MASTER COPY.

UNIVERSITY OF SOUTHAMPTON

REASON, MOTIVATION AND BEHAVIOUR: THE POSSIBILITY OF THE RATIONALITY OF UNCONSCIOUSLY MOTIVATED BEHAVIOUR

Les R.V. Burwood

TABLE OF CONTENTS

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ARTS

PHILOSOPHY

Doctor of Philosophy

REASON, MOTIVATION AND BEHAVIOUR: THE POSSIBILITY OF THE RATIONALITY OF UNCONSCIOUSLY MOTIVATED BEHAVIOUR

by Leslie Robert Victor Burwood

Historically there has been a link between unconscious motivation and Freudian theory. Freud's view was that unconscious mental states can be referred to in explaining certain types of behaviour. He believed that such behaviour was intrinsically irrational and was produced in a manner quite unlike the way in which ordinary conscious beliefs and desires bring about action. However, I argue that the concept of unconsciously motivated behaviour can be disentangled from Freudian theory and can be examined in its own right. Ultimately, it is shown that the concept of rationality is neutral between consciously and unconsciously motivated behaviour.

Much depends upon how the concepts of motivation and rationality are analyzed. Two main conceptions of motivation are identified, as are various conceptions of rationality, such as following one's interests, having reasons for one's actions and acting reasonably. Of the cluster of concepts examined only acting reasonably is shown to presuppose an awareness on the agent's part of some of his or her reasons for acting.

It is shown that among the various cases of unconsciously motivated behaviour, both that classified as partially and that classified as solely so, one can find all the differences which may be found in ordinary, consciously motivated behaviour. Sometimes so-called unconsciously motivated behaviour is, in fact, simply caused, and irrational; sometimes it is motivated, yet irrational; and sometimes, where the person is unconsciously motivated, the person may be said to act fully rationally.

NOTES FOR THE READER

I use single inverted commas to refer to words (rather than concepts), when a word is used in a non-standard sense, as a distancing or protesting device and for quotes within a quote. I use double inverted commas for direct quotation, for drawing attention to a concept and all other uses.

ACKNOWLEDGEMENTS

I would like to thank Guy Robinson, Tony Palmer, Jennifer Radden and Edward Roberts for their help. I would especially like to thank my friend Carol Brady for inspiration, for valuable suggestions, for typing the thesis and for reading through the typescript with me and helping me to express myself more clearly.

Les R.V. Burwood
August, 1981.

INTRODUCTION

I shall be concerned in this thesis with the concept of unconscious motivation and, in particular, with the question of the rationality of behaviour which is unconsciously motivated. This involves a cluster of philosophical questions: What is an intentional action? What is it to act for a reason? What is the basis of an evaluation of a person's action as 'rational' or 'reasonable'? The ultimate aim of the thesis is to show that the concept of rationality is neutral between consciously and unconsciously motivated behaviour.

It is my contention that the historical link between the concept of unconscious motivation and Freudian theory serves to sully and confuse subsequent discussion about rationality and unconsciously motivated behaviour. Confusion is especially created by the theoretical framework in the context of which Freud presented his hypothesis that unconscious mental states can be referred to in explaining behaviour and his own view of such behaviour as intrinsically irrational and as produced or caused in us in a manner quite unlike the way in which ordinary beliefs and desires occasion ordinary action. Moreover, it may be the case, as MacIntyre claims, that Freud shows us that more rather than fewer pieces of behaviour are rational.

For these reasons, I am proposing that when the question of the rationality of unconsciously motivated behaviour is considered, the latter concept can be disentangled from the web of additional Freudian theoretical hypotheses, such as those implicit in the theory of repression (in conjunction with which it is most usually and most naturally considered) and that it should be examined on its own. I shall argue that once concepts such as those of being reasonable or rational when one acts are analyzed in the light of descriptions of unconscious motivation which are free from psychoanalytic interpretation, then it can be shown that sometimes unconsciously motivated behaviour, both that classified as partially and that classified as solely so, is rational and, indeed, intentional in the fullest sense.

Freud's hypothesis of unconscious motivation intersects with and, thus, might be expected to throw light on and to be illuminated by traditional philosophical interests in two key ways. First, it involves the notion of intentional action and what it means to act for reasons or out of a motive. And, second, it can be related to the traditional philosophical debate on the nature of a rational action, where "rational" is used comparatively and evaluatively. The area of intersection mentioned and these two general conceptual discussions will be my concern in this thesis.

Despite his incalculable contribution to psychology, and less directly to philosophy through his explicit repudiation of Descartes's equation of mind with consciousness, Freud's approach to the evidence for the phenomenon of unconscious motivation was not, as I shall argue, one which could be germane to philosophical development in either of the two conceptual areas referred to above.

There are several reasons why this should have been so. First, Freud placed his discovery squarely in the context of the mechanistic, causal theoretical framework in which he had been disciplined as a physiologist, so that naturally he failed to see the full possibility of his claim that there are unconscious intentions and purposes. Freud took his observations almost exclusively from the behaviour of his neurotic and mentally disturbed patients and buttressed his explanations with the thesis that their irrational nature constitutes a defining characteristic of those mental processes which are unconscious. By so doing, he effectively excluded at the outset the possibility of evaluating unconsciously motivated behaviour in the light of its rationality. Such features of Freud's approach could only serve to mask the significance which the phenomenon of unconscious motivation held for philosophical discussions about reasons, intentional action and rationality.

A corollary limitation of the above ways in which Freud's conception of his discovery could have affected approaches to more purely philosophical questions is that real and conceptually important differences among the different ways in which unconscious

mental states can be shown to affect human behaviour would fail to be emphasized. Freud's interpretation of the concept of "unconscious motivation", in which seemingly irrational and illogical forces within us act upon our behaviour - as he thought - in the manner of simple efficient causes, conceals the differences (some of which are explored in the later chapters) which can be seen to exist between the different sorts of case he describes.

It is my plan in this thesis to investigate certain claims made in the context of the two philosophical areas mentioned above: (1) what is an intentional action and what it is to act for reasons or out of a motive, and (2) what it is to act rationally or reasonably. I then consider these areas in the light of the notion of unconsciously motivated behaviour. Partly because some of Freud's purely theoretical hypotheses are now in considerable disrepute in their own right and I do not desire the task of defending them, and partly because it may be thought, for the reasons I have just sketched, that Freud's own account of his discoveries comes as a rather mixed blessing, I shall attempt to do so clad as lightly in psychoanalytic 'metapsychology' as is possible. To the end of thus divesting myself, I shall throughout this thesis draw a sharp distinction between the purely empirical observations which Freud makes so acutely and describes so well and the theoretical conclusions which he draws from them. I shall use the former and attempt entirely to avoid the latter.

In the thesis I analyze notions such as reasons and motives, acting intentionally, and being reasonable or rational when one acts. I then examine these notions when characterized in the light of descriptions of unconsciously motivated behaviour which are free from any tinge of Freudian theoretical interpretation. I argue that it is certainly no longer obvious: <u>either</u> that all unconsciously motivated behaviour is irrational or unreasonable, <u>or</u> that unconsciously motivated behaviour always is simply caused or produced by unconscious mental states. However, I argue that the contrary account, sometimes suggested, is also false: that all unconsciously motivated behaviour fits neatly into the categories of action which is, respectively, rational or reasonable and motivated or intentional.

Rather, among the various kinds of unconsciously motivated behaviour emerge all the differences which we expect and find in ordinary, consciously motivated behaviour. We shall find that sometimes so-called unconsciously 'motivated' behaviour looks actually to be simply caused (verbal slips, for example, will be shown to be produced by something akin to mental causes) and is irrational; sometimes it is motivated or intentional, yet irrational. And, sometimes, although he or she may be unconsciously motivated in acting, a person may be said to act reasonably in the fullest sense.

These 'findings' suggest a general conclusion not unlike one which Freud himself drew in another context. On discovering that certain important mental functions like perception failed to be aligned with his distinction between Consciousness and Unconsciousness, instead of characterizing both 'parts' of the mind alike, Freud remarked that the characteristic of being conscious:

"......begins to lose significance for us."

(The Ego and the Id, Page 8)

In the same way, these possibilities which I will explore in my analysis indicate, with one exception, that the property or quality of consciousness may not be as significant for the concepts of "intentional action", "having reasons", and "acting rationally" as philosophers sometimes have supposed. Of the cluster of concepts analyzed, only "acting reasonably", it will be argued, presupposes an awareness on the agent's part of some of his or her reason(s) for acting. And, it will be shown that even so it may be possible to ascribe such awareness to an agent whose motivation is, technically speaking, 'unconscious'.

Before proceeding to offer some guide to the sequence of the arguments which I intend to develop, I must make one disclaimer. It will be noticed that nowhere in the following pages do I attempt to describe the conditions under which an explanation in terms of unconscious motivation might be regarded as a plausible one. This omission is deliberate. The whole question of the nature of the evidence supporting an explanation which posits unconscious mental states over competing explanatory hypotheses (like those in terms

of habit, for example) is not one which I can consider in this analysis. It may not even be a possible or useful exercise. For my purposes, it must be sufficient that there do seem to be cases, as I think there are, in which there is some prima facie plausibility for introducing explanations of behaviour in terms of unconscious motivation and unconscious mental states.

The flow of ideas and arguments in the thesis proceeds in the following way. In Chapter One, the technical Freudian use of the term 'unconscious' is outlined and it is contrasted with the non-technical uses we employ in everyday speech. The way in which the technical sense may be separated from its theoretical background and still distinguished from its non-technical, everyday counterpart is then explained. Chapter Two comprises an examination of the notion of unconsciously motivated behaviour. Freud's claims about unconscious motivation are considered and it is shown that among the various case studies of behaviour presented in psychoanalytic theory, only some can helpfully be described as motivated by unconscious mental states rather than being signs or symptoms of, or produced by, those states. In Chapter Three, the relationship between unconsciousness and the notion of having and acting upon a reason is examined with the intention of countering the claim that the idea of an unconscious reason or intention is one which is conceptually incoherent. Toulmin and Flew's significant work in extending and opening up Freud's interpretation of the notion of unconscious motives, intentions and purposes is described, and the suggestion is introduced that unconsciously motivated behaviour might be evaluated in terms of its rationality. Chapter Four provides an analysis of the notions of rationality and reasonableness with emphasis upon these concepts as they are introduced into evaluations of particular actions, and in Chapter Five the suggestion that unconsciously motivated behaviour might prove to be rational or reasonable is discussed in the light of that analysis. I argue that unconsciously motivated behaviour sometimes may be described as rational according to each of the different ways in which we commonly characterize that concept. Finally, I hope to discuss and dismiss the possibility that a

procedural policy advocating the wisdom of 'forgetting' may be drawn from my conclusions, and I hope to show that the only policy suggested by my arguments is one of striving towards self-knowledge and self-awareness, a policy which Freud himself urged so forcefully.

CHAPTER ONE

Summary

In order to undertake an investigation of the kind described in the Introduction, it is first necessary to defend the claim that a technical concept of unconsciousness would have any content left at all if we were to divest it of its theoretical background. It has been suggested, mistakenly, as I shall argue, that the technical concept must "stand or fall" (MacIntyre) with Freud's general theory. Further, that if this were so, then in the absence of its theoretical underpinnings the concept of an unconscious mental state would not differ significantly from our looser, everday use of 'unconscious' with its links with 'unwitting' and 'unnoticed'.

I begin by giving an account of the technical Freudian concept of unconsciousness, drawing attention to what must be for us its central feature - its implication that a person not only is unaware of a mental state which is "unconscious" but cannot even become aware of it (without analysis) - and showing the ways in which its explication is usually in terms of other aspects of psychoanalytic theory.

