehicle of choice as well as cognition."⁷⁸

Though we can feel the present moment of consciousness, we do so with difficulty and we can know about it more easily when it passes away. The feelings we have of our body are more vivid than those faint feelings we may have of the "I" through introspection. These bodily feelings are, therefore, "our personal identity."⁷⁹ On the basis of this more vivid bodily feeling, James concludes that the appropriations of the "I" are not so much "to itself than to the most intimately felt part of its present Object, the body, and the central adjustments, which accompany the act of thinking, in the head."⁸⁰ These bodily feelings are the "real nucleus of our personal identity" and even if the present "Thought were entirely unconscious of itself in the act of thinking," these bodily facts "would be a firm basis on which the consciousness of personal identity would rest."⁸¹ Thus, James believes that there is some sort of feeling that gives us our idea of personal identity or unity. The bodily feeling is more vivid and complete. The "I" is experienced through these bodily feelings, and it is nothing over and above such feelings. (It should be noted that when James says the "I" is not experienced

he means that the "I" is not something over and above what is experienced; it is not a substance.)

James concludes that the "passing Thought" is "the Thinker."⁸² The "I" has thus been described in purely empirical terms. The facts that make it up do not require a non-phenomenal Thinker. The "I," in fact, is based upon one of the entities James describes as the spiritual "self"--the one that is the "self of selves." The "self of selves" has the following characteristics: it was said to be permanent and our most intimate "self." It was said that the spiritual "self" in this sense is never partially experienced--it cannot be split; when this "self" is experienced in a moment of consciousness it is fully present in that moment, no less than in a lifetime of such moments.⁸³ Finally, the central and most vivid feelings of this spiritual "self" were said to be bodily. James was talking about this spiritual "self" as a "me," but now we see the full meaning of these characteristics. For this "self" can become an "I," the subject, the thinker. And this "I" is a necessary condition for one's being one and the same and fully a person. Hence, the characteristics of permanence, intimacy and wholeness.

While the "self" that James calls the "me" is the empirical or objective self, the "I" is the subjective self. The "I" is the element of the person that allows judgments to be made (it is the vehicle of jud-

gments) and that continues through time.⁸⁴ The "I" as the subject might be interpreted mistakenly as the real or essential self. This would be an attempt to fit James's position on the "I" into the integrated view, where the "I" would be represented by a consistency impulse. However, the "I" is not such a self; it is not the vehicle for a constant impulse for consistency. The "I" is not that kind of principle; it represents the unity of the self, but not consistency. The "I" is what makes the self one self, but it does not make the self an integrated, non-conflicted, self. The question of whether there is a unifying subject in the self is not the question of whether there is a constant consistency impulse. Just as it was wrong to see the spiritual self as a consistency impulse, it is also wrong to see the "I" as a consistency impulse.

James's "I" avoids the complications of the integrated theory's consistency impulse. One problem that the integrated theory had was that it was unable to give a nonarbitrary account of which beliefs or feelings were consistent with the person. Beliefs or feelings that were not cited seemed to have just as strong a claim to representing the self. With James's notion of the "I" however, this problem does not arise. Since the "I" is not a system, and in that sense it is not a self, its centrality does not raise this problem. The "I" simply allows the various beliefs or feelings to gain

expression. There still is a question concerning how attitudes get expressed for James, but since it is not the question of how one and only one set of beliefs or feelings gets expressed, the problem of arbitrariness does not arise.

The problem of expression that does arise for James is the question of how each "self" gets expressed. If this happens by means of a choice made by the "I," James's "I" might involve the same kinds of problems as the integrated theory's dominating consistency impulse, because we might think that the "I" is guided in its choices by a single consistent system of values. The implication of this would be that the "I" is the "true self," it represents consistency, and that James's view of the self can be interpreted as an integrated theory. But on the following interpretation, James avoids this problem. Since the mechanisms of motivation and choice are inherent in the empirical "selves" not in the "I," the determination of which "self" gets expression (or gains temporary dominance) occurs on the level of the empirical "selves." The "I" is just the vehicle of expression for a decision that has been made on the level of the "me's". One of my material "selves" may be dominant at one moment, ruling out expressions of my other empirical "selves." Or several empirical "selves" may gain expression through the "I," whether they are consistent or not. Somehow a decision is reached among

them, as to which gets expressed. The "I" is just the vehicle of expression and the mechanism that preserves unity (not consistency) of the person. The "I" can be seen as the locus and vehicle of policy-making. But it is not a policy-maker. Policies are based upon input from the "selves." The values that guide the policies it expresses can vary. They depend upon the "selves" that are dominant in the situation. Although the "I" is neutral in its role, it is possible that the "I" is a policy-maker in the sense that it surveys and ranks the different attitudes and comes up with a balance. (This notion of a surveying and ranking kind of policy-making will be discussed further, later in this chapter). The result could be expression of some of them while others are recognized but not expressed, or are only expressed minimally; it could be to ignore some attitudes and to express others; it could be to combine some attitudes or to eliminate some. But the policy it sets will depend upon the particular configuration of attitudes that are active in the situation. Some kind of policy is evoked from this configuration. The "I" simply carries out the policy therein evoked. This would accord with James's view that the "I" is active, not passive; it does have to rank attitudes and balance them.

The spiritual self is the "I" experienced as a "me." James talks about the spiritual self and the "I"

in the same way sometimes, and my interpretation of this is that at times he means the same entity. Insofar as it is experienced as a "self," the "I" is a spiritual self. But it is different from other "selves" in its link to, and in the characteristic it shares with, the "I." This makes it possible for us to also see James's spiritual "self" as the policy-maker. James says that the spiritual "self" is responsible for attention and choice. Thus, James's spiritual "self" is like his "I." He says that the spiritual "self" is responsible for choice and attention,⁸⁵ it has the ability to survey the other "selves," to cultivate some, to eliminate or ignore others.

A problem for the integrated theory closely related to the problem that defining persons in terms of consistency impulses is arbitrary, is the problem of reconciling the view that persons are characterized by this consistency impulse with the inconsistent activities of the self and with the lack of evidence for a constant consistency impulse. James's "I," not being a consistency impulse, does not require that what gets expressed through it be consistent. It is simply the vehicle through which whatever is there--constant or not--gets expressed.

The main problem of the integrated theory for which the Jamesian "I" suggests a solution, is that the integrated theory is unable to show why the different and inconsistent elements of the self do not split up

into multiple selves. The notion of a consistency impulse does not eliminate the multiple self problem, as we have seen, for it is arbitrarily chosen. It seems to exclude other impulses of the self and is forced to say that they are external, or that the inconsistency is an illusion. To rule that certain of these elements are external is not possible, given the difficulty of defining persons without either involving operations that might yield conflict (such as valuing, judging, identification, avowal) or giving an incomplete picture. To say that the different elements in persons really result in consistency is empirically false. The result is that on the integrated view we are unable to see how parts of the self that we think ought not to be excluded from the self can all be part of one whole. On the present interpretation, James suggests a solution to these problems with the "I." This is possible because his "I" is stripped of the characteristics of any system. Since the "I" is only a principle or mechanism of unity, the question of why it is chosen to "dominate," rather than other elements in the self, does not arise as it does in the case of a dominating consistency impulse. If the "I" were a full-blown self-like entity, with motives, beliefs and feelings, we would have the problem of trying to show how it relates to the other elements of the self that also seem to be "selves," and why it is the dominant one.

If the "I" were like a "self," with beliefs, feelings, values, motives, etc., we would have to ask why this particular "self" is singled out as the "I." If the "I" is a "self" there seems to be no difference between the "selves" that James calls the "me's" and the "I." Then there would seem to be two alternatives. Either the self that is said to be the "I" is arbitrarily chosen and does not necessarily deserve the title; nor does any other self, on the same grounds. The result would be a collection of selves with nothing to unify them, and therefore no reason to ascribe them to one person. The other alternative is to try to show that the system chosen as the "I" is a consistency impulse. But this will not do either. We have no evidence for such a constant impulse and a lot of evidence for diversity and inconsistency. That there is so much inconsistency in persons suggests that they cultivate or tolerate it and that they do not have an overriding drive to eliminate it. James's "I" cannot be a consistency impulse, because it is neutral. It does not bring a point of view to the "me's." Rather, it gets its views from the "me's." It is a principle and a vehicle, that in virtue of which the many "selves" (or attitudes) belong to one person, and through which policies are made and expressed.

James's view of the self suggests the major features of the fragmentary theory. In the fragmentary

theory the person is made up of a variety of attitudes. Different configurations and orderings of these attitudes can occur in one and the same self. The attitudes may be inconsistent. Thus, a single person may be many different "selves," and some of them may even be inconsistent.

James's most important contribution to an understanding of the fragmentary theory, for our purposes, is his notion of the "I." It suggests a way that we might construct a theory of how policy is made in a fragmentary self. This is important because a major claim of the integrated theory is that policy-making in a person must be guided by a consistency impulse. According to this view, policies not covered by such an impulse are external to the person. They are nearly as remote from the person as those that belong to other persons. Thus, the integrated theory suggests that the only alternative to its position is a multiple self view. James's "I" might be useful in showing that a fragmentary self can make policies and not be a multiple self. Thus, James's "I" might be useful in our attempt to show that self-deception requires a fragmentary self. As shown in Chapter II, self-deception requires a single self that can be in conflict, that can maintain conflict, and that can create conflict. Self-deception could not exist in a multiple self. It requires a self that can make policies and handle conflict, but possibly not via a consistency impulse.

James and Mead

James and Mead are alike in their emphasis upon the claim that persons are made up of different "selves." Mead's emphasis upon social roles is like James's views about the social self. Both philosophers make a distinction between the influence particular individuals, roles or attitudes may have over us, and that of groups, the community, or society. In this they show their appreciation of the complexity of the self. Mead takes this point a step farther, when he asserts that the incorporation of groups, the community, and society involves bringing systems into the self. In this sense he adds to our understanding of how "selves" may be seen as systems, and how the self as a whole is organized into a single self. Such systematization means that the self can survey all of its "systems" and that it can have a concept of how they affect one another. James adds to this vision of how the self sees itself as a single self with his concept of warmth. James focusses more on how we experience ourselves--on "self-feeling"--, whereas Mead tells how the self develops and shows its dependence upon society. Both perspectives add to our vision of a fragmentary self.

Although Mead claimsthat the "selves" that make up the "me" are internalized social roles, he allows for a "self" that is like James's spiritual "self" when he says that "there are parts of the self which exist

for the self in relationship to itself."⁸⁶ This is the "I." It is the self-reflecting "self." Like James's spiritual "self," it is also the thinking or judging "self," and it is the locus of our values and ideals. Finally, it is, for Mead, the "self" of the origin of action. Thus, Mead's notion of the "I" seems to include thought, ideals, and agency. It is unlike James's "I," which is a mechanism for unifying the self. Mead's "I" is a structural part of every "self."

But, he says, it is unlike the "me" in that we cannot experience it directly, but only in retrospect. In this it is different from James's "I." James thinks that the evidence is never conclusive about whether there is something purely spiritual and unexperienced about the self, but he emphasizes that as far as we know the "I" has no content to be experienced beyond those physical processes we experience when we experience the "self of selves";⁸⁷ "The words I and me signify nothing mysterious and unexampled-- they are at bottom only names of emphasis. . . The 'I' meaning for the Thought nothing but the bodily life which it momentarily feels."⁸⁸

Mead's "I" also differs from James's "I" in that it has functions that make it more than just a mechanism. Mead's "I" contributes the element of uncertainty to the conduct of the person. The "I" reacts to the "me" and the consequences are unpredictable.⁸⁹ The spontaneity, freedom and initiative of the "I" are the factors that

enable each individual to be a unique personality.⁹⁰

The important points to emphasize are that Mead sees the "I" and the "me" as contributing to behavior in different ways, and that he sees them as being two different phases of the self:

Taken together they constitute a personality as it appears in social experience. The self is essentially a social process going on with these two distinguishable phases. If it did not have these two phases there could not be conscious responsibility, and there would be nothing novel in experience.⁹¹

The fact that Mead sees the "I" as a phase of the self that manifests itself in behavior and that accompanies the "me" is reason to think that his notion of the "I" is like James's notion of the spiritual "self." Mead also says that the "I," the spontaneous part of the self, is manifested by "my response to the attitudes of others,"⁹²

This mention of spontaneity again brings to mind James's description of what the spiritual "self," upon which his "I" is based, feels like: he says he is aware of "furtherances and hindrances" in his thinking. Thoughts arrange themselves on the side of the original thought or against it:

The mutual inconsistencies and agreements. . . produce what seem to be incessant reactions of my spontaneity upon them.⁹³

Such spontaneous actions take the form of welcoming, opposing, etc. James refers to this experience as being of his "palpitating inward life."⁹⁴ Both James and Mead, in describing the spiritual "self" and the "I" respectively, are referring to the part of the self that res-

ponds to itself and that contributes spontaneity and novelty.

There is then a major difference between James's and Mead's "I," and an affinity between James's spiritual "self" and Mead's "I." One difference between their notions of the "I" is that Mead seems to pit the "I" against the "me" along the lines of inner vs. outer, self contrasted with society, and perhaps even true self contrasted with roles. This is not what James means by the "I," although it may be what he means by the spiritual "self." James's "I" is a principle or mechanism that is present in the self, whereas Mead treats it as another "self." For Mead, the "I" represents the self's response to the part of it that is conventional. We have seen that the "me" for James includes many different kinds of "selves," with different sources; they are not necessarily conventional. Thus, James's spiritual "self," like Mead's "I," may represent the self's response to the "me." What this means for each of them may differ, however, since James does not think that every "me" is a role internalized from society.

Further differences between James's and Mead's "I" are that Mead's "I" initiates action and is the source of the unique quality of each self,"⁹⁵ while James's "I" is the vehicle through which action gets initiated and it is his "me," not his "I," which is the source of both action and of what is unique. James's

"I" is the vehicle for the expression of choice and uniqueness, but the many "selves" of the "me" provide the content of the self. For James the uniqueness of the person is due to the different combinations of empirical "selves" it has. The Jamesian "I" is responsible for uniqueness only in the sense that by functioning as a unifier it allows the distinction of one individual from another, but not, as Mead seems to mean, as it distinguishes one aspect of the self (the "I") from the other (the "me"). Mead's "I" can also consist of

our most important values. . . . It is the realization in some sense of the self that we are continually seeking.⁹⁶

this resembles James's discussion of possible "selves" and the ideal social "self,"⁹⁷ and in fact Mead makes note of this affinity with James.⁹⁸ But what Mead means is that we are the values of society--the "me." But as an "I" we can contemplate and try to achieve other values.⁹⁹ James, on the other hand, does not locate our values and ideals in the "I," although they may only gain expression through the "I," due to the structure of the self and how the subject is related to the rest of the "selves." It is not due to the "I's" being a "self" or the locus of our values. All of the functions Mead assigns to the "I"--agent, source of the unique in the self, source of values--are attributed by James to the spiritual "self." For this reason, it seems

that they have the same idea about the structure of the self.

Mead does not mention anything about the "I" being a unifying principle. Unity in the self is provided for him by the systems we incorporate, although it is also due to the "organizing impulses" of the "I." But as we have seen, these are not the "I's" only impulses. It is capable of unifying the self and of disrupting it, via antisocial responses. Mead's "I" seems to be meant as a kind of structure present in or accompanying each "me," not as a unifier. This raises questions that are not raised by James's approach. Mead's view is unable to account for the notion of shifting among conflicting attitudes, whereas James's "I" as a principle of unity, does. Mead's view easily might be misinterpreted as the view that the "I" is a consistency impulse which responds to the various "me's". This sounds like an integrated theory, with the "I" functioning to integrate or systematize the "me's" according to the "I's" single set of ideals. But, as mentioned above, Mead is not committed to a view of the "I" as an integrated consistent system, so we do not have to take his view to be an integrated theory. What Mead does not explain is why such plurality should be considered one self, beyond the fact that it belongs to a single organism.

James's view, however, is more adaptable to an account of the unsystematic, diverse, and inconsistent

nature of the self, because it provides an explanation for shifting attitudes. The Jamesian "I" is the same over time, in a sense, but only in a sense that preserves diversity. It is each time a different thought, made up of different attitudes or different configurations of them. It can shift between attitudes and still represent the same person. Mead does not show how the "I's" functioning preserves the diversity of the self. Nor does he show how any other mechanism does this. He says that the individual organizes the parts of the "me" into a system that reflects the system of society, but as we pointed out earlier, this simply allows us to survey our attitudes. Mead does not show how such an overview can be taken by an "I" that is part of each "self." We want to know in virtue of what these different fragments (or "selvss") are one self. Mead does not give us a mechanism like James's "I." The most he provides is a single organism.

We saw that Mead's notion of the "I" can be confusing. It sometimes seems to be the subject of each "self," but he also sometimes gives it the function of systematizing the self. The question arises as to how the "I" in all of these different manifestations can be the same "I." How can we see Mead's self as a single person and not as multiple selves? The Jamesian notions of the spiritual "self" and the "I" help clarify this problem, although, as we saw, there is considerable

difficulty involved in sorting out what James means by these two concepts. Nevertheless, it appears that we can get a model that solves the problem of unity in the fragmentary self if we interpret Mead's "I" as being like James's spiritual "self." Then we have James's "I" to unify the self. There is reason to think that what Mead had in mind by the "I" is basically what James had in mind by the spiritual "self." Mead's "I" and James's spiritual "self" are not empirical "selves" and they are not principles or mechanisms of unity in the self. They are part of the structure of each "self"; they are the subjects of each "self" and they contribute reflection and spontaneity. We can experience ourselves from the perspectives of our spiritual "selves"-- thinking of ourselves as thinkers. But this is not a self, it is a way of focussing on an aspect of ourselves, of our "selves," of our faculties. The spiritual "self" (or Mead's "I") is not a principle of unity either, as explained in the previous section, it is the foundation for the unity of the self. It gets translated through the Jamesian "I" into expression and when it does it inherits all the "selves" of the past stream of thought. The difficulty we had in distinguishing James's spiritual "self" from his "I" is the same difficulty we had in dealing with the ambiguity of Mead's "I." James goes a step further in distinguishing two different concepts. Mead tried to make one entity do

the work of two. This is understandable, since the spiritual "self" and James's "I" are closely connected, the one giving rise to and forming the foundation of the other, and also because the "I" in James's view, is not really an entity.

"Selves" or Attitudes

descriptions of the self concerns their references to the different "selves" that constitute selves. This kind of terminology can be confusing. In previous discussions we have tried to distinguish the fragmentary theory from a multiple self view. But talk about our different "selves" suggests a multiple self view. Talk about roles and conflicting roles may be less confusing, but the least confusing tack to take would be to talk about attitudes, by which we mean to include things such as beliefs, feelings, impulses, values. The notions "self" and "role" suggest reference to motivational systems. It is difficult to distinguish the concept of motivational systems from that of persons.

We saw that Fingarette had a notion of what the fragments are that constitute the self that is somewhat fuller than reference to attitudes. His notion of "engagements" is of "rudimentary but unified complexes of reasons-motive-feeling-aim-means-and moralistic reaction." He talks about the "community of the self." He also refers to engagements as "projects." But his notion of engagements

comes close to suggesting the notion of "selves," especially when he talks about engagements as if they are motivational systems that can act independently of persons. He does this when he talks about engagements of the individual as opposed to those of the person.¹⁰⁰

Watson avoids the implication that projects are somehow independent of persons by stating that the fragments of persons are all motivational systems and that they do not act independently of persons although they might operate independently of or in contrast to persons' better judgment.¹⁰¹ But Watson's view leaves unanswered the question of how motivational systems are different from selves.

To avoid such problems, it is better to describe the plurality and inconsistency in the self in terms of attitudes, such as beliefs or feelings or impulses. In a similar spirit, talk about what conflicts and is denied or claimed in self-deception (in Chapter II) was confined to talk about beliefs and feelings.

The view that persons are made up of self-like entities has been criticized by de Sousa on the grounds that it implies a multiple self view.¹⁰² de Sousa is sympathetic to a fragmentary type of theory but he sees the importance of restricting the notion of the elements that conflict to attitudes.

de Sousa, in criticizing the view that certain kinds of conflict, like akrasia, imply that there is a

"'committee' of quasi-persons" within the single agent that is the person, calls the members of such a committee "homunculi" and defines "homunculus" as ". . . a simplified rational agent. . . which has at least minimal structure. . . ." ¹⁰³ Such a structure is a homunculus by virtue of being thought of by the person as internal. The person identifies with the homunculus or endorses it, and de Sousa adds, "one can endorse as a part of oneself only what can be viewed as an intentional system." de Sousa explains that

something is an intentional system, and has what I call "minimal rationality," if its behavior may be thought of as explainable in terms of wants and beliefs. . . . ¹⁰⁴

The thesis that de Sousa considers is that inconsistent action of the person implies homunculi. Thus, he says that:

putative homunculi, with their wants and beliefs, tend to constitute themselves into a person as a whole: in the most likely cases where we might want to talk about a person divided, it is the conduct of the person as a whole that will show erratic or inconsistent signs. ¹⁰⁵

de Sousa criticizes the homunculi view, however, by claiming that to make sense of homunculi as quasi-persons they would have to act independently of the person. Thus, the problem he points out is that in order to explain conflict in the person we cannot think of homunculi as agents independent of the person, but that in order to make sense of homunculi we must see them as independent agents. This paradox is what makes de Sousa reject the

homunculi view. He sees it as a view that falls apart into a multiple self view. Since, he thinks, what homunculi are posited to explain are conflicting projects of the person, he suggests that we should think of the self as being complex in that it has projects that conflict and that are on different levels, instead of thinking of it as multiple (i.e., made up of homunculi). This approach, de Sousa thinks, preserves the idea that it is the person who is inconsistent in akrasia, rather than the inconsistency being between several quasi-persons. de Sousa concludes that the self is neither multiple (what he sees as the homunculi view) nor integrated. Rather, he says, it is complex.

de Sousa thinks that it is better to think of the self as a complex policy-maker constituted by different levels of conflict among desires, beliefs, or feelings, than as a complex of different and conflicting integrated systems. One reason for his view, as already cited, is that he defines homunculus in such a way as to make it impossible that the person could be identified with the homunculus. They are like different selves. Another reason he gives is that the hierarchical-policy view places less stringent requirements on what can conflict (conflict may be simply between desires and not whole systems).

de Sousa thinks that his hierarchical-policy view better captures the kind of conflict that occurs in persons because it allows more for the

fluidity and interpenetration of projects. He sees the homunculi view as inadequate to account for this because he sees homunculi as stable intentional systems. de Sousa is sensitive to the fact that what is "central" in a person can change, it "is not measured by a single standard":

Complexity of involvement, long-term importance, and high-priority in the immediate future might all be cited, and are not equivalent.¹⁰⁶

de Sousa thinks that what is peripheral at one time may move to a more central position at another time.

I believe that in talking about different "selves" or roles Mead and James did not mean to imply a multiple self view. Rather, they meant to emphasize that single persons can have many attitudes. Calling the attitudes "selves" is meant to call attention to the fact that persons are connected to each of these attitudes, and insofar as there is this connection, the union of persons and their attitudes forms a system. From the point of view of each of these roles or attitudes persons can experience and act, and in that sense we can talk about them as being motivational systems. Mead and James did not work out the metaphysical implications and problems of calling the fragments of the self "selves." The philosophical problems and confusions that are generated by calling them "selves" outweigh the benefit of doing so. But in the present interpretation, the views of Mead and James are not changed by specifying that the

fragments are attitudes, or, more specifically, beliefs or feelings. We can still preserve the point, that is necessary to the fragmentary theory, that these beliefs or feelings all belong to and represent persons, by showing that the fragmentary theory is not a multiple self view. James's notion of the "I," as previously discussed, is helpful on this score.

How Policy Gets Made in the Fragmentary Theory

We have seen that a major criticism of the fragmentary theory is that it reduces to a multiple self view. We have seen that there is another way to view the fragmentary self, such that its parts are not full-fledged selves. In this view the parts of a fragmentary self are unified into a single whole, which is the self. This type of fragmentary view claims that the self is complex and pluralistic in that it contains conflicting elements that are not under the influence of a dominating-consistency impulse. In this view, a fragmentary self does not have to be seen as a multiple self. Since it is a single self, it can do things and it can make policies. Action and decision as a single self would not be possible if each of the fragments were a full-fledged self.