A comparison is then drawn between the Freudian concept and the everyday one described above, and it is argued that two features distinguish the non-Freudian concept: its inapplicability to discussions of unconscious reasons and motives, and, more importantly, its lack of implication for the person's ability or inability to become aware of states which are unconscious. I argue for the need for an account of the technical concept of unconsciousness which is at once free from additional theoretical hypotheses, and distinguishable from the everyday non-technical concepts with its implication that the unconsciousness in question is a merely contingent and alterable matter. The need for such an account is expressed in the requirement that the technical sense be adequately defined. A definition is given and some of the pitfalls involved in establishing such a definition

are enumerated and discussed. It is pointed out that inaccessibility to awareness cannot alone count as the defining characteristic of states which are unconscious in the technical sense, because not only are we said to become unconscious of ideas once conscious, so too, through the process of psychoanalysis, we are described as becoming conscious of unconscious ideas. It is shown that sometimes unconsciously motivated behaviour is defined by using the aspect of hindsight usually associated with post-therapeutic revelations of unconscious states, whereby unconscious states are those which we cannot be aware of at the time at which they are influencing our behaviour. Despite this use of hindsight, the plausibility of also describing cases of people who continue to act upon what seem to be unconscious motives even after having become aware of them, is postulated and subsequently established through argument based around and drawing on some of Freud's case studies. The nature of this awareness is, however, somewhat problematic.

Certain objections which might be directed at the examples I use are then discussed. The charge that descriptions of cases such as those used embody the paradoxical claim that some mental state is both conscious and unconscious at the same time, is shown to be dependent upon the mistaken assumption that "unconscious" and "conscious" are simple contraries, and "unconscious of" and "conscious of" univocal expressions. I argue that "being conscious of" signifies two different relations in which we stand towards our mental states: being aware of them, and being aware of them in a particular, distinguishable way. Thus, in the cases I describe, the person might still be said to have or be in the unconscious mental state affecting his or her continuing behaviour, even while being aware that that state may be ascribed to him or her and is, in that sense, conscious of it. The nature of the particular qualitative difference between awareness of ordinary mental states and awareness of unconscious mental states is then discussed. It is argued that the notion of "knowledge without observation" is unhelpful as a way of characterizing the former kind of awareness, and Freud's characterization of the latter in terms of the absence of an accompanying "feeling of familiarity" is discussed.

The difficulties of characterizing that kind of awareness which is an awareness of unconscious motivation is explained.

The notion of the agent accepting an explanation of his or her behaviour in these terms and yet the belief being strangely alien is elucidated. It is contrasted with both the more straightforward acceptance of an analyst's hypothesis, and with our ordinary conscious ways of knowing our mental states. It is then pointed out that, in any case, as long as the capacity unerringly to make the distinctions mentioned is present, no further detailed characterization is necessary of the qualitative differences between the kinds of awareness involved. Finally, a comparison is made between the technical concept of unconsciousness as I have defined it, and the allied one of "self deception". It is shown that the concept I am concerned with is, so to speak, a static one, and is for that reason to be distinguished from that of self deception, which primarily connotes an activity, rather than a mental state.

* * * * * * * * * * *

How are unconscious mental states to be defined? How do beliefs, wishes and feelings which are unconscious differ from ordinary conscious beliefs, wishes and feelings? I mean to investigate these questions by looking first at the way in which the expressions 'unconscious of' and 'unconsciously' are used as technical terms in the context of Freudian theory. It will then be possible to compare the latter use with the everyday uses of those expressions encountered in plain, non-technical discourse. In what follows I refer to "Freudian" (technical) concepts and contrast these with "Non-Freudian" (non-technical) concepts.

The Freudian Concept of the Unconscious

Freud's clearest theoretical statements concerning the concept of an unconscious mind are to be found in his so-called "meta-psychological papers"[1], in particular, in the two great 1915 essays, "Repression" and "The Unconscious", as well as in The Ego and the Id (1923).[2] At the time of writing the latter work, Freud understood his new system of classification into Ego, Id, and Superego processes to have transcended and replaced the earlier one into Consciousness, Preconsciousness and Unconsciousness. We find him writing, as was remarked earlier, that the characteristic of being conscious

> ".....begins to lose significance for us."
>
> (The Ego and the Id, page 8)

and in Freud's and the neo-Freudians' later advance into what has come to be called ego psychology, the distinction between the consciousness and unconsciousness of mental states received scant attention. Increasingly, Freud turned from an exploration of the mind in terms of the qualitative aspects of mental items or ideas, to an analysis in terms of different mental functions, such as perception and defence. The distinction between "conscious" and "unconscious" generally was not conceived in functional terms. Nor did the functional distinctions which Freud found useful map onto that between consciousness and unconsciousness in any way. Important ego functions, for example, were found to take place both at a conscious and at an unconscious level, to be characterized by the property of consciousness in one case and not in another.

However, it must be stressed that Freud's move into ego psychology represents merely an alteration in direction and emphasis. The distinction between Conscious, Preconscious and Unconscious remains a viable and fundamental one within the body of psychoanalytic theory despite the latter concern with a more functional analysis of the mind. The property of being conscious or not, as Freud himself remarks in the same passage from which the previous quotation derives:

> ".....is in the last resort our one beacon-light in the darkness of depth-psychology."
>
> (The Ego and the Id)

What, then, does Freud mean by the Unconscious? Briefly, his theory is that there are certain "instinctual impulses" or wishes, which are either (a) represented by ideas, or (b) manifested in affective states or feelings, which are themselves best seen as ideas since they are identified qua unconscious items, by their "ideational representation", or object[3]. These are somehow active within us in the sense of influencing our behaviour and conscious mental states, while at the same time, because of a particular resistance on our part to their content, they are not available to our conscious minds. These ideas are said to be unconscious; and just as we speak indifferently of an ordinary conscious idea as being conscious or as being an object of consciousness, so Freud speaks at one time of an idea's being unconscious and at another of its being such that we are unconscious of it. (In addition, Freud resorts to a somewhat regrettable spatial metaphor: we find him referring to these ideas as residing "in" the Unconscious and speaking of "the Unconscious" as if describing a physical place.) Not merely are we not aware of these ideas; we are unable to become aware of them, according to psychoanalytic theory.

Freud was prompted first to hypothesise such unconscious mental states during his studies of hypnosis and the condition then described as "hysteria" - studies which began in 1885 and culminated in 1895 with the publication, in collaboration with Breuer, of Studies in Hysteria[4]. By hypothesising that memories and ideas of which the patient was herself unaware remained active within her mind and produced her hysterical symptoms - just as, in

the condition known as "post hynotic suggestion", ideas introduced under hynosis and unknown to his or her conscious mind had been shown to affect the subject's subsequent behaviour - Freud found himself able to explain the peculiar features of this condition. Hysterical patients, Freud wrote:

> "suffer from reminiscences. Their symptoms are residues and mnemonic symbols of particular (traumatic) experiences."
>
> ("On Hysterical Mechanisms"[5], page 29)

The contrast drawn between not merely being unaware of ideas which are "unconscious", but being unable to become aware of them, is emphasized in the distinction which Freud makes between the "descriptive" sense in which mental life is unconscious, on the one hand, and the "systematic", or what later came to be called the "dynamic" sense, on the other, a distinction which is reflected in the terminology of 'Conscious', 'Preconscious' and 'Unconscious':

> ".....a psychical element (for instance, an idea) is not as a rule conscious for a protracted length of time. On the contrary, a state of consciousness is characteristically very transitory; an idea that is conscious now is no longer so a moment later, although it can become so again under certain conditions that are easily brought about. In the interval, the idea was - we do not know what. We can say that it was latent, and by this we mean that it was capable of becoming conscious at any time. Or, if we say that it was unconscious, we shall also be giving a correct description of it. Here 'unconscious' coincides with 'latent and capable of becoming conscious'....."
>
> and
>
> ".....The latent, which is unconscious only descriptively, not in the dynamic sense, we call preconscious; we restrict the term unconscious to the dynamically unconscious repressed; so that now we have three terms, conscious, preconscious and unconscious.....We can now play comfortably with our three terms, conscious, preconscious and unconscious so long as we do not forget that in the descriptive sense there are two kinds of unconscious, but in the dynamic sense only one."
>
> (The Ego and the Id, pages 4-5)

Thus, speaking descriptively, "unconscious" can be used to cover anything which is not an item presently in consciousness, anything of which we are not presently aware, including those things which are merely latent, or preconscious; while speaking dynamically, "unconscious" covers only those items of which we cannot become aware, the "repressed".

The process of "censorship" by which ideas are both expelled

and withheld, from consciousness, is described as "repression", and the ideas which have been inhibited thus are described as repressed material, or "the repressed". The mechanism by which repression is understood to work, and the theory underlying which items are selected for repression need not concern us in any detail here. But, briefly, Freud's idea is that there is a quantum of psychic or mental "energy" which accounts for the relative strength of various impulses, or ideas representing impulses. An idea representing an impulse is "strong" when it is highly "energized" or "cathected", or "invested", as Freud sometimes puts it, with a high degree of "libido" or "interest". The psyche is regulated by the so-called "pleasure principle", by which an organism functions so as to maximize pleasure and avoid pain and anxiety. Although the satisfaction of impulses is always pleasant, Freud hypothesises that when the satisfaction or awareness of such impulses is incompatible with others of the person's claims and purposes, and when the pain and anxiety caused by that conflict outweighs the prospective pleasure which the satisfaction of the impulse might be expected to bring, then:

> ".....the element of avoiding 'pain' shall have acquired more strength than the pleasure of gratification."
>
> ("Repression", Page 105)

and

> ".....as soon as an idea which is fundamentally offensive exceeds a certain degree of strength, the conflict takes on actuality, and it is precisely the activation of the idea which leads to its repression."
>
> ("Repression", Page 109)

So, by becoming unconscious of them, a person avoids those of his or her impulses the awareness of which would be more painful than pleasant.

Thus, "the repressed" might be supposed to be, and as I shall argue, is, roughly equivalent to "the unconscious". However, without offering explicit reasons why this should be so, Freud insisted that the class of mental items which are unconscious was broader than that covering "the repressed".[6]

The significance of the latter claim actually is fairly

slight; it rests on certain features of technical definitions arising in two different aspects of psychoanalytic theory. In the first place, Freud's view is that there are two different kinds of unconscious material: that which has never reached consciousness and that which has been conscious and has been expelled from consciousness. Thus, the "primal" repression of ideas which have never entered consciousness is distinguished from "repression proper", or what Freud sometimes refers to as "after expulsion": the repression of ideas which have once been conscious. Clearly, it may be said that only in the case of the latter phenomenon, strictly speaking, has the material actually been repressed. In addition, the distinction which Freud draws between descriptive and dynamic sense of "unconscious", the distinction underlying his introduction of 'conscious', 'preconscious' and 'unconscious' terminology, would allow that at least when we are speaking descriptively, there is a class of mental states which, while being "unconscious", are not "repressed", that is, those which are merely latent in the mind, or "preconscious". Thus, only in a narrow sense is it true that the two classes, "the unconscious" and "the repressed" are not completely co-extensive. And, as long as we treat Freud's primal repression as repression, and make use of the distinction between "unconscious" and "preconscious", then "the repressed" exactly corresponds to "the unconscious".

The evidence for claims about the unconscious is similarly linked with the theory of repression, in particular with the thesis of "the return of the repressed".[7] The latter thesis proposes that what is repressed continually pushes towards consciousness and sometimes breaks through and shows itself, somewhat disguised, in what are described as "substitute formations" and "symptoms" - in dreams and fantasies, slips, mistakes and omissions of memory, in humour and neurotic symptoms, all of which are quite as inexplicable to the person in whom they occur as they are to an observer. It was in order to explain such haphazard and apparently inexplicable phenomena as these, in addition to explaining the evidence indicated by hysteria and hypnosis, that Freud introduced the hypothesis that there must be unconscious forces at work.