Thus, the questions of how things get done and how policy gets made in a fragmentary self are important for establishing that a fragmentary self need not be a multiple self, as well as for establishing its

compatibility with self-deception. Policies can be the result of input from all the different and possibly conflicting beliefs or feelings. There are at least three ways that this can occur. One possibility is that beliefs or feelings are evoked by circumstances. This type of activity in a fragmentary self is not adequate for explaining self-deception because one of the conditions of self-deception is that it is purposeful. Self-deception is not just a response that is evoked in persons by circumstances or that occurs within us independent of our will. Rather, self-deception is something persons do on the basis of interests and motivations upon which they choose to act.

Another possibility for how beliefs or feelings gain dominance in the fragmentary self is strength. This view is not acceptable because it does not differ from a multiple self view. Different selves can conflict and survive on the basis of strength. Strength as a criterion is also not an example of purpose.

But a third way that action might originate in a fragmentary self is through a certain kind of policy-making. In this case, persons survey their beliefs or feelings, they rank them and they come up with policies that represent a balance. The results of each of the three methods of self-expression might be the same, but the latter procedure is both purposeful and presupposes unity. Policy-making might involve having

certain beliefs or feelings that initially dominate partly because the situation evokes them, but persons can choose which ones get expressed and can go against factors such as mere strength and circumstances. Persons can decide not to act on attitudes that press for expression more strongly than others--i.e., they can decide not to act on impulses they find most compelling or identify with most strongly--or not to let roles dominate that would be appropriate in a situation that would normally evoke them.

In this view, every attitude gets a "vote," or gets considered, and in that sense every attitude is represented in the policy. But not every attitude need be expressed in the policy. Persons may survey their attitudes and may decide a variety of things--from not expressing some of them, to expressing them all. Thus, if persons have a conflict between studying and getting exercise, and they come up with a policy of studying, this means that they could have consulted their different attitudes and have decided, in spite of one, to let the other one be operant. The exercise attitude does not have to be eliminated, and so, the conflict need not be eliminated. Persons may still want to exercise, but they can adopt a policy of studying. In adopting such a policy, persons might still express their ambivalence about it, even if only in the mild form of avowing that they also value exercise but are not letting that attitude

get expressed. It is also possible that the policy will be a compromise between attitudes and that they will get expressed but not fully, or that some attitudes are expressed not as fully as others.

At this point the question might arise (as it did when we discussed James's "I") as to whom the person really is and whether there is a policy-making entity that also has attitudes. Are persons nothing over and above some of their attitudes, or all of them; are they just the part that makes policies, or are they both? On the versions of the fragmentary theory that claim that attitudes get expressed as the result of strength or of being evoked by circumstances, the self is all of the attitudes. In the version of the fragmentary theory that we are considering for self-deception, the self is all of the attitudes and this includes policy-making or the policy-making entity if there were such a thing. Both views require a unifying principle in the self--like James's "I." So the entities that make up the self on each version of the fragmentary theory are not different. But their processes do differ. When policy-making occurs, the attitudes not only gain expression as impulses that are responding to external stimuli, or that are strong, they gain expression due to a policy that is the result of persons' ranking their attitudes. This means that persons take into account the claims of the different attitudes (one might ask oneself: how strongly do I identify

with these different attitudes?) and try to realize their aims. Maybe certain attitudes will appear to be more compelling in a certain situation and persons will give them a heavier vote. The process of policy-making involves surveying all the attitudes and balancing them. How they get balanced depends upon how they interact. This is different from their being merely evoked, in which case every attitude is not represented in the outcome--for the outcome is not a policy but a simple response.

But it is not different from the situation in which impulses dominate due to strength. The contention we are considering here is that if there could be a unified fragmentary self there could also be policy-making in such a self, and that the policy-making could be a kind of surveying and ranking as described above. But this contention is not easy to describe or defend. It is difficult to decide just how the policy gets made. In keeping with the solution borrowed from James's notion of the "I," the motivation of the policy comes from the attitudes, not from any super-imposed or higher-level policy-maker, not from the "I." This view differs from the view that has attitudes dominating merely due to strength, although strength is a factor in policy-making. Strength determines how thoroughly persons will consider attitudes in ranking them, but we want to say that it does not determine how big a vote they will get.

We want to say that strength of persons' identifications with certain attitudes can be an indication of how much they want to express those attitudes. But they can also choose conflicting attitudes to express that seem less compelling to them--i.e., they can choose to act against the path of least resistance. The point is that either attitude belongs to such persons, and that ranking makes it possible for strength to be a factor in choice and for it also to be overruled.

We saw earlier that the question of a policy-maker in the self touches upon the question of unity in the self. The fragmentary theory responds to the problem of unity by positing a single subject in the self that can perceive itself from the perspectives of different and conflicting attitudes. Like Mead's "I" and James's spiritual "self," this subject never appears unless it is connected to a "me." So it would be wrong to see the subject as the real self (everything else being merely things the self claims or rejects) and as being apart from the attitudes. There is a sense in which there is a subject connected to each attitude (because the person can identify with or express each attitude). And as a policy-maker (the surveyor of all the attitudes) one is also a subject. These subjects--or what seems like different subjects--are all the same, but they get manifested differently depending upon the content provided by the different attitudes. We want to be able to defend

the view that the policy making "entity," if seen as simply a subject, is not an attitude, but rather, it is the unifying mechanism in the self (like James's "I") through which attitudes (via policies) get expressed.

The point of this discussion is to clarify several things. First, the fragmentary theory claims that the policy-making "entity" is not another self, the true self, or a self over and above the person's attitudes. Nor is it an attitude. It would be more accurate to talk of policy-making in the self, rather than a policy-maker. This is because the policy-maker is the person, but the person is represented by all the attitudes, any attitude with which it identifies, and the person is also capable of surveying and ranking all of its attitudes. Thus, the person is not a different entity from the attitudes it ranks. But the view I am distinguishing this view from is one that says that the policy-maker is a self which imposes its values and either eliminates or transforms the other attitudes until there is no conflict with the policy-making self. This would be a different kind of policy-making in the self. The balancing and ranking kind of policy-making allows conflict and is done from different points of view, depending on the attitudes. The view the fragmentary theory wants to defend is that the policy-maker has no views or values of its own. It gets these from the attitudes which it ranks. The ranking is done according to the input from the

attitudes. Policy emerges from the balancing of these attitudes. This allows the "policies" of the "policy-making self" to shift. If the "policy-making self" were something over and above the attitudes it ranks, with its own system of values (or attitudes), it would represent a constant system of values, it would impose them when it balanced and ranked attitudes, and it would be trying to make the attitudes consistent with itself, and thus it would be a consistency impulse.

We have described what it would be for a fragmentary self to make policies. Making policies in a fragmentary self often involves handling conflict, because the attitudes of a fragmentary self are likely to be in conflict. And a fragmentary self is capable of having different policies of handling conflict, since there is no constant impulse to be consistent. Since self-deception involves creating or maintaining conflict, it should be something a fragmentary self can do. We saw that self-deception might not be something that an integrated self could do because an integrated self is limited to a certain kind of policy in respect to conflict. That is, it can only make policies that strive to eliminate conflict. But self-deception does not eliminate conflict and the fragmentary self is a kind of self that can make policies that do not eliminate conflict.

The Weakness of the Fragmentary Theory

Even though we have tried to sketch what policy-making is according to the fragmentary theory, there is a weakness here. It is difficult to get a clear view of how the surveying and ranking is done. More specifically, it is difficult to pinpoint the source of the criteria that are used to rank conflicting impulses. The claim of the fragmentary theory, based on the suggestion taken from James, is that the ranking comes from the impulses themselves. But no adequate picture has been given of how these impulses interact when purpose is involved. If persons can rank their impulses the ranking must be based on criteria other than the mere strength of the impulses. Ranking would not be necessary if strength were the only factor. But if there is an "I" over and above impulses that ranks them, it seems likely that it makes choices on the basis of its own criteria and this seems to imply that the "I" is a consistency mechanism, a "true self," a dominant and defining system of values. The plurality of the impulses and the neutrality of the "I," as suggested by James, were seen in our earlier discussions as being supportive of the view that policies can shift or conflict or exist in spite of persistent conflict. It may be that further work could unearth a way to strengthen this argument. But the present view is that this point of view, although a possibility, has not been adequately established.

Thus, one problem that the fragmentary theory has is in showing how purposeful policy can be made. If the fragmentary theory invokes some kind of consistency impulse to explain this, it then has the problem of distinguishing itself from the integrated theory.

The problem mentioned in this section will be taken up again in Chapter V.

Concluding Summary

In this chapter we have described the fragmentary self. We have explored the possibilities that such a self need not be multiple and it is not a special kind of integrated self (i.e., that the integrated and fragmentary selves are mutually exclusive). What is meant by distinguishing the fragmentary from the integrated self is that the fragmentary self is made up of attitudes that may conflict and that there is no constant consistency impulse to integrate them. The fragmentary self is not a multiple self. The many and conflicting attitudes can all belong to the person. Each attitude represents the same person, as do all of them together. Persons are capable of acting even though they are fragmented in this way. Policy-making can be done because persons can survey all of their attitudes, balance and rank them, and come up with decisions. The values that guide such policy-making may vary, so that what is policy at one time, may not be such at another time.

The policy-making mechanism is not a self, but a structural part of the self. It transmits persons' different attitudes, but it does not transform them, necessarily, as is required by the integrated theory, into a consistent unity. According to the fragmentary theory, unity is not an achievement of a self, whereas consistent integration is; while, in the integrated theory, to achieve unity is to achieve consistency. Thus, the fragmentary self is complex, in that it contains plurality and inconsistency, but it is also unified. This is the picture the fragmentary theory presents of the self.

In conclusion, however, after describing what is meant by a fragmentary self and exploring accounts the fragmentary theory might give for policy-making and handling of conflict, we find that the fragmentary theory seems to require a consistency impulse to complete its explanations. Thus, the fragmentary self and the integrated self may not be mutually exclusive after all.

Footnotes to Chapter IV

¹The fragmentary theory is also suggested by the work of Rorty, loc.cit.

But cf. Wollheim, Sigmund Freud (New York: The Viking Press, 1971), 210-211, who states that Freud has a principle of constancy in the "Scientific Project" which rules the mind and aims at keeping tension at zero. Later he revised this to minimal tension, but in Beyond the Pleasure Principle he returned to the original position and asserted that sometimes "the mind acts as though it could altogether eliminate tension, as though, in other words, it could reduce itself to a state of extinction." Whether or not Freud's constancy principle is a consistency impulse, and whether his notion of tension reduction means reduction of inconsistency, is not clear. But his accounts of conflict and fragmentation in persons still make him a strong influence on the fragmentary theory.

²Irving Thalberg, "Freud's Anatomies of the Self," in ed. Wollheim, 1974, 147-171, criticizes Freud for partitioning persons into "quasi-persons." He says that Freud personifies the instincts, the unconscious and the conscious mind, the ego, the id, and the super-ego, and that his explanations for dreaming, repression, hysteria, etc., imply divided persons, the fragments of which are personified. Because Freud's view amounts to a multiple self view, Thalberg thinks, his doctrines either cannot adequately explain psychological phenomena, or they are incoherent. Thalberg thinks that the solution is that we must assume the unity of the self. But the fragmentary theory does not deny the unity of the self. It tries to account for conflict and plurality in one self.

³Fingarette, 1969, 85.

⁴Ibid.

⁵Ibid., 87-8.

⁶Watson, loc.cit.

⁷Rorty, 1973a, 63. See also Daniel Dennett, "Conditions of Personhood," in ed. Rorty, 1976, 175-196. Dennett also emphasizes the inevitability of the normative nature of the concept of a person. Because of this, he thinks there is no room for the notion of irrationality in the concept of a person and there is always a difficulty in finding an entity that can instantiate the concept of a person. See also, Rorty, "A Literary Postscript," in ed. Rorty, 1976, 301-323.

Footnotes to Chapter IV, cont.

There, Rorty gives a history of the concept of a person throughout literature in which she emphasizes the contingency of the factors that determine the notion at different times and places and for different purposes. Trilling also discusses the question of the relativity of views of human nature. He points out that sincerity emerges as a value at a certain point in history, and with it comes a certain concept of what not being in self-deception is. On one view of non-self-deception, he says, sincerity is an essential moral quality. This view seems to be associated with the integrated view:

'To thine own self be true'--with what a promise the phrase sings in our ears! Each one of us is the subject of that imperative and we think of the many difficulties and doubts which would be settled if only we obeyed it. What a concord is proposed--between me and my own self: were ever two beings better suited to each other? Who would not wish to be true to his own self? True, which is to say loyal, never wavering in constancy. True, which is to say honest: there are to be no subterfuges in dealing with him. True, which is to say, as carpenters and bricklayers use the word, precisely aligned with him.