Although he takes pains to emphasize the similarities between unconscious and conscious mental states and processes in order to justify describing the unconscious states as "mental" or "psychological" at all[8], Freud also insists that there are certain qualitative differences distinguishing unconscious ideas and thought processes from ordinary conscious ideas and thought processes. Freud's concern here is with differences which may be summed up in his distinction betwee "primary process" and "secondary process" modes of thought and mental functioning.

Primary process thought is contrasted with secondary process thought; the latter is found in normal, logical adult thinking. Primary process thinking, at least in the version of the theory which is of interest to us, characterizes all unconscious thought processes.[9] It is described as non-logical (it embraces contradictions), given to "displacement" (that is, the substitution of one word or idea for another which in some respect(s) resembles it), and to "condensation" (that is, the compression which enables one word or idea to stand for and symbolize several different ideas which it in some respect resembles), unaltered by the passage of time or by changes through time, unimpeded by beliefs about the external world at all, untroubled by "negation, dubiety..... varying degrees of certainty.....",[10] characterized by a "mobility of cathexis", a tendency to make the various substitutions described (the activities of displacement and condensation) and to reorient "psychic energy", or feeling, appropriately,[11] and finally, pictorial and concrete rather than verbal and abstract.

In addition to drawing the distinction between primary and secondary process thought, Freud elaborates, in the later part of his essay on the unconscious and in The Ego and the Id, upon the distinction between the pictorial and concrete nature of unconscious ideas and the verbal and abstract nature of conscious (and preconscious) ones, suggesting that an idea's entering preconsciousness from the unconscious consists in its changing from being a mere picture to being a picture together with a "word association". His theory is that we can distinguish "the idea of the word" (a verbal or sound memory image) from "the idea of the thing" (the

concrete picture, or visual image). A conscious and an unconscious idea:

> ".....are not different records of the same content situated in different parts of the mind, nor yet different functional states of cathexis in the same part; but the conscious idea comprises the concrete idea plus the verbal idea corresponding to it, whilst the unconscious idea is that of the thing alone."
>
> ("The Unconscious", Page 147)

The visual thinking which characterizes unconscious processes, according to Freud, is incapable of conveying the relations between various idea elements; it is static and concrete. Only when the verbal ideas are added to their corresponding pictorial images, as occurs when repressed ideas reach preconsciousness, do the "intermediate links" between the various elements of the subject matter emerge. Only then do the ideas take on the quality of real "thoughts".[12]

I have now given a brief account of the psychoanalytic notion of unconsciousness. Let us compare this with the ordinary non-Freudian use of the terms 'unconscious of' and 'unconsciously' as they are found in everyday non-technical discourse.

The Non-Freudian Concept of Unconsciousness

Apart from the adjectival use of 'unconscious' describing the insentient and unthinking state of someone suffering the effect of a blow on the head, there is a standard current use of 'unconscious of' and 'unconsciously' which in some ways resembles and in other ways notably differs from the Freudian one with which we have been concerned until now. In what I shall call the "non-Freudian" use, we speak of people being unconscious of:

(i) their mental and bodily states; for example, "Unconscious of her own grief and fatigue she battled on".

(ii) what happens to them; for example, "Unconscious of the snub, she replied politely".

(iii) what they do; for example, "Unconsciously she reverted to the language of her childhood", and "Unconsciously she let the book fall from her hands".

(iv) certain aspects or ways of describing what they do; for example, "She settled herself into the chair and in doing so unconsciously she echoed the posture of the woman in the painting".

The distinction between (iii) and (iv) perhaps requires elucidation. Assuming that "settling herself into the chair" and "echoing the posture of the woman in the painting" count as doing one thing rather than two, it may be said that "echoing the posture of the woman in the painting" is one way of describing what the person referred to did, that is, sitting down. The contrast to which I wish to draw attention through this example and the ones described in (iii) is that between:
(a) carrying out some action or doing something while being totally unconscious of having done so, ("The girl is totally unaware that the book has fallen"), and
(b) carrying out some action or doing something when only some or one of all the possible descriptions pertaining to that which is done elude the person in question.
Although not unaware that she has seated herself, the girl in (iv) is ignorant of having done so in such a way as to echo the posture of the woman in the painting. In the sense of "unconscious" under discussion, not only is it true that some descriptions under which we can place the things we do may be unknown to us, so too may the fact that we have done anything at all.

There are two ways of distinguishing this Non-Freudian sense from the Freudian sense of 'unconscious' and 'unconsciously'. First, "unconscious reasons" and "unconscious motives" are strictly post-Freudian concepts. So note that being "unconscious of", or not knowing what one is doing, in the two different ways described in (iii) and (iv) above is to be distinguished from being unconscious of or not knowing why it is that one does what one does. In the non-technical, Non-Freudian usage under discussion we do not speak of being unconscious of the reasons or motives for our actions. And this is the significance of the qualification in (iv). Although it is possible to distinguish beliefs about what one does from those about one's reasons or motives for what one does, sometimes a description of what one does is cast in terms of the agent's goals and ends, and thus it reveals the reason(s) or motive(s) underlying the action. ("Writing a thank-you letter", for example, is a way of describing what a person does which

reveals at least one of that person's reasons or motives in doing it.) Thus, the exception in (iv) which stated that we can be "unconscious" of only certain of the various aspects or ways of describing what we do is that of descriptions in terms of motives and reasons. These we cannot be said to be unconscious of in the Non-Freudian sense. "Unconscious of what she was doing, she finished the thank-you letter" is possible, of course, but it would mean that the whole task was undertaken unconsciously if "unconsciously" were used in the latter way - mechanically, distractedly or without close attention. It would merely be a case of unconsciousness of something done, in the sense of (iii). "Unconscious that what she wrote was a thank-you letter, she formed her words with care" makes little sense as long as the Non-Freudian sense of "unconscious of" and "unconsciously" is intended. Such a description could only imply the technical, Freudian use of the expression.

So, in the Freudian sense of "unconscious" we can speak of reasons and motives as unconscious; in the non-technical sense we can not. The second way in which the Freudian and Non-Freudian uses of "unconscious of" and "unconsciously" can be distinguished is as follows. Although there are nuances of difference between the notion of being unconscious of something, in the Non-Freudian sense, and failing to notice or to attend to that thing, nevertheless there are very strong links between the Non-Freudian sense of unconsciousness and the idea of a simple failure of attention.* In some cases, it is precisely a failure of attention which can be seen to have resulted in the unconsciousness in question. This is clearly the case when we say, for example: "The excitement of the

---

*In connection with the above-mentioned nuances of difference between being unconscious of some object in the sense under discussion and failing to notice or attend to that object, it is interesting to note that while in this sense "being unconscious of" often means something very close to "not noticing", we are said to be unconscious of by no means all that we fail to notice. The concept "unconscious of" seems to be restricted to roughly those things which we might have been expected to have been conscious of or to have noticed. To say that a person was unconscious of something is to suggest that there is something unusual or peculiar in the thing's having failed to attract the person's attention. When it is natural that the object should not have received attention, then we should not naturally speak of being unconscious of it.

occasion so absorbed her attention that she remained unconscious of her fatigue", or "In her eagerness to exchange reminiscences, she failed to notice that she had unconsciously reverted to the language of her childhood". In other cases, it is at least true that no more than an act of attention is required to remedy or dispel the state of unconsciousness.

The Non-Freudian sense of 'unconscious' contrasts with the Freudian sense in this capacity to become aware of the previously unconscious state. In the Non-Freudian use of "unconsciously" and "unconscious of" there is no suggestion as to the person's capacity or incapacity to make conscious or attend to his or her hitherto unconscious state. In none of the examples given earlier illustrating this use is the implication that any barrier might be placed in the way of exercising the act of attention or introspection required to make the state in question an item of present consciousness. So, one aspect of the Freudian concept of unconsciousness which distinguishes it from the Non-Freudian use of "unconscious of" and "unconsciously" is its emphasis on the person's inability to become aware of the states in question. The distinction may be put by saying that in the case of Non-Freudian unconsciousness, it is only contingently the case that we are unconscious, while their being unconscious is a necessary feature of Freudian unconscious mental states.

Summary of Differences Between the Freudian and Non-Freudian Concepts of Unconsciousness

We can now sum up the differences between the Freudian concept of unconsciousness described in the previous section and the non-Freudian concept just discussed. The Freudian use of "unconscious of" and "unconsciously" differs from the Non-Freudian one:

1. in implying that our unconsciousness is a necessary and not a merely contingent feature of our mental state, and
2. insofar as they take different objects. As well as taking all the objects described in (i)-(iv), the Freudian sense of "unconscious of" and "unconsciously" extends, as we have seen, to cover reasons and motives.

In the Freudian sense, we may be unconscious of our reasons and motives in two different ways. First, we may be unconscious both of having acted in a certain way and of the reason why we so acted or the motive for our action. And, second, we may be unconscious of the reason or motive for the action although aware of having acted in a certain way. Neither description is possible if we are merely speaking in the Non-Freudian sense of "unconscious of" and "unconsciously".

The Need for a Definition to Distinguish the Technical from the Non-Technical Concept

It may be supposed that were we to rid the Freudian concept of unconsciousness of its theoretical underpinnings, having an unconscious mental state would come to nothing more than having a mental state of which one was unaware. That is to say: all the so-called Freudian uses of "unconscious of" and "unconsciously" actually would reduce to what we have been calling Non-Freudian uses, and there would be no such thing as being necessarily unconscious of some state. Clearly, if this were the case, then distinguished from the body of psychoanalytic theory the concept of unconsciousness would have very little philosophical interest. Yet, many philosophers of psychology apparently have assumed that it is possible to give a definition of "unconscious of" and "unconsciously" so as to distinguish the Freudian from the Non-Freudian use. Or, more often perhaps, it has been assumed that the difference between the two classes of phenomena is sufficiently obvious to obviate the necessity of a formal definition at all, other than the cursory and patently inadequate kind which states simply that unconscious mental states are those of our mental states of which we are not conscious.[13]

Now Freud, as we have seen, presupposes the validity of at least two fundamental cornerstones of psychoanalytic theory in introducing the concept of the unconscious. First, he uses the theory of repression to describe the origins, development and dynamics of unconscious ideas and processes; and, secondly, he refers to the distinction between primary process and secondary

process thinking to characterize the qualitative difference between consciousness and unconsciousness. Equipped with these additional theoretical tools, it is certainly possible for Freud adequately to distinguish his own from the Non-Freudian use of "unconscious of" and "unconsciously", and to offer a watertight definition of his own use. Indeed, as we have seen, he may do so in a number of ways. He may define unconsciousness in terms of the theory of repression, in terms of the theory of primary process thinking, or in terms of his theory about the presence and absence of the "verbal idea" or "word association".

It has been argued, for example by Alasdair MacIntyre,[14] that the dependent relationship described which obtains between Freud's concept of the unconscious and his other theoretical concepts is not one which it would ever be possible to sever. If this were so, then the usefulness and the very life of the Freudian concept of unconsciousness would rest squarely upon the conceptual and theoretical viability and validity of psychoanalytic theory as a whole - the concept of the unconscious would "stand or fall" (MacIntyre) with the general theory.

In opposition to the above view, however, I wish to consider the possibility that the notion of a state of which a person is necessarily unconscious in the sense described can be detached from any other aspects of Freudian theory and yet can be distinguished adequately from the Non-Freudian sense of "unconscious of" and "unconsciously". To this end, I shall attempt in the second half of this chapter to give a formal characterization of what ought perhaps now to be called the "technical" rather than the "Freudian" concept of unconsciousness - a characterization which is independent of other Freudian theoretical concepts in the way described.