Trilling, p.4.

This idea of non-self-deception seems to be similar to Fingarette's view, in which self-deception is a violation of integrity. But Trilling points out that sincerity is devalued in recent history. It is not complex enough and denotes naivete or innocence. Contemporary thought is concerned with authenticity, not sincerity, a more complex notion that takes into account the difficulty of knowing and being oneself, and the complications contributed by social pressures, roles, and inner psychic conflict. It seems that Fingarette takes such factors into account, especially since he emphasizes Freud. But his ultimate judgment of self-deception seems to be based on the values of sincerity and the notion of a true self. Thus, Fingarette is modern in his acceptance of self-deception as a phenomenon that reflects the normal workings of the mind, and that is not paradoxical. But he is not modern in his judgment of self-deception.

Footnotes to Chapter IV, cont.

- ⁸Rorty, 1973a, 65.
- ⁹Ibid., 67.
- ¹⁰Ibid.
- ¹¹Fingarette, ibid.
- ¹²Frankfurt, 1976.
- ¹³Fingarette, 1975.
- ¹⁴Freud's developmental theory implies continuing inconsistency too. See Freud, loc.cit., and Phillip Rieff, Freud: The Mind of the Moralist (Garden City, New York: Anchor, 1961), Ch.II.
- ¹⁵From this point of view, Kant's claim, that as rational beings we really will our maxims to be consistent, is arbitrary. See Immanuel Kant, Groundwork of the Metaphysic of Morals, trans.J.J.Paton (New York: Harper Torchbooks, 1956).
- ¹⁶Marcel Proust, The Sweet Cheat Gone, trans. C.K. Scott Moncrieff (London: Chatto & Windus, 1949), 249, quoted in Derek Parfit, "Later Selves and Moral Principles," in ed. Alan Montefiore, Philosophy and Personal Relations (London: Routledge & Kegan Paul, 1973), 141. Cf. Freud, 1917e; Wood, 211; Trilling, 61, 121-123; and Parfit, 1973.
- ¹⁷Rorty, 1972.
- ¹⁸Rorty, 1972, 1975.
- ¹⁹Rorty, 1973a.
- ²⁰Mead, 135.
- ²¹Ibid., 138.
- ²²Ibid., 140.
- ²³Ibid., 174, 175, 182.
- ²⁴Ibid., 142.
- ²⁵Ibid.
- ²⁶Ibid.
- ²⁷Ibid., 155.

Footnotes to Chapter IV, cont.

²⁸Ibid., 154-155.

²⁹Ibid., 158-159.

³⁰Ibid., 158.

³¹Ibid., 159.

³²Ibid.

³³Ibid., 174.

³⁴Ibid., 182.

³⁵Ibid., 174, 182.

³⁶Ibid., 173.

³⁷James, 173.

³⁸Ibid., 292.

³⁹Ibid., 292-293.

⁴⁰Ibid.

⁴¹Ibid., 293.

⁴²Ibid.

⁴³Rorty, 1973a, 63.

⁴⁴Marcel Proust, Within a Budding Grove, vol.1, trans. C.K.Scott Moncrieff (London: Chatto & Windus, 1967), 349, quoted in Parfit, 141. Parfit has altered the translation somewhat. Parfit points out that people change over time and that personal identity is a matter of degree. This idea is related to the fragmentary theory, in that it takes seriously the idea of persons consisting of different "selves."

⁴⁵James, 315.

⁴⁶Ibid., 293.

⁴⁷Ibid., 294.

⁴⁸Mead, 154-155.

⁴⁹Erving Goffman, The Presentation of the Self in Everyday Life (Garden City, N.Y.:Doubleday Anchor, 1959), 49.

Footnotes to Chapter IV, cont.

50 James, 294.

51 Ibid., 295.

52 Ibid., 294.

53 R.D.Laing, Self and Others (Chicago: Quadrangle Books, 1962), 171. Cf. Erving Goffman, Frame Analysis (New York: Harper & Row, 1974), 59.

54 James, 296.

55 Ibid., 297-298.

56 Ibid., 297.

57 Ibid., 305.

58 Ibid., 291.

59 Ibid., 332.

60 Ibid., 331.

61 Ibid., 330-331.

62 Ibid., 333.

63 Ibid.

64 Ibid.

65 Ibid.

66 Ibid.

67 Ibid.

68 Ibid., 335.

69 Ibid.

70 Ibid.

71 Ibid., 336.

72 Ibid., 337.

73 Ibid., 338.

Footnotes to Chapter IV, cont.

- ⁷⁴Ibid.
⁷⁵Ibid.
⁷⁶Ibid., 339.
⁷⁷Ibid., 340.
⁷⁸Ibid.
⁷⁹Ibid., 341.
⁸⁰Ibid.
⁸¹Ibid.
⁸²Ibid., 342.
⁸³Ibid., 298-299.
⁸⁴Ibid., 371.
⁸⁵Ibid., 447; 452-453.
⁸⁶Mead, 142.
⁸⁷James, 301.
⁸⁸Ibid., 341n.
⁸⁹Mead, 175.
⁹⁰Ibid., 177, 200.
⁹¹Ibid., 178.
⁹²Ibid., 174-175.
⁹³James, 299.
⁹⁴Ibid.
⁹⁵Mead, 196-201.
⁹⁶Ibid., 204.
⁹⁷James, 315-316.

Footnotes to Chapter IV, cont.

⁹⁸Mead, 204.

⁹⁹Mead, 214.

¹⁰⁰Fingarette, 1969, 85-88.

¹⁰¹Watson, ibid.

¹⁰²Ronald B. de Sousa, "Rational Homunculi," in
ed. Rorty, 1976, 217-238.

¹⁰³Ibid., 217-218.

¹⁰⁴Ibid., 221.

¹⁰⁵Ibid.

¹⁰⁶Ibid., 224.

Chapter V

THE FRAGMENTARY THEORY, THE INTEGRATED THEORY AND SELF-DECEPTION

We turned to the fragmentary theory in Chapter IV because it seemed in Chapter III that the integrated theory would have difficulty explaining the fact that self-deception involves the creation and/or perpetuation of conflict. The fragmentary theory is a view of persons that does not require that persons are motivated by an impulse for consistency. Therefore, it seems more hospitable to the possibility that persons are conflicted and that they can create and/or maintain conflict. But we also saw in Chapter III that one thing for which the integrated theory could provide an explanation was the initial motivation for self-deception. In Chapter II, it was established that self-deception grows out of situations in which persons have beliefs or feelings that they do not want to have. Not wanting to have beliefs or feelings that one has can lead to trying to eliminate them, and trying to eliminate dissonant beliefs or feelings can be seen as trying to become consistent. This alone seems to be in line with the claim the integrated theory makes that persons are motivated by impulses to become consistent. But

in addition, the impulse for consistency would entail an aversion to conflict, and so the integrated theory could explain the subsequent self-deceptive denial and masking of these undesired beliefs or feelings as being motivated by an aversion to conflict.

Even though the integrated theory seems, for the reasons just cited, to provide an explanation for self-deception, this seems paradoxical because self-deception involves the perpetuation or creation of conflict. This is the striking fact about self-deception--it involves conflict in such a way that it seems that toleration of conflict in persons is natural and common. Thus, the fragmentary theory seemed more likely to be the description of the self that would fit in with the fact that persons self-deceive. But the fragmentary theory will not be suitable unless it can give an account of the motivation for self-deception, and the integrated theory seems to be able to do this.

The Motivation for Self-Deception

The motivation for self-deception is a problem for the fragmentary theory because it emphasizes plurality in persons and the absence of any consistency criterion for beliefs or feelings. It would seem to follow that persons do not object to any of their beliefs or feelings and that conflict is not necessarily associated with tension. Conflict, according to the fragmentary theory, simply means plurality. Tension

might arise when it is impossible to do more than one thing at once and one has incompatible impulses. This would be inconvenient. Tension might also arise when society requires persons to be single-minded, to act in unconflicted ways, to commit themselves to consistent policies. Social situations and demands might evoke conflict-resolution, suppression of conflict in favor of monolithic policies, and artificially induced impulses to become consistent. Persons might try to induce in themselves the habit of trying to become consistent for practical reasons such as social appropriation and convenience. As Rorty points out, the degree to which persons care about being integrated depends upon the concepts of responsibility that are accepted by society.

These are the kinds of situations that could give rise to tension about conflict in persons according to the fragmentary theory. But such tension would not come from a fundamental and constant inner impulse to be consistent based on an aversion to conflict that characterizes persons regardless of social pressures. And even this kind of tension is not really explained or allowed by a fragmentary theory. Basically, the fragmentary theory seems to provide no reason why persons would mind having any of their beliefs or feelings. Why would the mother not want to believe her son is dead or guilty? Why would the secretary not want to hate her boss? Why would Freud not want to see the train?

From the point of view of the fragmentary theory, it would seem that persons might find certain beliefs or feelings unpleasant, but such an attitude would not be connected necessarily with impulses to eliminate them. Rather, one could just accept one's feelings and beliefs along with one's feelings of displeasure.

Something seems to be missing from the fragmentary theory that is needed to explain the motivation for self-deception. Why would persons deny their beliefs or feelings self-deceptively according to the fragmentary theory? Two possibilities arise in connection with this problem.

The first possibility is that the situation that gives rise to self-deception is not a case of wanting to eliminate conflict. The fragmentary theory might try to argue along these lines. The situation that gives rise to self-deception, as explained in Chapter II, is that persons have beliefs or feelings that they do not want to have. They try to eliminate the beliefs or feelings and they fail. Then they deny them self-deceptively. Why do they deny them self-deceptively? The integrated theory says they do so ultimately because they want to eliminate conflict. But the fragmentary theory might say that they deny that they have the beliefs or feelings simply because they find these beliefs or feelings painful. There is at work no desire to become consistent; there is no attempt to eliminate conflict.

It is simply an attempt to eliminate painful beliefs or feelings.

The integrated theory could respond to this argument, however, in the following ways. First of all, the fragmentary theory has given no explanation for why persons find the beliefs or feelings painful. The fragmentary theory might reply that no explanation is needed. Beliefs or feelings are painful because they are painful. No further explanation is required. Or, the fragmentary theory might claim, many explanations are possible, explanations having to do with situations persons find unpleasant, and not referring to objections persons might have to inner conflict. The mother finds painful having the belief that her son is dead because she misses him. The secretary finds hating her boss painful because it jeopardizes her job. Freud found seeing the train he wanted to miss painful because he wanted to miss it. The fragmentary theory claims that no more of an explanation than these sorts of explanations is necessary or justified.

The integrated theory can counter again with the argument that the situation that gives rise to self-deception, not wanting to have beliefs or feelings that one has, is the same as wanting to eliminate the beliefs or feelings, and that this amounts to wanting to eliminate conflict and to become consistent. The mother believes her son is dead. She does not want to believe this.

Not wanting to believe it leads her to try to eliminate the belief. Having this belief conflicts with her not wanting to have it and her attempt to eliminate it. Thus, she is in conflict, and trying to eliminate the belief is the same as trying to become consistent, to eliminate conflict.

The fragmentary theory responds at this point with the argument that eliminating conflict would be a result of persons' eliminating undesired beliefs, but that it is not part of persons' motivation. Persons in such situations are simply trying to eliminate undesired beliefs--they are not trying to eliminate conflict. That is just a side-effect.

But the fragmentary theory has not explained its acknowledgment that there is conflict here and that conflict is causing tension. According to the fragmentary theory, persons can respond to the tension caused by conflict by accepting it. But we seem to have a different notion of conflict when we link it with impulses that result in eliminating it. If the fragmentary theory acknowledges that the pain of having certain beliefs or feelings can lead persons to try to eliminate them, it acknowledges this different notion of conflict. But it has not explained how it can acknowledge this.

One might think that the fragmentary theory and the integrated theory differ on this issue in the following way: that the argument of the fragmentary

theory commits it to the view that eliminating undesired beliefs or feelings and eliminating conflict are two different things; and that according to the integrated theory, they are one and the same thing. But this interpretation, which might be put forth by integrated theorists, may be objectionable. One possibility is that denying beliefs or feelings because they are painful is different from denying them because they are painful and one wants to eliminate the conflict between having them and having pain about them. On the basis of this distinction, the integrated theory claims that what self-deceivers are doing is the latter, that the fragmentary theory cannot explain this, and that if persons were only doing the former this would not be self-deception. But the other possibility is that doing the former is the same as doing the latter and that the fragmentary theory is overlooking the analytic connection between these two kinds of denial. The integrated theory might have trouble maintaining this claim, because there does seem to be at least a conceptual distinction between trying not to believe or feel something because it is painful, and trying not to be in conflict because it is painful. This is so, even though the result of each endeavor might be the same-- the elimination of conflict. (The question remains as to whether self-deception can be the result of either impulse.) The integrated theory ought to claim that the fragmentary theory cannot explain

either phenomenon.

The line of argument open to the integrated theory is that what self-deceivers object to in the situation that gives rise to self-deception--having beliefs or feelings that they do not want to have--is not just that they have certain beliefs or feelings, but that they are in conflict. It is possible that what self-deceivers cannot tolerate or choose not to tolerate is that they are in conflict. This would explain why they mask their beliefs or feelings and claim to be single-minded. If there were no additional stress experienced by persons as a result of having conflicting attitudes, they might choose to accept the fact that they have conflicting attitudes. The mother might feel that she does not want to believe her son is dead while she continues to acknowledge that she has this belief. What would make her deny the belief self-deceptively? One explanation is that she finds her distress about this belief, in conjunction with having the belief, so intolerable that she wants to mask the conflict. What leads to self-deception is her intolerance of her conflict about having the belief. Otherwise she could simply accept an unpleasant belief.

This argument, put forth by the integrated theory, is more satisfying than the previous argument, because the previous argument relied upon the assumption that not wanting to have a belief or feeling was the same

as trying to eliminate it and that trying to eliminate it was the same as trying to eliminate conflict. Thus, it built the motivation to eliminate conflict into the situation of having beliefs or feelings that one does not want to have, when, in fact, the elimination of conflict is only a result of trying to eliminate undesirable beliefs or feelings. In addition, this argument was problematical because it did not explain masking of the undesirable beliefs or feelings or of the conflict about having them. But now we have an explanation for the masking involved in self-deception: it is that self-deceivers will not tolerate the tension-bearing conflict that can accompany having beliefs or feelings and not wanting to have them. The reason persons do not want to have the beliefs or feelings (reasons might be such things as loss, guilt, etc.) might explain why the conflict is intolerable, but it is plausible that it is ultimately desire to eliminate conflict because conflict is intolerable that motivates self-deception.

Thus, it appears that the most plausible explanation for self-deception is that conflict is unacceptable to those who self-deceive. It follows that the fragmentary theory cannot explain self-deception. Self-deception is not simply an attempt to get rid of undesirable beliefs or feelings. More than that, it is a masking of the fact that persons have these beliefs or feelings

and that they do not want to have them. The reason self-deceivers mask these facts from themselves is, most plausibly, that they have an aversion to conflict. The fragmentary theory shows that self-acceptance of conflict is possible, but it cannot come up with an explanation for the motivation for self-deception.

Another Look at the Integrated Theory and the Conflict Involved in Self-Deception

We have just seen that the integrated theory is more adequate than the fragmentary theory in explaining the motivation for self-deception. But the fragmentary theory does seem to have a better explanation for the maintenance and/or creation of conflict that is involved in the ongoing stage of self-deception. Let us look again at the integrated theory to see what it might say about this.

It could be that the integrated theory's claim is that when persons are aware of conflict they try to eliminate it and that self-deception does not entail that the perpetuation and/or creation of conflict is intentional. Perhaps only the attempt to eliminate conflict that gives rise to self-deception, and the masking of conflict, are intentional.

The fragmentary theory might persist in claiming that maintaining and creating conflict is purposeful. From the point of view of the fragmentary theory, the view of conflict as being purposeful is plausible,

the argument might go, because it captures the nature of ambivalence better, and the aspect of self-deception that involves a claim that fully represents persons (as explained in Chapter II), regardless of whether the beliefs or feelings claimed are held in deep or shallow senses. The argument here is that if persons maintain beliefs or feelings, these beliefs or feelings are maintained purposefully. (Of course, the question remains unanswered as to whether, in self-deception, persons create conflicting beliefs or feelings; the other possibility is that they are created as a side-effect of trying to mask undesirable beliefs or feelings.) But to establish that there is conflict and that conflict involves inconsistent beliefs or feelings to which persons are committed, is not to establish that conflict is purposeful.

The fragmentary theory might attempt to strengthen its claim that being in conflict in self-deception is purposeful by pointing out that self-deceivers endorse what they deny by continuing to believe or feel it, even though they deny it; and they endorse what they avow. The mother continues to believe that her son is dead but she also has created a new belief--that he is alive. That persons subscribe to beliefs or feelings is taken to be evidence that they do so purposefully. But this is an unwarranted assumption. The origin of beliefs or feelings does not have to bear on whether

they are held purposefully. Persons can have beliefs or feelings that they do not have purposefully. They can have beliefs or feelings that they do not want to have.

However, in the case of self-deception, it might be thought that the fragmentary theory has a point after all, since what persons claim to believe or feel when they are in self-deception has been described in Chapter II as being motivated beliefs or feelings. They believe or feel what they claim because they want to.

The integrated theory could respond to this by pointing out that although the creation or perpetuation of the self-deceptively avowed beliefs or feelings is purposeful, continuing to have the beliefs or feelings that persons deny in self-deception is not purposeful. In fact the initial goal of self-deceivers is to not have the denied beliefs or feelings, the integrated theory claims.

The fragmentary theory could respond to this with the claims that creating new beliefs or feelings that conflict with ones that one already has is properly described as creating conflict; that persons know that they have the denied beliefs or feelings (why else would they deny them?), and that there may even be occasions on which persons self-deceive because they want to maintain both the denied beliefs or feelings

and the new ones--in which cases, in essence, their goal is to be in conflict.

In answer to the first claim, the integrated theory could claim that in creating new beliefs or feelings that conflict with ones one already has, one is not purposefully creating conflict if one does not know what one is doing. And the integrated theory could maintain, in answering the second point, that persons do not know what they are doing when they do this. One of the most striking things involved in self-deception is masking from oneself what one is doing. And finally, if there are occasions on which persons self-deceive for the very purpose of wanting to have conflicting beliefs or feelings, why do they bother to mask one of them? Why do they not simply create conflicting beliefs or feelings and accept this?

The fragmentary theory could answer this last point with the claim that the purpose of self-deception in such cases could be convenience, reduction of stress caused by social situations, desire to be able to adapt to different and inconsistent situations, etc. But why is there stress if persons do not have an impulse to resolve conflict? Otherwise they could just act single-mindedly, not bothering to mask inner conflict.

Another argument the fragmentary theory might try to use is that since, in the situation that gives rise to self-deception--having beliefs or feelings that

one does not want to have--there is an alternative to self-deception (namely, self-acceptance,) this implies that self-deception is a choice. Self-acceptance of conflict is, at least, not a situation that involves creating more conflict, whereas self-deception often is. Thus, the self-deceiver purposefully creates conflict.

But the integrated theory responds to this by pointing out that persons who accept conflict do so by default, since they have failed to do what they wanted--to eliminate conflict. Acceptance of conflict does not entail that persons do not have an aversion to it. Furthermore, the argument that being in conflict is purposeful because persons really have each of the conflicting beliefs or feelings, and thus, they endorse each of them, fails because it does not do justice to the fact that persons may have conflicting beliefs or feelings and not want to have both of them. Persons can want to be conflicted. The argument of the fragmentary theory implied that conflict is always purposeful. It rules out the possibility that persons are capable not only of being in conflict but of having an attitude towards being in conflict of either acceptance or rejection.

The question seems to be whether having an attitude of acceptance is something persons do only after failing to eliminate conflict--as a kind of resignation.

This is the view of the integrated theory. The conflict that is created or maintained in self-deception might also be explained as a side-effect of integrative impulses that fail, and not as purposeful. Self-deception would be especially illustrative of the integrated theory because it would show that not only have integrative impulses been at play and failed, but that persons are so intolerant of conflict that they mask it.

What the fragmentary theory describes, it turns out, is a stage or phase persons go through, but not the motivation for the policies that result in such phases. What the fragmentary theory describes--that we can be in conflict--is not incompatible with the integrated theory.

What, then, should we say about the views suggested by James and Mead in Chapter IV, to the effect that there is no consistency impulse in constituting the self? Must we deny the possibility that there might be no conflict-resolving motivation governing the process of constituting the self? A possible answer is that perhaps initial periodic fragmentation is part of the process of becoming integrated, of integrative impulses. It may be that when persons constitute themselves they do not think about consistency, necessarily, as a criterion for that with which they identify. But if conflict arises they will then try to resolve it, for they will feel tension.

The views of James and Mead were cited to show that integrative impulses do not have to govern identification, but we did not show that when conflict arises such impulses do not emerge.

But to allow that persons can identify with things, have beliefs and feelings, etc. that are not consistent, still seems to weaken the claim of the integrated theory. Although the integrated theory cannot be expected to maintain the claim that persons are never in conflict, it cannot allow that persons knowingly create or maintain conflict. This claim is tricky because there still seem to be cases in which persons knowingly maintain or create conflict, and in the case of self-deception, the question of awareness is paradoxical. Self-deceivers must know and not know what they are doing. It was the original claim of this thesis that procedures of self-deception are purposeful, even those that are "unwitting." But to claim that the integrated theory allows conflict as a side-effect of integrative impulses, or perhaps as a misdirected but purposeful attempt at resolving conflict, seems to weaken the claim of the integrated theory. But the claim still contrasts significantly with the fragmentary theory, which denies that there is an antipathy towards conflict and a general motivation towards consistency in persons.

Finally, there is the "true self" problem. In Chapter III, we criticized the integrated theory on the

grounds that it arbitrarily assumed that only the consistent resolution of conflict represented persons. In that view conflict of beliefs or feelings means that only one side of the conflicting feelings or beliefs, or some transformation of both of them until what persons believe or feel is consistent, truly represents persons. So in the case of the mother who denies her belief that her son is dead, the denied belief is her true belief and the belief she claims--that he is not dead--is not really her belief. She claims it mistakenly. And the secretary who denies that she hates her boss and claims to like him really hates him and does not like him. And when Freud claimed to miss seeing the train, he really saw the train, and his claim not to see it was false. One way that we saw that this way of seeing self-deception was confusing was in examining Fingarette's view of self-deception. There we saw that self-deception was described as occurring when one is attracted to an engagement that conflicts with one's governing values.

This could mean that what persons deny in self-deception is not their true belief or feeling. This is one way the integrated theory might try to explain self-deception. On this view, self-deceivers are untrue to themselves because they are not legitimately represented both by the beliefs or feelings they deny and by those they claim. Or, the integrated theory may claim that self-deceivers are not legitimately represented by either

what they deny or by what they claim, but only by some consistent resolution of the conflict. But it turns out that this goes against the interpretation given above--where what is denied is the true belief or feeling. The point is that if the integrated theory tries to say that one side of the conflict is identified with the true self, this claim will be arbitrary. Furthermore, we argued in Chapter II that there was good reason to suppose that self-deceivers believe or feel both what they deny and what they feel.

The assumption that the integrated theory could not accept this feature of self-deception was based upon the assumption that a view that sees the self in terms of impulses for consistency could not identify the self with conflicting beliefs, feelings, or impulses. But now it can be seen (as explained below) that the integrated theory does not have to claim this. The different impulses, beliefs, or feelings that conflict can be truly representative of persons, while integrative impulses are also truly representative of persons. Fingarette's view that self-deception involves being attracted to engagements that go against one's governing impulses is acceptable only in the sense that it means that persons have beliefs or feelings that they do not want to have. The view his claims suggested--that the denied feelings or beliefs are not part of persons, and that the beliefs or feelings claimed are claimed falsely--is now found to be not only unwarranted,

but not necessarily implied by the integrated theory. The integrated theory, as we discussed it in Chapter III, was essentially claiming that persons do not self-deceive, rather individuals, or regressed, or dissolved persons do. The kind of beings the integrated theory saw as self-deceivers is the kind of beings described by the fragmentary theory when it described persons. In the earlier interpretation of the integrated view, the beliefs and feelings that are denied do not properly belong to persons. Sometimes the claim also seems to be that the beliefs or feelings that are claimed are claimed falsely, so they do not properly belong to persons either. The result is that neither of the conflicting beliefs or feelings belong to persons, and persons are represented only by some consistent resolution of this conflict.