Apart from the purely theoretical challenge which encourages such an enterprise, several considerations, some of which were raised in the introduction to this thesis, have influenced my

decision to approach the concept of unconsciousness in this way. The first is that there is some evidence which appears to cast into doubt the theoretical and conceptual validity of certain other of the constituent concepts making up the general theory of which the concept of the unconscious is a part.[15] To claim this is not to make the absurd claim that, say, the concept of repression could be given up easily. Nevertheless, it is to say that it would be wiser not to let the validity of the technical sense of the unconscious depend on relatively controversial Freudian theories, concepts and empirical claims. Thus, to dramatise our endeavour, we might speak of 'saving' the concept of unconsciousness. In addition, there would appear to be some justification in exploring a concept which, while beginning as a purely technical and theoretical one, has found a place in ordinary discourse. I wish to test the validity of using this concept in the absence of the theoretical underpinnings which first spawned it. The latter point is linked with the following one. The notion of a mental state which is correctly ascribable to a person, while an ascription of the capacity for awareness of the same state is at the same time witheld, is one which alone would appear to have the utmost significance for and application to several of the most pressing problems which emerge in the philosophy of mind. For this reason, it should be hoped that the concept could be helpfully introduced into discussions of philosophical problems uncluttered by any further theoretical baggage.

In the section which follows, I shall explore some of the difficulties which arise when we attempt to follow such a procedure and to give a definition of the technical concept of unconsciousness when it is abstracted from the additional theoretical hypotheses in terms of which it is presented to us by Freud. It is important to notice that unless a definition can be found which is adequate to distinguish the technical from the non-technical use of "unconscious of" and "unconsciously" - as I think that we do eventually succeed in doing - then the concept would not hold so great a philosophical interest.

Before proceeding to the next section, I must make one further remark about the general method and approach employed in the following pages and in all conceptual analyses throughout this thesis. While, for reasons which I hope are now apparent, I intend that all subsequent discussion should be free of reference to psychoanalytic theoretical concepts, I shall nevertheless continue to cite Freudian writing and appeal to descriptions of cases offered by Freud when making reference to empirical facts about the evidence for unconscious mental states and unconscious motivation.

It is, of course, possible to distinguish Freud's purely empirical claims from his theoretical analyses of those claims. And psychoanalytic theory offers not only the classic statement of the case for the existence of unconscious mental states and unconscious motivation, but also the most thorough empirical description of these phenomena. It goes without saying that the thesis owes so much to Freud's work. The substantive work of Freud and his followers provides the material about which I raise and try to answer some important philosophical questions.

Unconscious Mental States: The Problem of Providing a Definition

What is clearly for us the significant aspect of the technical concept of unconsciousness is that, as we have seen, it involves not only a lack of awareness of those states on the part of the person whose states are said to be unconscious, but an inability to become aware of them. We may treat their being inaccessible to awareness or consciousness as the defining characteristic of unconscious mental states.

But, the latter way of characterizing unconscious mental states in terms of their inaccessibility to awareness is, as it stands, insufficient for our purposes. For, not only does it seem possible that the same state, for example the belief that P, which is now unconscious, might have been a conscious state at some earlier time (that is, at time t_1 I might have believed P consciously, while at time t_2 I had come to believe P unconsciously),

but, in addition, the very process of cure in psychoanalysis, and in much psychotherapy as well, is said to enable us to become aware of our previously unconscious mental states. Yet, if its inaccessibility to the person's awareness is the defining characteristic of an unconscious mental state, how can this be possible?

There are a number of ways in which the definition of unconscious mental states may be modified to deal with the difficulty mentioned. It has been proposed by Alisdair MacIntyre[16], for instance, that since the therapeutic recovery of unconscious mental states employs methods quite unlike the effortless calling up of ordinary mental states which are merely latent in the mind (or preconscious, to use Freud's terminology), and the reflection which precedes the ordinary giving of reasons for our beliefs and actions, the definition might better read

> (A) Unconscious mental states are those of our mental states of which we are unaware and of which we cannot become aware without the special aid which therapy provides.

The difficulty with the latter qualification is that it fails to account for the possibility of the practice of self-analysis. It leaves too loose the specification of the nature of the aid which therapy provides. But it must do so just because the nature of the aid is only loosely definable. Various activities are said to be able to disclose unconscious motives, beliefs and desires, including self-analysis. We can imagine a person time t_2 discovering unaided the hitherto unconscious reason for the action which he or she undertook at time t_1. Indeed, Freud explicitly admits of this phenomenon.[17] So, this particular qualification to our definition is an unhelpful one.

Instead, it may be suggested that although it is possible for a person to become aware of his or her unconscious mental states, it is not possible for that person to become aware of them at the same time that they are influencing the conscious mental states or behaviour which they are introduced to explain. Our coming to acknowledge our unconscious mental states usually is pictured as a form of hindsight. At some time after the behaviour

which the unconscious mental states motivated or in some way produced, or the occurrence of the conscious mental state which they occasioned, we become aware of them as the explanation of that behaviour or state. And it may be supposed that this temporal qualification applies unfailingly to situations in which such awareness of unconscious mental states occurs. If this were so, we might modify the definition to read

> (B) Unconscious mental states are those of our mental states of which we are unaware and of which we cannot become aware, at least at the time at which the behaviour which they influence and which they are introduced to explain, is taking place.

Yet, there would appear to be a weakness with this formulation of the definition also. For, prima facie, there is no particular conceptual difficulty with the notion of a person somehow learning of his or her unconscious motive or reason for doing X (or for being in state Y) through therapy, self-analysis or simply through having been told it, and thus no longer being unaware of it, and yet continuing to do X (or to be in state Y) and continuing to act upon (or behave or be in some state out of) his or her "unconscious" state, motive or reason.

Plausible cases of a person doing something out of, or being affected by, an unconscious mental state of which that person is at the time conscious are not difficult to frame.

Case 1: Let us consider a simple case of straightforward action. A man may experience nothing more than a conscious wish to avoid stepping on the cracks in the pavement, and yet may hypothesise that there is some unconscious motive underlying this wish, and may discover it. Perhaps it is an unconscious fear of the cracks based upon some symbolic identification between them and another object, for example, menacing female genitalia. There are actually two distinct versions of the case as it has been described so far:

(i) On the one hand, the person might experience no conscious inclination concerning the cracks in the pavement at all. If his avoidance of them were drawn to his attention before the light

of insight dawned, he would deny the existence of any intention in this behaviour and charge his actions to coincidence or habit.
(ii) On the other hand, he may (consciously) either have known that he liked to avoid the cracks but believed (falsely) that his crack skipping reflected a whimsical identification with Christopher Robin, or have believed it to be a mere fancy, without reason or significance of any kind. Whichever version of the case is presented, the important point is that after his recognition of the likely connection between his crack skipping and his sexual fears, it seems possible to describe the man as continuing to be affected by his unconscious mental states while walking on pavements and continuing to avoid the cracks.

Case 2: Suppose that after a session with her analyst a woman hypothesises that her tendency to be attracted to, and to develop relationships with older men, reflects her Oedipal attraction to her own father, an attraction which, at the level of conscious experience, she does not feel. Believing this to be a likely explanation for her behaviour, she nevertheless continues to exhibit the same tendency and to act upon it. Again, corresponding to (i) and (ii) above, there are two slightly different situations which might fit this general description:
(i) The woman may be aware of a consciously experienced preference for older men for which, until her moment of insight, she could give no reason, or for which she could give no reason which actually had moved her (she might falsely have believed their savoir-faire to be the source of their appeal), or
(ii) Both preference and reason for it may have been unconscious. She may have found features other than those related to their maturity to be the attractive aspects of her men friends, so that her preference for older men would itself have had to be discovered by observation of her past behaviour, rather than by an introspective appeal to her own tastes.

The latter case is complicated because "being attracted to older men" is a catch-all phrase covering a whole behaviour pattern, which comprises some behavioural tendencies and some affections as well as standard cases of actions. Feelings of

attraction, for instance, are primarily passions, whose onset, quality and duration are not factors over which we seem to exercise much, if any, control. We must restrict ourselves to some particular action which makes up part of the behaviour pattern of being attracted to older men, for example the action of accepting flowers and invitations from them. Then we can say that our woman continues to act thus (that is, to accept flowers and invitations) after her moment of insight, and continues to be unconsciously motivated in doing so, even while she is aware of the unconscious motive or state moving her. If, on the other hand, we choose another aspect of the behaviour pattern and deal with the apparently passive state of wanting to accept flowers and invitations from older men, we may say that after her insight our woman continues to be affected in this way, even as she acknowledges to herself the reason why she is so affected.

Cases 1 and 2 have, as I have said, a prima facie plausibility. On the face of things, it would seem incorrect to suppose that merely knowing of the unconscious mental states moving him or her would make it impossible for a person to (continue to) act upon, or be affected by those states. Yet, these very examples might be thought to invite certain objections nonetheless.

Before examining the kinds of objections mentioned and the ways in which they can be countered, one further observation must be made about the nature of the examples themselves. Notice that in each of the cases described, the unconscious state or motive under discussion is something more than a mere wish or desire taking as its object a description of the action undertaken, or state experienced. Thus, although it seems possible to describe a case in which a person does X even while being aware that the "unconscious" reason which explains his or her action is some further belief P (for example, the belief that the cracks represent female genitalia) and desire Q (to avoid such menacing objects), it is not possible to describe a case in which a person does X while aware that the "unconscious" reason which explains his or her action is the wish to desire to do X. The wish to do X can only be introduced once into an explanation of a particular

action X. And, because a person's doing X, assuming that the action is voluntary and intentional, entails that he or she may be said in the ordinary conscious way to want or desire to do X, then, merely to introduce an unconscious wish to do X is not in any way to add to the explanation of why the person did X. So, only cases of the kind described in Cases 1 and 2, where the unconscious state of which the agent is said to become aware is some additional reason or motive over and above the simple wish to do the action in question, are ones which it is possible to present in this context.

Let us now turn to the kind of objections that might be raised against examples such as those described. The first such objection precludes the possibility of that kind of case on purely a priori grounds. The objection rests upon the temptation to argue, as some people discussing unconscious mental states as motives have done,[18] that our unconscious beliefs and desires are so "totally opposed" and contrary to our consciously held beliefs and desires that, were they to become conscious, we would never espouse them or act upon them. For example, the desire to commit incest might be one such belief.

Two different formulations of this kind of objection may be distinguished. The above formulation, which is the stronger one, states that for conceptual reasons internal to psychoanalytic theory it cannot be possible to find cases of the sort which we have sketched. A weaker form of the objection would be that while it is not conceptually impossible, as a matter of fact we never find cases such as the ones described.

However, in my view, neither version of the objection has much to recommend it. The stronger formulation seems to reflect nothing more than a slavish (and possibly mistaken) adherence to Freudian hypotheses about the causes of repression, viz., that only those ideas and wishes which continuously are highly painful and anxiety provoking to the agent ever are unconscious - Freudian hypotheses to which I am explicitly remaining uncommitted. Moreover, the factual assertion underlying the weaker formulation of the objection is cast into doubt even by certain passages in

Freud's own writing, in which we find hints that unconscious wishes and ideas sometimes might come to be accepted, and not merely acknowledged, by the agent's conscious self.[19] And the very possibility of describing cases of the kind we have seems to cast such an hypothesis further into doubt.

And yet an air of paradox which pervades certain ways of describing our cases deserves further comment. If we describe either case as one in which a person is acting upon or being moved by an unconscious motive or state while at the same time being conscious of that same motive or state, our description appears contradictory. If the agent is conscious of the motive or state, it may be supposed, then the motive or state is no longer unconscious at all. If he or she continues to act upon, or to be affected by such a motive or state when it becomes conscious, then it becomes a conscious motive for his or her action or the conscious state affecting him or her, and the person is now consciously motivated or affected rather than unconsciously so. Once the agent can be said to be conscious of the motive or state then he or she cannot be said to be unconscious of it.