But the present interpretation of the integrated theory allows that persons can be represented by what is denied and by what is claimed. That is, it allows that persons can have conflicting beliefs or feelings. How can it allow this if it posits a constant impulse for consistency? It does so with the special explanation it gives for masking. More specifically, it does so by allowing that masking of conflict is consistent with the consistency impulse. The reason self-deceivers hide from themselves that they are in conflict and that they are denying they are in conflict is because persons have an aversion to conflict and they make constant

attempts to eliminate it when they are aware of it. If they cannot eliminate it and the aversion is strong enough, they will attempt to mask it. But what they mask is a genuine part of themselves. So the integrated theory is compatible with persons' being in conflict and with their inadvertently perpetuating or creating it. It is compatible with the kind of plurality in the self that is pointed out by the fragmentary theory. We need not continue to think that the view that there is a consistency impulse that characterizes persons implies that anything short of consistency does not truly represent persons.

Thus, in self-deception, persons have beliefs or feelings that they do not want to have. This is a conflict. If they fail to eliminate the conflict and they attempt to mask it, they do so because they have an aversion to conflict. It is not simply that the beliefs or feelings that they do not want to have lead self-deceivers to mask them. These beliefs or feelings, plus persons' desires not to have them, and their intolerance for this conflict, lead to masking. A consistency impulse, which includes an aversion to conflict, motivates self-deception, and in this sense the integrated theory seems to explain self-deception. However, the self-deceptive denial, which is based on interests in denying conflict, involves the perpetuation of the conflict, if it does not also lead to the creation of more conflict.

This was the part of self-deception that the fragmentary theory seemed to explain. But it is possible that the creation and perpetuation of conflict is an inadvertent result of self-deceptive masking. It is not the goal of self-deceivers to create or perpetuate conflict; rather, their goal is to mask conflict that they have been unable to eliminate and of which they are intolerant. The conflict they inadvertently perpetuate is the very conflict they purposefully try to eliminate. The conflict they create is likely to be undesirable to them. This explains why they mask it. Therefore, it is not likely that it is purposefully created. We now see that the fragmentary theory does not explain the motivation for self-deception, because it cannot explain how persons could object to conflict. Insofar as the fragmentary theory describes the possibility of conflict in persons, it is compatible with the integrated theory. Persons can be in conflict. But as a theory of persons' attitudes and impulses towards conflict, it differs from the integrated theory and it is incompatible with self-deception.

Summary

We started out with the problem of trying to explain the element of conflict in self-deception. In particular, we wanted to know what kind of attitude persons have towards inner conflict and what kind of

theory of persons, vis à vis conflict, would be compatible with and explanatory of self-deception. Self-deception was chosen as an example of inner conflict. As it was defined in Chapter II, persons in self-deception have conflicting beliefs or feelings. This led to speculations about what kind of attitudes persons have towards conflict. According to one view of the self, the integrated theory, which is described in Chapter III, persons constantly try to eliminate conflict. At first this view seemed both appropriate and inappropriate for self-deception. It seemed appropriate in that it was consistent with the motivation involved in the situation that gives rise to self-deception. That situation is that persons have beliefs or feelings that they do not want to have and they try to eliminate them. Trying to eliminate them can be seen as trying to become consistent. It is because this attempt fails that persons turn to self-deception.

Up to this point in the process, the integrated theory seemed appropriate for self-deception. But the next stage of self-deception involves the perpetuation or creation of conflict. Initially, self-deceivers are either doubly conflicted in the sense of having inconsistent beliefs or feelings and not wanting to have them, or they are simply conflicted in the sense that they have beliefs or feelings that they do not want to have. In the former case self-deception involves perpetuating conflict. If one loves and hates one's boss and does

not want to hate him, and if this situation leads to self-deception, the result is that one will continue to both love and hate one's boss. In the latter case, self-deceivers create conflict. If one hates one's boss and does not want to hate him, and if one self-deceives, the result is that one both hates and does not hate one's boss. The argument that self-deceivers believe or feel what they claim self-deceptively, as well as what they deny, was presented in Chapter II. The view there was that self-deception involves conflict and that conflict means that persons really have the beliefs or feelings on both sides of each pair of conflicting beliefs or feelings.

It may seem strange to point this out. It may seem that this is obviously what conflict means--that persons have different beliefs or feelings that conflict. But the view that emphasizes consistency motivations in the self tends to suggest the implication that in conflict certain beliefs or feelings do not really belong to persons. We are reminded of Frankfurt's distinction between the internal and the external, Fingarette's distinction between individuals and persons, and Watson's distinction between the valuing and the desiring systems of persons. Since, in such views, a consistent resolution of conflict really represents persons, one might think that the implication is that beliefs or feelings that conflict are not regarded as all be-

longing to persons.

The view that conflicting beliefs or feelings do represent persons seemed incompatible with an integrated theory which seemed to imply that because persons are identified with consistency impulses they could only be truly represented by consistent beliefs or feelings. Thus, the integrated theory seemed to have a different view of conflict than that presented in the description of self-deception and by the fragmentary theory. Thus, we turned, in Chapter IV, to a description of the fragmentary theory, which emphasized plurality in the self, which implied the likelihood of conflict, and which did not posit a constant consistency impulse. We argued that a fragmentary theory is not a multiple self and we tried to show that the fragmentary theory can explain how policies can be made when there is conflict, without appealing to a consistency impulse. There was difficulty with this last claim. We discussed the possibility of a surveying and ranking type of policy-making that would not eliminate conflict, but it was difficult to understand this kind of policy-making without appealing to a consistency impulse. This was due to the fact that surveying and ranking seemed to require choice and criteria for choice, and to have one policy and not another seemed to require criteria for choice that are consistent.

With this reservation about the fragmentary theory in mind, we set out in this chapter to examine its compatibility with self-deception, and we found that although it describes the situation of being in self-deception, in which conflict is maintained or created, it does not explain the motivation for self-deception. We also saw that the conflict involved in self-deception is not incompatible with the integrated theory. The integrated theory is easily interpreted as a view that suggests that persons cannot be in conflict, but having a constant impulse to be consistent does not really imply this. The integrated theory can acknowledge that persons can be in conflict and that they can be truly represented by conflicting beliefs or feelings. Therefore, the integrated view can acknowledge that self-deceivers have conflicting beliefs or feelings. It need not take the view that self-deceivers are making false claims in the sense that they are claiming to believe or feel what they do not believe or feel. Rather, the integrated theory can acknowledge what was pointed out in Chapter II about self-deception, and what the fragmentary theory seemed to be able to acknowledge too --that the dishonesty of self-deception is that self-deceivers claim to be single-minded when they are not. It is not that the secretary is claiming a simple falsehood when she says she likes her boss, but that her claim is incomplete, because she also hates him. The

mother is not claiming a simple falsehood when she says she believes her son is dead, or not guilty. Her "error" is that she is failing to present the whole picture: that she also thinks he is alive or guilty. Freud was not wrong when he said he did not see the train. But the whole truth was that he saw it and did not see it. Finally, the integrated theory explains the masking of undesirable beliefs or feelings as a project that grows out of the consistency impulse. Persons have an aversion to conflict and when they cannot eliminate it, they might try to mask it from themselves. This is compatible with our view that self-deception involves presenting oneself to oneself as if one were unconflicted about one's feelings or beliefs.

Chapter VI

CONCLUDING REMARKS

Originally this project involved the assumption that the integrated theory and the fragmentary theory were mutually exclusive theories about how persons handle conflict. That assumption is challenged by the end of the present work. But it may still be arguable on some level. Originally, the integrated theory was defined as the view that persons have an aversion to conflict which manifests itself in their trying to eliminate it. The fragmentary theory was defined as the view that persons do not always try to eliminate conflict -- they do not always have an aversion to it and they can accept it, seek it out, or create it. Going by these definitions, both theories could explain the interpretation of self-deception that is spelled out at the end of this project in Chapter V. Self-deception requires that persons have an aversion to conflict that motivates the self-deceptive denial, and the fragmentary theory as defined above allows that persons sometimes can have an aversion to conflict. But the fragmentary theory was rejected in Chapters IV and V on the grounds

that it does not allow that persons ever have an aversion to conflict. This shift in the conception of the fragmentary theory is due to the fact that, in the end of Chapter IV, the fragmentary theory seemed incoherent when it claimed that sometimes people do not have an aversion to conflict. There it was argued that it is possible that an aversion to conflict underlies all other attitudes--it is a constant born of our capacity for rationality and of the unity of the self. The idea of a person who, at times, is devoid of any concern for lack of conflict is incoherent. Even those cases that seem to be such on the surface seem to require an underlying aversion to conflict.

It was suggested in the end of Chapter IV that the fragmentary theory as originally defined cannot account for the sometime aversion to conflict, unless it appeals to an overall aversion. Thus, it may be that we should not allow that persons can have an aversion to conflict only sometimes. Thus, the only coherent version of the fragmentary theory that contrasts with the integrated theory is the view that persons never have an aversion to conflict. As such, it certainly cannot account for self-deception.

The original version of the fragmentary theory seemed, thus, to be incoherent. In another interpretation, to be discussed below, the fragmentary theory and the

integrated theory seem not to be mutually exclusive theories, but views that emphasize different things about persons.

Originally, however, the fragmentary theory and the integrated theory were taken to be, possibly, different theories. The prevalence of conflict in persons, as pointed out by James and Mead and others, caused us to consider the possibility that people do not mind it and that they might even seek it out. On the other hand, notions about identity, rationality, judgment, and the development of the person as discussed by Fingert, Frankfurt, and Watson, suggested the integrated theory, according to which conflict is associated with tension and desire to eliminate it. It was only through discussion and elaboration of these views that it became apparent that the integrated theory was compatible with persons being in a great deal of conflict and that it is even compatible with self-acceptance and creation of conflict (if these things are seen as existing against a basic background of aversion to conflict and perhaps failed attempts to avoid or eliminate it). And the fragmentary theory (as pointed out in the end of Chapter IV)--or any theory of the self--seems to require a consistency impulse for policy-making, unity, and handling of conflict. The integrated theory and this version of the fragmentary theory, thus, seem to emphasize different characteristics of persons that, prior

to careful examination, can seem to conflict. But the evidence for each of these so-called theories turns out not to be mutually exculsive. This is one interpretation.

Another way of looking at this problem is to say that the fragmentary and the integrated views are not really different theories but that they are different emphases that perhaps interpret the same phenomena in somewhat different ways. The crucial case is that in which the fragmentary theory claims that persons do not have an aversion to conflict (while it allows that sometimes they do). According to the integrated theory, the phenomena that merit this description on some level (cases in which persons acknowledge that they are conflicted, and claim and seem not to mind it, or cases in which persons seem to purposefully create conflict) are possible, but ultimately there is an aversion to conflict even in these cases. The integrated theory might claim this using the argument that a sometime aversion to conflict is not possible and that personhood requires such an aversion as a fundamental. When conflict avoidance or resolution fails, persons may resort to strange means--such as self-deception--in which more conflict is apparently inadvertantly created. If this is possible, it is also possible that, in order to reduce tension, and in the face of the inability to eliminate or avoid conflict, persons might choose to try to cultivate a tolerant

attitude towards it. This might extend to persons actually trying to create conflict in themselves, and to adjust to being conflicted without minding it. Such phenomena are, according to the integrated theory, learned adaptations to difficult situations. They are things persons do in spite of their overwhelming aversion to conflict.

What is the evidence for such an aversion? This is a difficult question. The fragmentary theory and the integrated theory allow that the same phenomena exist. Both agree that there could be a man who claims that he both approves and disapproves of monogamy, and who does not mind being in this conflict. But the integrated theory goes a step further and claims that originally the person did mind such conflict, or perhaps that the aversion to it is repressed or ignored temporarily. He does not mind his conflict only against the background of minding it.

The two views do not disagree about some of the data but they involve different conceptions of what ultimately motivates persons. If the fragmentary theory were to say that there is no ultimate aversion to conflict in cases like the one cited above, the integrated theory would say that such cases do not exist.

Therefore, if these are two mutually exclusive views, the integrated theory says that what the fragmentary theory describes--that sometimes we do not have an ultimate aversion to conflict--does not exist. This is how the views might differ but then they would not differ significantly, on these grounds, for self-deception, for then both views could still claim that self-deception is an example of aversion to conflict.