This apparent paradox notwithstanding, our cases seem, as we have seen, to be perfectly plausible. The paradox, as I shall now try to show, is more apparent than real. It rests upon the mistaken supposition that "being conscious of" a mental state, and "being unconscious of" a mental state, are contradictories. In fact, as we shall see, the expression 'being conscious of' is not univocal at all. There are two distinguishable interpretations of the expression and only the unexceptional one, which is compatible with "being unconscious of", is implied in the cases described.

My cases do seem plausible and this, I would suggest, is because it would always be possible intuitively to distinguish - within those of our beliefs, desires, motives and reasons of which we are aware and which we ascribe to ourselves - between our ordinary conscious states of which we had always been aware, on the one hand, and those of our unconscious states of which we had

merely become aware, on the other. Thus, the crack skipping man in Case 1 would have been able introspectively to distinguish such an ordinary conscious reason for behaving as the one attributed to him of wishing to play at being Christopher Robin (in order, perhaps, to amuse a small child accompanying him), from an unconscious motive or reason of which he happened to become aware, such as the unconscious identification of the cracks with female genitalia. The quality of his awareness or consciousness of these two different sorts of reason or motive would strike him as different in each case. That is, he is not aware of it as a motive or as his motive; he is not aware of the fear that he might be swallowed back into the womb. He is aware, however, of the unconscious motive explanation.

Nor, in this case of the crack skipping man, would this capacity to distinguish between the two different reasons or motives for his behaviour be based on the peculiar, incoherent nature of the unconscious beliefs and desires attributed to him, the intrinsic oddness of the idea of such an identification and such a fear. To turn to my other example, had the woman in Case 2 been moved by an ordinary, conscious lustful identification of her ageing suitors with her own father, with whom, let us say, she had quite recently had an actual incestuous relationship, then her awareness of it still would be distinguishable from the awareness she had after analysis of her unconscious Oedipal feelings; the latter awareness being an awareness that she had been so moved.

In the same way, our awareness of the unconscious springs of our apparently passive states or affections would also be distinguishable from our awareness of the conscious springs of these states. Believing (a) that I am annoyed because I have failed to finish some work in time is distinguishable qualitatively from believing (b) that unconscious feelings of rivalry and competition have fed my annoyance. And so is (c) believing that I am annoyed because of the irritating noise of children playing outside, and (d) believing that I am annoyed because unconconsciously I am regretting having failed to finish the work in time. The reasons for our actions and so too the objects of our affections which are known to us in the ordinary conscious way are always

distinguishable, intuitively, from any reasons for our actions and states which we might hypothesise as being their unconscious movers.

Thus, the expression "being conscious of P" (or "being aware of P") can be shown, as was suggested earlier, to be ambiguous:

1. Saying that a person is conscious or aware of a belief or desire of his or hers may be taken to mean that that person knows or believes that that state is to be ascribed to him or her. And, in addition,

2. It may be taken to mean that the agent's awareness of the belief or desire in question presents itself in a qualitatively distinguishable way.

If the latter condition obtains, then the person may still be said to hold the belief or desire unconsciously in the technical sense, even while being aware that the state may be ascribed to him or her: ascribed, that is, either by others or by a theory which he or she accepts in general terms.

Armed with this intuitive distinction we can now insist that while the people in Cases 1 and 2 can be said to have known about the unconscious mental states influencing their behaviour, they were not completely conscious of them; technically speaking they remained unconscious mental states. In order that we avoid the apparent contradiction mentioned, it must be the case that there is something qualitatively different about our awareness of our unconscious beliefs and desires. As long as this distinction holds, it can be argued that the examples may be described without contradiction. And, at the same time, it can be shown that a slight modification of our definition of unconscious mental states will overcome the difficulties in its previous formulation which the examples were initially introduced to illustrate.

We need to modify the definition by adding a qualification to the effect that the awareness required be of the appropriate kind. Tentatively, then, we can say that:

(C) Unconscious mental states are those of our mental states of which we are unaware and of which we cannot become aware, except in the qualitatively distinguishable sense which we may call "awareness X", at the time at which the action or state which they influence, and which they are introduced to explain, is taking place.

But what is the nature of "awareness X"? What is the peculiar qualitative difference between the awareness people have of their ordinary conscious beliefs and desires and the awareness they have of their (previously) unconscious mental states? While we can accept intuitively the possibility of knowing of and acting upon unconscious mental states which do not present themselves in the same way as our ordinary conscious reasons for acting, it may be supposed that ensuring that this intuitive distinction is requisitely comprehensive would require ideally that we show some concrete characterization of the difference upon which it is based. I shall argue presently that the latter demand is misdirected, but before doing so, I wish to examine in a little more detail the directions in which such a demand may be expected to lead us.

In trying to characterize "awareness X" we might look first at a criterion commonly introduced to characterize our awareness of the reasons for our actions and of our own conscious mental states. We are said to know our reasons and our own mental states "immediately" or, in a more recent gloss of that expression, "non-observationally",[20] by which has traditionally been understood both that we are aware of our conscious mental states and reasons without grounds or evidence and that our knowledge of them is infallible. The pitfalls of this interpretation are legion and well documented. We do seem to be highly fallible concerning our conscious mental states - perhaps about whether we have them in any particular case, certainly that they take the objects we imagine and function as reasons in the way that we suppose. Thus, any attempt to characterize the difference between our awareness of our ordinary conscious mental states and "awareness X" in terms of the infallibility of our knowledge of the former, or the feeling of certainty we seem to enjoy in relation to our judgements about them, would be both misplaced and question begging.

Limiting ourselves, then, to the former interpretation of the idea of immediate knowledge - that is, knowledge without observation - we can see that if there is an intuitive distinction between knowing why one acts in the ordinary way, on the one hand, and having an hypothesis about the unconscious motive for what one does, on the other, then we might well have a way to characterize the difference. In the latter case, a person would go about establishing the existence of his or her unconscious mental states just as one might establish the existence of another person's ordinary conscious mental states or as one might discover one's purely behavioural traits. Either the person would be told, or he or she would form an hypothesis on the basis of the evidence provided by his or her knowledge of psychological theory and of his or her past habits, psychological characteristics, person history and the like. In the former case, on the contrary, appeals to such facts do not occur: a person seems just to know the reason why he or she acts.

Thus, it might be supposed that we need merely alter our definition of unconscious mental states to read:

> (D) Unconscious mental states are those of our mental states of which we are unaware and cannot become aware except when the awareness is acquired 'observationally' by reference to evidence or grounds.

But a further difficulty arises. We well know that some beliefs which initially were known observationally become, in time, merely by dint of their familiarity, part of our store of non-observational knowledge. This is true of certain beliefs about ourselves such as those concerning our own names, for example. In the same way, we can imagine a person reaching a state of familiarity with his or her unconscious mind which would obviate the necessity of his or her carrying out any reasoning or inference in order to assert the presence of some unconscious mental state. Thus, in a sense, some of his or her unconscious beliefs and desires might be said to have become, for that person, non-observationally known.

There is an obvious reply to this claim, however. It may be objected that neither knowing one's own name, nor knowing of

one's unconscious mental states would count, strictly speaking, as cases of non-observational knowledge. For, in each case, although at some particular time of assenting to them we do not hold our beliefs on the basis of the evidence, nevertheless
(a) there is evidence available which might have been adduced in support of such beliefs, and
(b) originally, when we first acquired such beliefs, we would have done so on the basis of evidence.
Moreover, it may be said, when we know something non-observationally in the strict sense, we do so because there is no evidence which could be adduced in support of it. Knowledge of our conscious mental states and of the reasons for our actions is said to be like intuitive knowledge in this respect: we have such knowledge without grounds because it neither has nor requires grounds. Thus, it may be proposed that the class of unconscious mental states of which we are aware is distinguishable from the class of our ordinary conscious mental states in that only our knowledge of the latter counts as what might be called "original non-observational knowledge".

But the above refinement, too, is unsatisfactory. For there remains the possibility that some unconscious mental state of which we are aware might originally have been a conscious state and might, thus, have originally been known non-observationally. For example, I might have felt X at time t_1 and hence entertained the conscious belief that X at time t_2 non-observationally, and this would count as a case of knowledge which was originally non-observational. At time t_2, however, I might come to hold my belief that X unconsciously, and then at time t_3 my belief might again have 'entered' my conscious mind. Now, if we bring in the element of originality described earlier, we seem committed to judge my belief at time t_3 according to its initial status as originally known non-observationally to the person whose belief it is.

We are thus left without a useful interpretation of immediate or non-observational knowledge. Neither if we read it to mean simply "known without grounds or evidence", nor if we read it as "known without grounds and groundless", will this characteristic be sufficient to exclude the class of cases of unconscious

mental states of which the agent is aware from the class of ordinary conscious mental states.*

In discussing the clinical phenomenon of "resistance", Freud himself makes the observation that merely becoming aware of his or her previous unconscious mental states during the course of analysis does not by itself put a person in the same relation to them as that person enjoys in relation to his or her ordinary conscious mental states. But, Freud's characterizations of the nature of the qualitative difference are, for our purposes, regrettably few and casual.[21] He remarks that the immediate effect of revealing some hitherto repressed and unconscious idea to a patient is negligible:

> ".....our telling him makes at first no change in his mental condition.....it does not remove the repression nor undo its effects, as might be expected from the fact that the previously unconscious idea has now become conscious."
>
> ("The Unconscious", Page 125)

---

* There is actually an additional difficulty which may be supposed to arise with the introduction of a stipulation about qualitative differences in awareness in terms of observationality and non-observationality, which can best be brought out if we consider again the second of the two cases described and, in particular, the second formulation of it (Case 2 (ii)). This is the case where the woman has been ignorant both of her preference for older men and of what has been called, in contrast, "her reason" for her preference. If we conceive our present task as one of showing wherein lies the difference between the case of a consciously motivated woman aware of her attraction to older men, on the one hand, and that of an unconsciously motivated woman who is ignorant of her preference, on the other, and if we are proposing that the reason for her selection of companions may in each case be said to be this preference itself (rather than being the reason for the preference, which has hitherto been referred to as "the reason"), then it seems worth pointing out that the woman's consciously motivated preference might be based upon observation as much as would an inference about her mental states drawn by the unconsciously motivated woman. There is not always a datum of conscious experience to tell us of our own preferences. It may be necessary for the consciously motivated woman to discover by observation the reason for her behaviour in the cases where her "reason", viz. the preference, has not hitherto been an unconscious one, in a similar way to the way in which the unconsciously motivated woman might attempt to discover the reason for her preference. Now it may be said, in reply, that the preference, whether conscious or unconscious, is not a reason: it is merely a disposition, and to propose it is never to explain but merely to restate with greater generality. But, it does seem worth pointing out that we do cite preferences, both conscious and unconscious, as reasons for certain behaviour.

What does Freud mean by the removal of the repression and the undoing of its effects? He goes on to give a technical account of the nature of this phenomenon:

> ".....there is no lifting of the repression until the conscious idea, after overcoming the resistances, has united with the unconscious memory trace. Only through bringing the latter itself into consciousness is the effect achieved."
>
> ("The Unconscious", Page 125)

But such an account is not of help to us since it fails to describe the qualitative difference between an idea which has been united with its unconscious memory trace and one which has not. In another passage in which he describes having been asked to advise analysts who complained that revealing to the patient his or her unconscious states made no difference to that patient's behaviour or mental state, Freud is a little more explicit:

> "One must allow the patient time to get to know this resistance of which he is ignorant, to 'work through' it, to overcome it..... Only when it has come to its height can one, with the patient's cooperation, discover the repressed instinctual trends which are feeding the resistance: and only by living them through in this way will the patient be convinced of their existence and their power."
>
> ("Further Recommendations in the Technique of PsychoAnalysis" (1914), Page 375)

Linking the preceding remarks with an earlier passage from the same paper in which Freud speaks of:

> ".....(convincing himself that) after his resistances have been overcome, the patient no longer invokes the absence of any memory of them (sensation of familiarity)* as a ground for refusing to accept them." (op. cit., page 369)

we can perhaps conclude that even when acquainted with the interpretation of his or her behaviour in terms of a particular unconscious mental state or memory, the patient at first resists the interpretation literally, by simply denying that it is correct, and (and this is the part which is of importance to us here) 'resists' the interpretation by failing to feel towards the putative unconscious state as he or she would towards an ordinary conscious state, that is, by failing to feel that

---

* My underlining

"sensation of familiarity" - a sort of acceptance - which characteristically accompanies mental states of which we are conscious in the ordinary way.*

The problem which confronts us now is this: how much can be made of this characteristic feeling in our search for a qualitative criterion for distinguishing ordinary conscious states from those of our unconscious states of which we are aware? It has been suggested by Peter Alexander that on becoming aware of them as a result of analysis the patient comes to and remains feeling towards his or her unconscious mental states "as a stranger".[22] This expression is very promising as an attempt to give a qualitative description of the difference in awareness to which we have been referring. But it is ultimately metaphorical - in finally admitting that the hitherto unconscious reason is the reason why he or she acted, felt or behaved in a particular way, the patient is admitting to some relationship of ownership to it. In the most obvious physical and causal sense, it is his or her reason since it is the reason or motive which, at least in part, brought about the action or affection in question. Thus, the sense in which it is not his or her reason but rather is a reason which is strange and alien, seems to require further elucidation.

Can it be that a characterization of the patient's unconscious mental state of which he or she is aware as a strange and alien 'feeling' or 'seeming' indicates an actual datum of experience? Could the claim be that there is a distinctive feeling associated with an awareness of these states - something, perhaps, if we follow Freud's suggestion in the last passage quoted, along the lines of the absence of that "feeling of familiarity" which

---

* It ought perhaps to be emphasized again that I am interested in Freud's empirical claims here, and not his theoretical hypotheses. Whether or not the concept of resistance actually explains the peculiar quality of the patient's awareness of these states need not concern us. I am interested simply in Freud's empirical claim that the state of familiarity is present in the one case and not in the other.

William James claims accompanies our memories of our own experiences?[23] I do not believe that this 'special' awareness of unconscious states can be so characterized as a feeling state. The sort of awareness we are trying to capture, and with which Freud was clearly familiar, is minimally (but not only) an awareness that..... That is, an awareness on the part of the agent that he or she is unconsciously motivated. But more than that it is a sort of intuitive acceptance as a result of the belief presenting itself in a distinguishable way. It might be said that the agent accepts, and not merely acknowledges, that he or she is indeed unconsciously motivated, but yet the belief is, and remains, strangely alien. That is, it is not fully embraced at a conscious level. It is difficult to characterize more accurately. However, it is not, at this stage, of crucial importance to attempt further clarification. Few of our subsequent arguments depend on any definitive characterization. It is, in any case, unimportant because any unconscious mental state which is deduced while we are still doing whatever it is which the state is introduced to explain either will become a conscious motive, reason or state, or else will strike us, intuitively, as different. If it becomes a conscious state (motive or reason), then we no longer have a case of at once being aware of and acting upon, or being moved by an unconscious state (motive or reason) at all. It becomes the conscious reason or motive for our acting or being affected. Let me illustrate this point by returning to the examples we discussed previously.

In Case 2 the woman was described first as accepting invitations and flowers from her older suitors while, at the same time, knowing that her wish to do so derived from an unconscious Oedipal attraction to them. In this case, the woman experienced no conscious Oedipal attraction despite her "knowledge" of her unconscious attitude. However, we can imagine that the case might change, so that what were unconscious Oedipal feelings of which she happened to be aware, might become conscious feelings. Consciously, she might come lustfully to identify her older suitors with her own father. But as long as she did do this, then what was an unconscious state would have

become an ordinary conscious one. (And the case would be the same as the one described earlier where the woman's actual incestuous relationship with her own father meant that what might in an extended sense still be described as her 'Oedipal' motivation, occurred on a thoroughly conscious level.) Similarly, in Case 1, on becoming aware of why he wished to skip the cracks in the pavement, the man might either (a) continue to do so while aware of the unconscious reason for his wish to do so, but without embracing it as the reason, or (b) he might so embrace it. If the former case (a) were to obtain, then his avoiding the cracks might still be said to be based on, or occasioned by, an unconscious reason, despite the reason's now being one of which he was actually aware. Only in the latter case (b), however, has the reason 'become' his conscious reason.

Either the person's state appears distinguishably alien and different to him or her despite being acknowledged as the state which occasioned the particular action or affection, or it does not. If the latter (that is, it is not distinguishably alien) is so, then it is acknowledged not only as being the state occasioning the behaviour, but also as being, in some sense, the person's own reason for behaving thus. It is not possible for the person to be mistaken in his or her judgement that the state is of one kind or the other. We are infallible at least with regard to our knowledge that a mental state is a conscious one, rather than an unconscious one of which we happen to be aware, even though we are not infallible with regard to the actual relation between our behaviour and either of these kinds of reason. (I may be wrong in supposing that "my conscious reason X" actually moves me to act, just as I may be wrong in supposing that "unconscious reason Y of which I am aware" did so.) And, presumably, the explanation of the small degree of infallibility which we do enjoy is simply that being "my conscious state" merely means appearing to me in a distinct way which is intuitively obvious to me. Since we are fallible with regard to judgements made about the relation between our ordinary conscious states and the behaviour they occasion, then their only characteristic feature is their appearance or phenomenal presentation, and we cannot be wrong about that. And, because we must be able unerringly to distinguish conscious from unconscious states, no further characterization of the quality of our awareness in either cases could be required.

So, in the troublesome cases of our seeming to be able to be aware or conscious of our unconscious mental states - the cases which looked to be a hindrance to our definition of 'unconsciousness' in the technical sense - we still have a way of characterizing the difference between unconscious mental states and ordinary ones, even when the former unconscious ones are known to the agent at the time at which the behaviour or state which they are introduced to explain is occurring. We are thus in possession of a definition of unconscious mental states which is sufficient to distinguish the technical sense of 'unconscious of' and 'unconsciously' from the looser, Non-Freudian and non-technical sense in everyday use. The definition:

> (E) Unconscious mental states are those of our mental states of which we are unaware and of which we cannot become aware at the time at which the behaviour and states which they are introduced to explain are occurring, except in a way which is always qualitatively distinguishable from the way in which we are aware of our ordinary conscious mental states.

Before turning to a more detailed examination of the concept of unconscious motivation, I wish to make some observations about the relationship between the concept of unconsciousness as I have attempted to elucidate it in this chapter and the allied concept of self-deception.

Unconsciousness and Self-Deception

How does our technical concept of unconsciousness relate to that one which has received so much attention in recent philosophical literature, the concept of self-deception? First, we must ask what is meant by the term 'self deception'. By way of giving a rough, working definition, we might say that 'self-deception' describes the activity of putting oneself into a position of entertaining some proposition while at the same time believing a contradictory proposition to be true (and thus necessarily being aware, in some sense, of the contradiction).

Any account of self-deception must adequately distinguish that phenomenon from several others which bear some similarity to it and which may be confused with it. There are, particularly,

(1) Cases where a person asserts P or so acts as to imply the assertion of P while believing not-P or not believing P. This, of course, is the deception of others, or lying. Closely related and again distinguishable from self-deception is:
(2) The case where a person asserts that P or acts so as to imply the assertion of P, or inwardly makes assent to P in order to persuade, comfort, encourage, etc., him or herself, while believing not-P or not believing P. The latter kind of behaviour is best described as pretending to oneself. Another kind of case which an account of self-deception must exclude is
(3) That of entertaining contradictory beliefs completely unwittingly, or 'unconsciously' (in the non-technical sense introduced earlier in this chapter), when the beliefs held are merely so complicated or so full of subtle implications that the contradictions inherent in them are not apparent to the person holding them.

The three cases which I have outlined are, in varying degrees and ways, related to the phenomenon of self-deception. Deceiving others and pretending or lying to oneself both merge into and, indeed, often become self-deception. And, we may describe the motive for pretending or lying to oneself as the wish to deceive oneself. At time t_1 I may tell myself something which I do not believe to be true, pretending to myself in the hope of deceiving myself in doing so. And at time t_2 I may have succeeded in deceiving myself. But pretending to onself is not always an example of or an attempt at self-deception in that way: day-dreaming and wishful thinking often may be instances of engaging in the former without being an instance of the latter. So the two activities and the two motives are distinguishable.

In our previous discussion, the notions of consciousness and unconsciousness have been understood as properties of states of mind or as states of a person vis-à-vis particular states of mind. In the same way, it seems possible to speak of a person as being in a state of self-deception in relation to one of his or her mental states. For instance, the state of self-deception vis-à-vis a particular belief may be supposed to be distinguished

from the state we are in when lying or pretending to ourselves, or lying or pretending to others, about that belief, in that in the case of self-deception, we both believe P and believe not-P, whereas in the cases of lying and pretending to onself we merely assert that either P or that not-P is true while believing the opposite to be true. And the state of self-deception appears to be distinguishable from the phenomenon described in (3) above (holding contradictory beliefs unwittingly) because, while in both cases the agent believes two contradictory propositions, only in the case of self-deception is he or she also aware of the contradiction.

However, more commonly, self-deception is understood dynamically, as a mental activity or what might be described as a motive for a particular action or procedure, and in that respect the concept of self-deception runs more closely parallel to the dynamic Freudian concept of repression than it does to the static concept of unconsciousness which we have been attempting to capture in our definition.

Even though the dynamic concepts of repression and self-deception are clearly related, it is not obvious that the phenomenon of repression in the strict sense implied in the Freudian theory, and the activity of self-deception are identical, as sometimes has been supposed.[24] Some cases which might be described as indicating self-deception and, indeed, many of the examples which Freud cites, imply the presence of less than such full or total repression as would correspond to, and be taken in Freudian theory to result in, unconsciousness of the kind implied in our definition.*

In fact, it is perhaps generally true that cases of self-deception are more commonly met with than strict cases of repression, and that the class of cases of "unconscious mental states" designated by our definition may be a relatively small one.

---

* Freud's own concept of "suppression" is designed to handle such cases.

First and foremost, "self-deception" is treated as a description of something which the agent does, a procedure engaged in (we say that person X had deceived himself into believing P, for example), or one intended by engaging in other manoeuvres (we say that by doing A, person Y tries to deceive herself), and not a state which, at a particular time, the agent is in, at all. And, because the concept of self-deception is not primarily a static one, in the way that that of unconsciousness is, it is not as well suited to a philosophical analysis of the sort to which we wish to subject that of unconsciousness, (although it is, of course, possible to evaluate the moral worth of the activity of self-deception, as some philosophers have done).

For this reason, at risk of introducing what may, in fact, be, as I remarked earlier, a very small class, I wish to restrict my discussion to the concept of unconsciousness and to the class of mental states fitting the description in our definition and, except inasmuch as they may be supposed to be identical, I shall not deal with those mental states described as indicating or resulting from the apparently more widespread phenomenon of self-deception.

To conclude the first chapter: my approach so far has been to examine the possibility of a notion of unconscious motivation which will stand on its own and which does not depend upon the correctness of either Freudian theories or empirical claims. I have argued that a Non-Freudian concept can, indeed, be identified and I have used some plausible examples of unconsciously motivated behaviour to refine the concept. Because of that, I have dwelt at length on the coherence of the examples. I have made a plausible distinction between:

(1) consciously knowing why one acts in an ordinary way, and

(2) consciously having an hypothesis about the unconscious motives for what we do.