But in this view of the fragmentary theory, as mentioned earlier, the fragmentary theory is incoherent and it can only become coherent if it is changed into an integrated theory or an extreme fragmentary theory (claiming that persons never have an aversion to conflict). This is why it was rejected as an explanation for self-deception--of its two coherent forms the only one that can explain self-deception is the one that does not differ from the integrated theory. The extreme form of the fragmentary theory cannot explain self-deception because the element of masking, which is essential to self-deception, seems to be best explained by an aversion to conflict.

Thus, what remains of the fragmentary theory if the incoherent claim that we only sometimes have an aversion to conflict is eliminated, is the extreme view that persons never have an aversion to conflict. This view cannot be compatible with self-deception. But as

we saw in Chapter II, in the early part of this investigation, the motivation for self-deception was not obvious (it was not clear whether the conflict that self-deception created was created purposefully). And so it was plausible to explore the extreme fragmentary theory's suggestion that self-deception was an attempt to create conflict.

We might say that the integrated theory claims that either there are no attempts to create conflict or that such attempts are reactions to a basic aversion to conflict that cannot succeed. The prevalence of conflict that was discussed in Chapter IV does not count as evidence for the fragmentary theory and against the integrated theory, as was suggested in Chapter IV. Although the fragmentary theory and the integrated theory are not mutually exclusive theories, they are based on different interpretations of the same phenomena. For example, the integrated theory may describe conflict situations as involving aversion to them, while the fragmentary theory would leave this out. These two views might be useful ways of emphasizing different conceptions of human nature, and these different emphases might have different consequences. The integrated theory, in emphasizing aversion to conflict, can be mistaken for a view that does not recognize persons' capacities for being in and dealing with conflict. The fragmentary

theory, in emphasizing conflict, can be misinterpreted as a view that does not recognize a basic aversion to conflict. Self-deception is a phenomenon that is illustrative of these two different emphases and of how the aversion to conflict can, almost paradoxically, result in more conflict.

At first it seemed that the integrated and the fragmentary theories were two profoundly different philosophical theories about the self. After working through these problems in this dissertation, however, it appears that the fragmentary and the integrated theories do not conflict and that they only differ in terms of what they emphasize. It should be recalled at this point that we talked about three different versions of the fragmentary theory. The present remarks apply to only the version that claims that as a matter of fact sometimes we avoid conflict and sometimes we do not. This view, we have argued, is compatible with the view that even when we do not avoid conflict we still have a basic aversion to it. (The other versions of the fragmentary theory are that people never avoid conflict or that sometimes when people do pursue conflict they do not have a basic underlying aversion to it.) The philosophical task has been to show that although the fragmentary theory and the integrated theory appeared to be two conflicting theories, they do not conflict.

Since each of these views places emphasis on different aspects of the way persons are involved with conflict we might ask about the merits of each approach. When it comes to understanding the motivation for a phenomenon like self-deception, the integrated theory seems to involve the appropriate emphasis, because what is essential to self-deception is masking, and this, as we have explained in Chapter V, is best understood as being motivated by an aversion to conflict. But when it comes to understanding the condition persons are in when they are in self-deception, the emphasis on conflict that we get from the fragmentary theory seems appropriate.

Perhaps the point is that these two views are mistakenly polarized and to get a complete picture with accurate emphases we need to be reminded that they do not exclude one another. What is important here is that irrationality occurs against a background of rational aims and requirements. A similar point is made by Rorty in "Belief and Self-Deception," (p.303). The facts that persons are so very conflicted and that they seem to resemble multiple selves can be taken as starting points for questions about whether persons are indeed split up into different parts, and how this can be. But if we argue that it is conceptually impossible for a single unified person to be split up into different parts (as if the different parts were different motiva-

tional systems with allegiances to different poles of the conflict), we can make the further claim, which is illustrated in the present dissertation, that conflict in a person is only possible because the person is unified, aims at consistency, and has an aversion to conflict.

As Rorty points out (and paradoxical as it might seem), we are capable of a huge range of irrationality and conflict. It is possible that we are irrational and conflicted most of the time. As long as we have the concept of belief (which entails truth orientation and consistency) we can misapply it.

In considering the fragmented conditions persons might find themselves in and the conflicting attitudes they might have, it is possible to lose sight of the fact that there is an underlying integrating impulse. Such an impulse may not be empirically discoverable. It may have to be evoked or taught. It may be very difficult to determine which fragment, attitude, role, or identity of a person is the one that can or should pull all the other ones together, and it may be that, for various kinds of circumstantial reasons, some persons are incapable of such integration. Perhaps most persons are incapable of it. It is probably a matter of degree. Some persons may have more frequent and longer periods of more pronounced integration than others.

It does not seem to be the case that self-deception necessarily increases the disintegration of persons even though it is irrational (it involves conflicting attitudes) and thus promotes fragmentation. Inability to function or lack of a sense of oneself as being a single self do not have to be a function of self-deception. This might rather depend upon the sorts of things persons are deceiving themselves about and how well this fits in with their social situation, the expectations others have of them, and other such considerations. Self-deception is often an adaptive measure and from this point of view it does not seem to be irrational. Condemnation of self-deception may not always be appropriate since the kind of honesty and authenticity that would be required without it might be immobilizing. Self-deception can be harmless to others, and the greatness of soul required by those who might recommend that we aim at eliminating self-deception might be the biggest self-deception of all. We are rational beings because (for one thing) we have a concept of belief, which means that we think it contradictory to claim to have a belief and to simultaneously deny its truth. We have an aversion to conflict. But we are not perfectly rational beings and we are thus capable of being conflicted, inconsistent, in self-deception, irrational, when we are trying to do just the opposite. Trying to be consistent is no guarantee of success. Strange as it may seem we often create the

very conditions we are trying to avoid. Thus, it is possible that we are often irrational and conflicted, but this does not mean that we do not have a basic aversion to conflict.

Finally, then, it appears that the question of whether there is a basic aversion to conflict in persons is a conceptual question rather than an empirical one. The philosophical task has been to clarify the possible misunderstandings and to show why the fragmentary theory and the integrated theory are not profoundly different theories about the self. Empirical evidence will not establish one or the other as it can be interpreted either way. We can describe a case of conflict without appeal to aversion to it if we do not go to that level or analyze back far enough. Or we can describe conflict always with an acknowledgment that there is a basic aversion to it. Any attempt to show empirically that there is no such aversion can always be countenanced with the conceptual argument that this is ultimately incoherent.

It might be helpful, at this point, to look again at some of the examples of self-deception to which we have referred throughout this dissertation, with the purpose of gaining a deeper understanding of the claim that there is a fundamental aversion to conflict in persons. When we talked about the

attitudes that conflict we restricted ourselves to talking about beliefs and feelings, for the sake of simplicity. We saw that conflict and self-deception raised somewhat different points, in some respects, depending upon whether the conflict or self-deception were about beliefs or about feelings. But in our conclusion we want to claim that persons have an aversion to conflicting attitudes in themselves, whether the attitudes are feelings or beliefs. First let us look at some of the examples of conflict of belief.

The mother who knows her son is dead but who self-deceivingly denies this and thereby claims that he is alive is maintaining conflicting beliefs. The belief that he is dead persists but she distances herself from this belief, via self-deception, and she consciously thinks of herself as believing that he is not dead. She can only maintain these conflicting beliefs through self-deceptive masking for, being a person means that she has a concept of belief from which it follows that if she thinks she believes something, she cannot think that it is also not true. Thus, she cannot simultaneously think that her son is dead and that it is not true that he is dead. The masking that takes place in self-deception shows that in self-deception the person has an aversion to conflict, but does it show that cases of non-self-deception also illustrate

an aversion to conflict? Is it conceivable that this mother could say, non-self-deceptively, "I know that I believe my son is dead and that I also believe he is alive, I know that I am maintaining conflicting beliefs, and that this is irrational, contradictory and even incoherent, but I cannot deny that it is true of me that I have these conflicting beliefs." She might also add "I will do nothing to try to resolve this conflict and I will allow or encourage myself to switch back and forth between these beliefs in different situations, as it is convenient to me." Would it be justifiable to claim that there is a fundamental aversion to conflict in such a case? Would it not be that such a case provides empirical evidence against the claim that persons always have an aversion to conflict?

Even in such a case, the present argument is that there is an aversion to conflict. For when the mother says she knows she has the conflicting beliefs that her son is alive and dead, it could be argued that she is not really endorsing these beliefs simultaneously. Rather she is reporting to us some information about her attitudes and behavior patterns. She is reporting that she is prone to self-deception in this matter and that she is not trying to eliminate this propensity. Her not trying to eliminate the propensity need not be taken as evidence for the lack of

an aversion to conflict. It is likely that she is expressing resignation at being unable to resolve the conflict. She may recognize that her aversion to accepting her son's death is so strong that she, at times, cannot control herself and actually, because of this aversion believes he is not dead. She also recognizes the demands of rationality and the meaning of the concept of belief and, thus, she recognizes that he is dead and that she cannot claim to have both beliefs simultaneously in a full-fledged way. To say "I know I believe p and not-p" one must be hedging about at least one of the beliefs. One must be holding back endorsement and one must be merely reporting that something strange is going on-- something that cannot quite be, without masking.

Do the same remarks hold for other cases of conflict of belief? In the case of the mother and her son's death, the fact of the death is so final and definite that maintaining the conflicting beliefs that he is dead and not dead seems to be obviously impossible. But what of some of the other cases in which there were conflicting beliefs. We spoke of the mother (Rorty's Mrs. Faith Beloyal) who was deceived about her son's guilt, and the mother who was deceived about her son's musical ability. In these cases, at first glance, it might seem that it is possible, without

incoherence, to maintain "I believe my son is guilty (competent in music) and that he is innocent (incompetent)." One is either dead or alive, but when it comes to guilt or competence we might be talking about different senses or degrees of each. This line of thinking is erroneous, however, and these cases are to be analyzed just as the case of conflict about death was analyzed. The point here is that what the person is claiming conflictingly is that he or she has and does not have certain beliefs. The question is not whether the son is guilty or not guilty in some sense, but what beliefs the mother thinks she has about his guilt. Does she claim to believe he is guilty or not? The same argument holds for the death and competence cases.

An example that involved conflict of intentions was the case cited from Stekel in which he "accidentally" untied a woman's robe as he shook her hand. He told himself that he was intending to shake her hand when he was really intending to untie her robe. Here we have a case of two different beliefs being claimed about one's intentions. But do these intentions conflict? Couldn't they both coexist without masking or tension? Couldn't a person intend to do both--shake hands and untie the robe? Certainly such intentions and beliefs about them need not conflict. Whether or

not they do conflict depends upon the world view and values of the person in question. Two such intentions might be unacceptable to the person or inappropriate in the situation. A social situation that calls for a handshake is not likely to be one that calls for simultaneous undressing of the person whose hand is shaken. Given such conditions such intentions conflict and Stekel's belief that he has the one intention is incompatible--for him--with his belief that he has the other one. This example is very important because it emphasizes that we are talking about psychic conflict in this dissertation and that psychic conflict can be relative and is not to be confused with logical contradiction. On the other hand what is common to this example and the others previously cited is that the person is claiming that he believes that one thing is his attitude (in this case, his intention to shake hands) while his actions reveal that something else (that he regards as incompatible with the first thing) is also his intention.

Similarly, in the case of conflicting beliefs, one claims to believe something (that one's son is dead, guilty, competent) that one regards as incompatible with something else that one believes. It happens that in the case of belief the beliefs one regards as incompatible are also conceptually incoherent if claimed together.

Similar remarks hold for the cases mentioned earlier of the adulterer who thinks he is being constructive in being honest with his wife while he is really trying to cause pain; and the adulterer who claims that there is no change in his marriage because he is being honest about his adultery. These men claim to have certain attitudes that they would regard as incompatible with the attitudes they deny having (but continue to have). The adulterer who claims to have constructive intentions has an aversion to thinking of himself as also having hostile intentions. The adulterer who claims that his marriage is unchanged has an aversion to also recognizing that a significant bond of trust has been altered.