I have tried to show how we can describe cases in which a person is acting upon or being moved by an unconscious motive or state (and therefore not completely conscious of it, in one sense), while at the same time being conscious of that same (unconscious)

motive or state. The paradox, I have argued, is more apparent than real. It rests on the mistaken belief that "being conscious of" a mental state and "being unconscious of" a mental state are contradictories. It is the case, however, that the expression 'being conscious of unconscious motivation' has two main senses (1 and 2 below). In order that the apparent paradox is avoided, it must be the case that there is something qualitatively different about our awareness of unconscious beliefs and desires. That this is so is, interestingly enough, found implicitly in Freud's own straightforward descriptions of patients. There are qualitatively different types of (conscious) awareness of unconsciously motivated behaviour:

(1) A simple acknowledgement of an unconscious motive where the agent or patient acknowledges an analyst's hypothesis.

(2) An intuitive acceptance - as a result of the belief presenting itself in a qualitatively distinguishable way (perhaps, for example, after analysis) - of the fact of unconscious motivation; and, where the agent is, indeed, still unconsciously motivated and yet the belief (the hypothesis in 1) is and remains strangely alien.

(3) A full embrace by the agent at a conscious level of the motivation involved. In this case the behaviour ceases to be unconsciously motivated.

The first sense of awareness is unimportant for my purposes. It is the second sense of awareness which is of interest to me in my definition of unconscious mental states. The qualitatively different sense of awareness involved is difficult to characterize, but it is real enough. It is not a feeling, but an intuitive acceptance.

Thus, we have now arrived at a working definition of unconscious mental states (definition E on page 34), which I will employ in subsequent chapters. I now turn to a more detailed examination of the concept of unconscious motivation.

Notes and References

1 The metapsychological essays are reprinted in Freud: General Psychological Theory, edited by Philip Rieff (Collier Books, N.Y., 1963). All page references are to this edition.

2 The Ego and the Id, by Sigmund Freud, translated by Jean Riviere; revised and newly edited by James Strachey (The Hogarth Press and the Institutes of Psycho-Analysis, London, 1962). All page references are to this edition.

3 See "Repression" in metapsychological essays, op.cit., Page 110.

4 In The Standard Edition of the Complete Psychological Works of Sigmund Freud, edited by James Strachey (Hogarth Press, London, 1953-1964).

5 Collected Papers of Sigmund Freud (Hogarth Press, London, 1957).

6 See "Repression", page 106; "The Unconscious" in metapsychological essays, op. cit., Page 122.

7 See "Repression", pages 108, 111.

8 See "The Unconscious", page 118.

9 As Wolheim has observed (Freud, Fontana Modern Masters, London, 1971, pages 154-155), Freud actually introduces two different and incompatible analyses of "primary process" thinking. Sometimes he treats it as the characteristic of all unconscious processes - at other times as the characteristic disguise which unconscious ideas don in order safely to enter consciousness. Only if Freud used the former analysis would he have a criterion for distinguishing unconscious mental states in the way which I have implied. Yet, if he did, then problems face him in the apparent rationality of much unconscious thinking; and, in fact, the evidence in some ways favours the latter interpretation.

10 See "The Unconscious", page 134

11 Ibid.

12 (a) Ibid., page 147

(b) Freud's analysis of bringing something to (pre) consciousness has an echo in a recent attempt to characterize the same activity in terms of "spelling out". In his discussion of the phenomenon of self-deception (Self-Deception, Routledge and Kegan Paul, London, 1969), Fingarette argues that becoming conscious of some mental states is best analysed in terms of exercising an essentially verbal skill which he labels "spelling out". See Chapter 11, especially pages 39-47 and page 121.

13 For example, that given in the Encyclopedia of Philosophy, edited by Edwards (MacMillan, New York, 1967).

> "Those mental states of which the individual is not aware while they occur to him."

14 For example, by MacIntyre, The Unconscious (Routledge and Kegan Paul, London, 1958), who asserts that the inter-dependence noted here between Freud's various theoretical concepts is an unbreakable one:

> ".....Freud's psychoanalytic technique no less than his doctrine of the mind depend on certain key theoretical concepts which can only be understood in terms of each other. This mutual interdependence of concepts.....is nothing new in the history of science. But, clearly, a comprehensive theory whose concepts are thus interwoven stands all the more in need of justification as a whole. And the whole concept of 'the unconscious' stands or falls with this general theory."
>
> (Page 13)

15 See, for example, criticisms and rejection of the Freudian concepts of psychic energy and repression in Colby, Energy and Structure in Psychoanalysis (Ronald, New York, 1955); Gill, "The Present State in Psychoanalytic Theory", J. Abn. Soc. Psychol., 1959; Holt, "A Review of Some of Freud's Biological Assumptions and Their Influence on His Theories", in Psychoanalysis and Current Biological Thought, edited by Greenfield and Lewis (Univ. of Wisc. Press, Madison, 1965); and in Freud and Psychology, S.G.M. Lee and Martin Herbert, eds. (Penguin, 1970).

16 See MacIntyre's description of:

> ".....a set of memories so painful that they cannot be brought into consciousness by any normal means."
>
> (Page 26, op. cit.)

17 For example, Freud gives the following list of the possible fates of once unconscious wishes:

> "The patient's personality may be convinced that it has been wrong and may be led into accepting it wholly or in part: or the wish itself may be directed to a higher and consequently unobjectionable aim (this is what we call 'sublimation'): or the rejection of the wish may be recognized as a justifiable one, but the automatic and therefore inefficient mechanism of repression may be replaced by a condemning with the help of the highest human mental functions - conscious control of the wish is attained."
>
> (My underlining here)
>
> Page 368, "Further Recommendations in the Technique of Psychoanalysis, Recollection, Repetition, and Working Through"(1915).

18 See, for example, the definition proposed by Collins ("Unconscious Belief", J. Phil., 1969), in which he argues that a person's belief is called unconscious because that person denies that he or she has (that is, believes) it. If there is compelling reason for "He believes that P" and "He candidly denies that he believes that P", then P is an unconscious belief. By contrast, see: Peter Alexander, "Rationality and Psychoanalytic Explanation", Mind, 1962, Page 338:

> "We can never know what it would be like to act with (the typical unconscious reason) in mind since it does not make sense to talk of acting with unconscious (in the technical sense) reasons in mind. The typical unconscious reasons are not the sorts of reasons which would lead to that sort of behaviour if we were conscious of them. The shocked reaction 'Good gracious, is that why I did it? I should never have done it if I'd known' is typical and says more than the speaker, and perhaps Freud, usually realizes. The whole point of the theory is that neurotics behave as they do because they fool themselves completely about certain things, but we cannot fool ourselves completely and be aware that we are fooling ourselves. Unconscious reasons are not just possible conscious reasons for the behaviour in question; they would not be regarded as reasons for it if they were conscious. (And) this does not mean simply that the patient would see the reasons as disreputable if they were conscious, but that he would see them as inadequate."

And Mullane, "Psychoanalytic Explanation and Rationality", J. Phil., 1971, Pages 423-424:

> "The reason why (the neurotic) gives up his neurotic behaviour rather than adopt his hitherto unconscious beliefs as operant, conscious ones and then proceeds to continue to behave as if he had is that these unconscious beliefs cannot operate as his (conscious) beliefs. Insofar as he is fully conscious of the nature of his reasons and the nature of his behaviour, he cannot behave for bad reasons, i.e., he cannot behave in self-defeating or neurotic ways. His rejection of these reasons takes place because they are totally and necessarily incompatible with and alien to his conscious or (as they are sometimes called) 'real' desires."

19 Speaking of one of his obsessive patients he remarks that:

> "The interpretation of the symptoms was discovered by the patient herself in a flash, without guidance or interference from the analyst....."
>
> (Page 275, General Introduction to Psycho-Analysis, Pocket Books Edition, N.Y., 1971)

(My underlining)

20 See Anscombe, E., Intention, (Blackwell, Oxford, 1957)

21 Nor is the distinction which MacIntyre draws in discussing this phenomenon, between becoming aware of an idea and becoming aware of a proposition, a helpful one, although remarks he makes are interesting:

> "To bring an idea in to conscious it is not enough for someone simply to become aware of the idea. An analyst, for example, may tell a patient something that the patient has forgotten because he has repressed the idea. But although the patient has been introduced to an idea which is in his unconscious he has not thereby made the idea conscious. To do that, he must bring out into consciousness his own unconscious idea. ('You still fear your father', the analyst may say. But the patient's reception of this, his agreeing to it even, is different from his becoming aware of his own idea of his father as an object of hate.) 'Ideas' therefore are to be identified with propositions. They are discrete mental entities which although they may be verbalized correctly in a certain form of words are still sharply distinguishable from other ideas which are similarly verbalizable."
>
> (Page 34, op. cit.)

But MacIntyre's distinction is not helpful. It merely reiterates the fact that the intuitive distinction which we have been discussing is indeed present. If I understand him correctly, MacIntyre is suggesting that when we are aware of any ordinary, conscious mental state, we are aware both of a proposition, viz. that the state is truely ascribable to us, and we are directly aware of the state itself, or, in the case of beliefs, presumably of the proposition there expressed. In contrast, when we become aware of unconscious mental states we are merely aware in the former way. We are aware that some state (in this case, the unconscious one), is applicable to ourselves. However, such an analysis of an awareness of an ordinary conscious mental state is artificial. While it is sometimes possible to distinguish these two separate objects of our awareness of our ordinary conscious mental states, very often the distinction fades. "Being aware that P" is ambiguous as between expressing each kind of act of awareness, and this ambiguity, it seems to me, is informative. In fact, having ordinary conscious mental states does not always consist in having each distinct act of awareness - sometimes feeling X or wanting Y consists merely in being aware that X is ascribable to me, or that I want Y - it is certainly not the case that we infer the one proposition from the other, as this kind of model of awareness is often taken to presuppose (I know that or am aware that I feel X about P on the basis of my being aware of my feeling). The distinction between being aware of an idea and being aware of a proposition, while it applies in certain cases of awareness of emotions (although apparently not to all - I might be aware of no particular feelings in becoming aware that I

feel jealous), and of wants and desires, fades when we speak of awareness of beliefs: "being aware of P" is merely eliptical for "being aware that I believe that P" and "being aware that I believe that P" presupposes "being aware that P".

22 Peter Alexander, op. cit.

23 William James, Principles of Psychology, Page 201, (Cleveland, Fine Editions Press, 1948).

24 For example, by Fingarette, op. cit.

CHAPTER TWO

Summary

In this chapter the concept of unconscious motivation is analysed. In general, it is accepted that motives comprise desires, wants and wishes and that some comprise, in addition, beliefs about achieving end states. Further, that a motive explanation is one that says that an action is in some way due to a want or desire. At a general level this approach is not especially controversial. Problems arise when we ask: how do motives give rise to action? This is addressed in attempting to see how it is that unconscious motives might be said to give rise to action.

In tackling the concept of unconscious motivation, first, the concept of ordinary conscious motivation is shown to comprise two distinguishable models, of which only one is as precise and purposive as to imply the presence of particular 'goals' or 'ends' on the part of the person to whom the motivation is ascribed, and beliefs about the likelihood of the behaviour in question to bring about those goals or ends. Explanations citing motives are contrasted with various other kinds of explanation and, particularly, attention is drawn to the way they differ from explanations citing signs and symptoms. It is shown that a person can *have* a motive without being motivated by it. That is, merely knowing that a person *has* a motive for doing X and has done X, is *not* sufficient to enable us to explain the action by *reference* to the motive.