Finally we cited the example of Freud missing his train because he failed to notice it, even though he saw it. This case bears a strong resemblance to the case of the mother and the death of her son, in that in order to claim to perceive something it seems to be a requirement that one not claim not to see it. It seems to be a conceptual inconsistency to say one sees the train and does not see it. Hence it would be difficult to understand what might be meant if Freud were to say that he knows that he sees and does not see the train unless this were understood in the manner suggested above for understanding the mother who

says "I know I harbor conflicting beliefs about my son's death, and that I switch back and forth between them." In such a case, Freud would not be actually claiming to perceive and not to perceive the train simultaneously. Rather, he would be reporting about one of his attitudinal and behavioral patterns.

In the above discussion we see what is meant by having a fundamental aversion to conflict of beliefs (and of intentions in the example from Stekel) as illustrated in some of the examples referred to in this dissertation. We now ask whether the same claim can hold for feelings. Since conflicting feelings do not seem to involve a logical or conceptual inconsistency, it is not immediately apparent that the same claim does hold. We are used to the idea that love and hate coexist simultaneously in persons, for example. This may not even seem strange to us and we may come to recognize it as the way people are. There is no apparent contradiction involved. Thus, in the case of the secretary who deceives herself about hating her boss, we see that she has an aversion to acknowledging her hate and to saying "I know I both love and hate him." But it does not seem incoherent for a person to make such an acknowledgment.

There are several reasons why we might argue that conflict about feelings involves a basic aversion to conflict, just as in the case of conflict of belief.

First of all, what is important is that the person finds the feelings incompatible. If the person sees the situation this way, then, of course, there is an aversion to conflict. And we might wonder what conflict could mean, if not that the person will see the conflicting entities--whether they are beliefs or feelings--as incompatible. And if they see the situation this way they will have an aversion to saying "I acknowledge that I both love and hate my boss and I don't care that I have this conflict, I will do nothing to try to eliminate it, I may even try to increase it." The person might say something similar to this in an expression of acceptance of the conflict but only on the basis of an ultimate aversion to it.

Another point is that it is possible that feelings are based upon beliefs or that they are intimately connected with them, so that whatever applies to conflicting beliefs would also have to apply to conflicting feelings. The reasons the secretary might hate her boss are beliefs she has, such as: "he treats me unfairly," "he ought not to do so," "he's a rotten person;" and feelings of hate might be accompanied by or responsible for beliefs about one's intentions

or policies, such as: "I will not let him take advantage of me," "I will not do anything for him," "I do not want to help him," "I want to harm him," etc. Such beliefs about what one thinks about a person and what one desires or intends to do, are potentially incompatible with other beliefs one might have about the same issues. So feelings may conflict, if only because of their connection with beliefs.

Another point is that it is likely that persons have a fundamental aversion to having conflicting feelings, because having such feelings is inconvenient. The angry secretary is a good example. It might jeopardize her job to hate her boss, and it might make her job even more complicated to both love and hate him. The erratic behavior or the strain of avoiding it would potentially cause trouble with others.

The objection might be posed at this point that although having conflicting feelings might at times be inconvenient, it might also, at times, be convenient. In these latter cases, the claim might continue, there would be no aversion to conflict. First of all, what might count as a case of it being convenient to have conflicting feelings? If one is trying to please different groups or receive benefits from different situations it might be convenient to keep conflicting feelings alive so that one can switch from one to the

other when the appropriate beneficial situation arises. But the convenience of such conflict in this respect does not obviate the likelihood, or perhaps necessity, of basic psychic inconvenience. It is likely that it is still psychologically difficult to maintain or accept such conflict and convenience gained in this way may be costly. Divided allegiances in oneself means tension. It requires maintaining (if not arduously cultivating) an alertness to the necessity to switch to the right frame of mind for different circumstances. A related kind of self-discipline is probably a necessary ingredient in any mature person's approach to life, but maintaining a number of points of view with no solidly grounded central one would seem to involve more of a strain than sticking to a single frame of mind in spite of clashing circumstances. One might object that such a psychic attitude is non-adaptive and inconvenient. Maybe it is--but it still may reflect a basic aversion to conflict, and the difficulty of the so-called adaptive attitude might be evidence for this basic aversion. Inner integrity seems to be a stronger position than fragmentation, but whether this is true or not, and which position is more adequate or adaptive to the circumstances in which persons find themselves, do not affect our claim that there is a basic aversion to conflict. However, if inner integrity is a stronger

position, then we could say that conflict threatens strength and is thus something that causes tension and that we want to avoid.

One final point is that an objection might be made to the whole approach of this dissertation to the effect that there really is no issue because conflict has been defined in such a way as to include the claim that persons have a basic aversion to it. In claiming that this is a conceptual question rather than an empirical one, the present position appears to be vulnerable to this objection. The objection is that all of the cases cited as counterexamples to the claim that persons have a basic aversion to conflict are rejected, it seems, on the grounds that what is meant by conflict includes aversion to it.

The defense of the present view is that one can meaningfully ask whether there is an ultimate aversion to conflict because there are cases that, on some level and in some sense, do not involve an aversion to conflict. There are different levels of conflict and aversion (or non-aversion) to it. The woman who recognizes that she both loves and hates her boss and that she is unable to reduce or eliminate the conflict does well to accept the fact that she is in conflict. If acting on the motivation of trying to eliminate it has failed, it will cause more tension to continue to act on it. On

this level, her acceptance of her conflict reveals an absence of an aversion to conflict. But this is only so because of a pre-existing and basic situation in which she did have an aversion to conflict. The philosophical task is to show that there are these different levels, to sort out the empirical and conceptual questions, and to show what bearing the empirical data has on the conceptual question. The philosophical task is to show that there is only a relative and contingent sense in which there is an aversion to conflict.

In summary, then, in this chapter we have discussed the differences between the integrated theory and the fragmentary theory, the significance of their different emphases, and the kind of evidence that exists for each. We also have discussed the significance of self-deception and these two views about the self in illuminating a basic precept of rationality and its relationship to irrationality and we have argued that rationality and unity in the self require a basic aversion to conflict.

BIBLIOGRAPHY

- DENNETT, DANIEL. "Conditions of Personhood, " in Amelie Okensberg Rorty, ed., *The Identities of Persons*. Berkeley: University of California Press, 1976, pp. 175-196.
- DE SOUSA, RONALD, B. "Self-Deception." *Inquiry*, XIII, 1970.
- _____. "Rational Homunculi," in Amelie Okensberg Rorty, ed., *The Identities of Persons*. Berkeley: University of California Press, 1976, pp. 217-238.
- FESTINGER, LEON, "Cognitive Dissonance," *Scientific American*, 207:4, October 1962, pp. 93-102.
- FINGARETTE, HERBERT. *The Self in Transformation*. New York: Harper Torchbooks, 1965.
- _____. *Self-Deception*. London: Routledge & Kegan Paul, 1969.
- _____. "Self-Deception and the 'Splitting of the Ego'," in Richard Wollheim, ed. *Freud--A Collection of Critical Essays*. New York: Anchor Books, 1974, pp. 80-96.
- FRANKFURT, HARRY G. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy*, LXVIII:1, Jan. 14, 1971.
- _____. "Identification and Externality," in Amelie Okensberg Rorty, ed., *The Identities of Persons*. Berkeley: University of California Press, 1976, pp. 239-251.
- FREUD, SIGMUND. (1901b, 6/1). *The Psychopathology of Everyday Life*, trans. by Alan Tyson, ed. by James Strachey, New York: W.W.Norton & Co., 1965.
- _____. (1916, 12/257). "A Note on the Unconscious in Psychoanalysis." *Collected Papers*, IV, London: Hogarth Press, 1950.
- _____. (1915, 14/161). "The Unconscious" in Joan Riviere, trans. *Collected Papers*, IV, London: Hogarth Press, 1950.

- _____. (1915d, 14/143). "Repression." Joan Rivers, trans. Collected Papers, IV. London: Hogarth Press, 1950.
- _____. (1917e, 14/239). "Mourning and Melancholia." James Strachey, ed. The Standard Editions of the Complete Psychological Works of Sigmund Freud, London: Hogarth, 1953-.
- _____. (1921c, 18/67). Group Psychology and the Analysis of the Ego. Standard Edition.
- _____. (1923b, 19/3). The Ego and the Id. J. Riviere, trans. J. Strachey, ed., New York: Norton, 1962.
- _____. (1925h, 19/234). "Negation." Standard Edition.
- _____. (1930a, 21/59). Civilization and Its Discontents. trans. and ed by James Strachey, New York: Norton, 1969.
- _____. (1940a, 23/141). An Outline of Psychoanalysis. trans. and ed. by James Strachey, New York: Norton, 1969.
- _____. (1940e, 23/273). "Splitting of the Ego in the Process of Defense." Standard Edition.
- GARDINER, PATRICK. "Error, Faith and Self-Deception." Proceedings of the Aristotelian Society, LXVIII, new series, 1969-70.
- GOFFMAN, ERVING. The Presentation of Self in Everyday Life. Garden City, N.Y.: Doubleday Anchor, 1959.
- _____. Frame Analysis. New York: Harper & Row, 1974.
- HARMON, GILBERT, "Knowledge, Reasons and Causes," Journal of Philosophy, LXVII: 21, Nov. 5, 1970.
- HAMLIN, D.W. "Self-Deception." Proceedings of the Aristotelian Society, supp. vol. XLV, 1971.
- JAMES, WILLIAM. The Principles of Psychology. Vol. I., New York: Dover Publications, 1950.
- KANT, IMMANUEL. Groundwork of the Metaphysic of Morals. trans. H.J. Paton, New York: Harper Torchbooks, 1956.
- KIERKEGAARD, SOREN. Purity of Heart. New York: Harper Torchbooks, 1956.
- LAING, R.D. Self and Others. Chicago: Quadrangle Books, 1962.

- MEAD, GEORGE HERBERT. *Mind, Self, and Society*. ed. Charles W. Morris. Chicago: University of Chicago Press, 1962.
- MOUNCE, H.O. "Self-Deception." *Proceedings of the Aristotelian Society*. supp. vol. XLV, 1971.
- PARFIT, DEREK. "Personal Identity." *Philosophical Review*. LXXX, 1971.
- _____. "On the Importance of Self-Identity." *Journal of Philosophy*, XXI, Oct. 1971.
- _____. "Later Selves and Moral Principles," Alan Montefiore, ed. *Philosophy and Personal Relations*, Montreal: McGill--Queens University Press, 1973.
- PENNER, TERRY. "Thought and Desire in Plato," in Gregory Vlastos, ed. *Plato--A Collection of Critical Essays II: Ethics, Politics, Philosophy of Art and Religion*. Garden City, N.Y. Doubleday Anchor, 1971, pp. 96-118.
- PEARS, DAVID. "Freud, Sartre and Self-Deception." in Richard Wollheim, ed., 1974, pp. 97-112.
- REIFF, PHILIP. *Freud, The Mind of a Moralist*. Garden City, N.Y.: Doubleday Anchor, 1961.
- RORTY, AMELIE OKENSBERG. "Belief and Self-Deception." *Inquiry*. XV. 1972.
- _____. "Persons, Policies, and Bodies." *International Philosophical Quarterly*. XII, 1973a.
- _____. "The Transformation of Persons." *Philosophy*, XLVIII, 1973b.
- _____. "Adaptivity and Self-Knowledge." *Inquiry*, XVIII:1, 1975.
- _____. "A Literary Postscript." Amelie Okensberg Rorty, ed. *The Identities of Persons*. Berkeley: University of California Press, 1976.
- _____. ed., *The Identities of Persons*. Berkeley: University of California Press, 1976.
- SHAPIRO, DANIEL. "Self-Deception." (unpublished Ph.D. dissertation, Dept. of Philosophy, City University of New York).

- SZABADOS, BELA. "Rorty on Belief and Self-Deception."
Inquiry. XVII, 1974, pp. 464-473.
- THALBERG, IRVING. "Freud's Anatomies of the Self," in
Richard Wollheim, ed. Freud--A Collection of Critical
Essays. New York: Anchor Books, 1974.
- TRILLING, LIONEL. Sincerity and Authenticity. Cambridge:
Harvard University Press, 1971.
- WATSON, GARY. "Free Agency." Journal of Philosophy. LXXII:
8, 1975, pp. 205-220.
- WILLIAMS, B.A.O. "Deciding to Believe," in B.A.O. Williams,
Problems of the Self. Cambridge: Cambridge University
Press, 1973.
- WOLLHEIM, RICHARD. Sigmund Freud. New York: The Viking
Press, 1971.
- WOOD, DAVID. "Honesty." in Alan Montefiore, ed. Philosophy
and Personal Relations. Montreal: McGill-Queens
University Press, 1973.