Intentional action is defined as action for which there are reasons, and it is argued that since merely wanting or having a *whim* to undertake the action in question counts as having a reason, not all intentional action can be said to be motivated. A whim is characterized as being unconnected with any other of the agent's desires. It is argued that only when we regard an explanation as citing *all* the conditions necessary for bringing about an action that we can be said to explain by mentioning the want to do X in answer to the question, "Why did P do X? *when X*

is a motivated action. Merely by citing the agent's desire or wish to act thus we cannot be said to explain a motivated action in the sense of giving a reason for it, although we do explain a whim in that sense by doing so.

Freud's theory of unconscious motivation is discussed and it is shown that while some of the phenomena which Freud cites seem to require the description of motivated by unconscious mental states, other behaviour would appear to be better described by saying that it is a sign or symptom produced or caused by such states. It is argued that only by postulating an additional or secondary unconscious mental state - a wish to express the unconscious belief or wish first postulated - is this latter class of behaviour transformed into behaviour which appears to be unconsciously motivated. Different kinds of unconscious motivation are distinguished, and in the light of these distinctions a definition of unconscious motivation is offered. The thesis of overdetermination is discussed and defended, and three kinds of explanations are distinguished, corresponding to the three ways in which unconscious mental states can be shown to motivate behaviour.

So as to clarify at the outset my preferred approach to the analysis of motivation I state, at present without justification, but for reasons of clarity and help to the reader, the following views which I hold. I express my views in ten propositions:

1. I use 'motive' as the central term: not 'reason';
2. Following Davidson, I claim that intentional action is a class for which there are reasons;
3. Some intentional action is unmotivated;
4. Some action for which there are reasons is unmotivated;
5. All motivated behaviour is intentional;

Therefore:

6. For all motivated behaviour there are reasons;
7. An action done on a whim is one for which there is no motive: but while it is unmotivated, it is done for a reason (that is, the whim itself);

Therefore:

8. An action for which there is a reason, but no motive, is an action done merely because the agent felt like it (that is,

had a whim);

Therefore:

9. An unmotivated killing, say, may still have been an intentional action;

Importantly:

10. The <u>reason</u> for any motivated action is just the motive with which it is undertaken.

* * * * * * *

In Chapter 1, I discussed the meaning of 'unconscious of' and 'unconsciously' both as these expressions are used in technical explanations associated with psychoanalytic theory and in their Non-Freudian uses. It was remarked that the technical use of these expressions differs from the non-technical one in two important respects:
(1) Only in the technical sense are we treated as necessarily unaware of that of which we are said to be unconscious; and
(2) Only in the technical sense are we said to be unconscious of reasons and motives.
I tried to show that there was justification for driving a wedge between being conscious and being aware in a special sense.

It is to the second feature of the technical concept which I now wish to turn. The question of the exact way in which unconscious mental states might be supposed to relate to and affect behaviour and ordinary conscious mental states in what is loosely labelled unconscious motivation, is one which requires careful attention. As well as considering what has been intended by uses of the expression 'unconscious motivation', we must examine what could be meant by it, from where the evidence for claims about unconscious motivation might be derived, and whether, evidence or no, the notion of motivation which is unconscious is conceptually coherent at all. Let us begin by considering what 'unconscious motivation' might be expected to mean, and to do so, let us cast a glance at the more familiar notion of "motivation" in general by asking what it means to say that a person's behaviour is in the ordinary way, or consciously, motivated.

Section 1: Conscious Motivation

Motives and Motivation

There are two somewhat different conceptions or models of motivation. On the one hand (1), to say that a person's behaviour is motivated sometimes means merely that it is to be seen and understood in the light of some motive(s), when by a motive is meant some emotion or attitude such as guilt, generosity, love,

remorse, pride, greed, jealousy, etc., the very concepts of which have built into them directedness towards certain characteristic and recognizable sorts of behaviour. When they function as motives such emotions and attitudes are associated with familiar patterns of feelings and tendencies to act and react. We understand what is meant when it is said that P acted out of jealousy because we are familiar with what might be described as the jealousy syndrome: the complex pattern of feelings, wishes and behavioural reactions, both voluntary and involuntary, which are occasioned by particular (perceived) states of affairs.

On the other hand (2), a motive sometimes is conceived of more narrowly and more precisely as:
(i) some desire or wish whose object is either a goal or end state of affairs for the achievement of which some particular behaviour or action can be regarded as a means, or that behaviour or action itself as conceived under some particular description, together with:
(ii) a belief that this particular behaviour or action is likely to produce that state of affairs.

It is possible for the desired goal or end state of affairs to be that same behaviour as seen under some particular description in the following way. The motive underlying my behaviour may be politeness or good manners. In this case, my goal need not be some further state of affairs likely to be produced by my behaviour, but merely my behaviour <u>qua</u> polite behaviour. Of course, it may be that I behave politely in order to impress those around me, although then I should at least be said to have an <u>additional</u> motive in my wish to impress. But, while in the case first described, I act solely in order to be polite, nevertheless, my behaviour can be regarded as the means to some end - the end, in this case, of being polite.[1]

It may be supposed that the class of motives described in (1) reduces to the class described in (2) - that we can rephrase sentences about motives of the sort listed in (1) into sentences about desires and wishes and hence goals and ends, making a complete

"translation" without change of meaning.[2] It is true that in many cases, this "translation" is possible. We can say, for instance, that the motive for P's killing the man who killed his father was revenge. Or, equally well, we can say that P killed the man because he wanted to avenge his father's death, and that the state of affairs in which his father's death had been avenged was the goal or end state of affairs for which killing the man might be seen as the means.[3] Or, we can say that in acting out of friendship for someone P was wanting to bring about that person's well-being or happiness. In other cases, such a "translation" seems forced, and sometimes we even are hard pressed to identify the appropriate desire(s) or wish(es) constituting the agent's goal or object in acting. If a person renounces some pleasure out of guilt, takes his or her own life in despair, or ignores someone in the street out of snobbishness, it may be supposed that, while in each case that person was acting irrationally unless he or she had some particular goal in so doing, that person may have so acted in the absence of such a goal.

At the risk of sounding somewhat awkward and unnatural it is nevertheless possible to give a schematic analysis of the more recalcitrant concepts such as guilt, despair and snobbishness, identifying the kinds of wants, desires and wishes (and aversions) which make them up. (Wants and desires which must be distinguished, of course, from simply wanting to do whatever it is that is done. If a person P acts snobbishly in doing X then, although it may not be true that P necessarily wanted to act snobbishly, it is true to say of P, assuming that the action was intentional at all, that she wanted to do X. We shall return to the significance of these kinds of wants in a later section of this chapter.

For example, if we say that P's ignoring Q was motivated by snobbishness, then we may mean all or some of the following:

P wanted to avoid speaking to Q and being seen with Q;
P wanted Q to see her and to have a certain belief about P's superiority;
P doubted her superiority and wanted to establish it by acting as if she were superior, and so on.....

If we ascribe snobbishness when none of the kinds of desires and wishes mentioned would be ascribed, then our correct understanding

of the expression 'snobbish' might be questioned.

But to say that there are always constituent desires and wishes is not to say that there are identifiable goals, aims and ends. It is not even to say that there are wants. Since the distinction suggested between wants and wishes is not clearly upheld in ordinary usage where my wanting X is alternatively described as my having a wish or desire for X, the contrast it signifies might best be brought out by emphasizing that wishes, unlike wants, may be "idle". An idle wish is one held in the absence of any further beliefs about the means of effecting or acquiring the object or state of affairs desired. The object of a wish, but not of a want, may be an impossible or unprocurable state of affairs and may be believed to be so by the person even at the time at which the wish is correctly ascribed to him or her.[4] We may analyse despair, for example, in terms of a wish for the whole world to be other than the way it is, but to fail to include the belief that the world could not be changed, would be to miss the key element which distinguishes despair from other motives. Thus, the despairing person who takes her life may be said to wish that everything were different, but she could certainly not be said to want it, in the sense implied, nor to have the alteration of something in the world as her goal or end in acting.

This is what is at the crux of the inclination to distinguish motives of the sort described in (1) from those of the sort described in (2). While:

(1) all motives may be shown to comprise desires and wishes, it is only true

(2) of some motives that they comprise, in addition, beliefs of the particular, optimistic kind described in (2)(ii), about the relation between the behaviour in question and the object of the wish or desire. Some motives comprise no such beliefs, others comprise only negative beliefs about the relation, for example, the belief that since the object of the wish is impossible or unprocurable, no action could bring it about. And (2)(i) belongs with (2)(ii) rather than with (1), because the concepts it employs - "goal", "means", "ends" - themselves suggest the presence

of positive beliefs of the sort described in (2)(ii). To speak of an action's goal is to imply a certain degree of confidence in the possibility of achieving that goal. (It is often also, as a corollary, to imply a more detailed conception of the goal or end in question.) To conclude, it seems that motives of type (1): the more general sort characterized by emotions or attitudes, are not always reducible to the narrower sort (2): whose object is some end state, and sometimes together with a belief about the production of the end state.

Contrasting Kinds of Explanation

To give an account of some behaviour in terms of motives or motivation is, then, to offer a certain kind of explanation of that behaviour, an explanation which cites one or several of that agent's emotions, attitudes, purposes, goals, wishes or wants. We can contrast an explanation in terms of motives or motivation with a number of different sorts of explanation which might be offered in answer to the question: "Why did F do X?" when it is asked about some particular action. If we ask, "Why did F sign the petition?", for example, a number of different kinds of answer suggest themselves:

(i) She saw that P had signed and wanted her name to be associated with that of a famous radical.

(ii) She signed on a whim - she just felt like signing it.

(iii) She always signs petitions.

(iv) She had had too much to drink.

While not an exhaustive list of the ways of answering the question, (i)-(iv) illustrate four important kinds of answer or explanations which might be given. Notice that only (i) is an answer which may be said to cite a motive or to describe motivation. None of (ii), (iii) or (iv) could be said, as such, to describe motivated behaviour at all, although they all describe behaviour which may be said to be intentional. In the cases of (iii) and (iv), explanations in terms of motives are not excluded by the explanation given (although, as has been observed, neither (iii) nor (iv) is as such an example of such an explanation). For on receiving (iii) as an answer, we may ask an additional question: "Why does F always sign petitions?". If the answer is that she does so on principle because of a sentimental attachment to what

she regards as expressions of democratic behaviour, then we might say that she had a motive in so doing. But if the answer merely is that she has an unreasoned habit of signing petitions, then we should not. However, as it is stated, (iii) is an explanation in terms of habit and is to be contrasted with one in terms of motives. So too is the explanation given in (iv), which simply cites one of the causal factors explaining why F signed the petition.

It might be said in passing that a fuller explanation, which is not incompatible with that expressed in (iv), might have been that F wanted to impress P as being, herself, a radicial. By contrast, (ii) describes the sort of unreasoned want or whim with which we sometimes are affected and for which there are no further explanations, or at least none in terms of ordinary conscious motives and reasons. "She had had too much to drink" might be a plausible explanation of the whim but is merely a cause of the signing, as we have said. It is not a reason why the person might have done so in the sense in which that expression is being used here - the sense in which it is contrasted with "cause". (A more detailed account of the latter distinction is supplied in chapter three.)

Signs and Symptoms

In addition to contrasting the way in which a motive explains with other ways of explaining behaviour, we can contrast the way behaviour is regarded when it is seen as the outcome of a motive with the way it is regarded when it is seen as a sign or symptom, or as revealing or reflecting, another phenomenon. Thus, we can contrast the answer given in (i) ("She saw that P had signed and wanted her name to be associated with that of a famous radical") which is one expressing motivation, with:

(v) Her signing the petition was a sign of (was a symptom of, revealed, reflected) P's complete domination over F, since she did not really believe in that cause at all.

Wants and wishes, too, of course, can be described as being revealed or reflected in signs or symptoms, for example, "Her interest in your career is a sign of her own wish for more meaningful employment".

To say that something is revealed or reflected in some behaviour or that some behaviour is a sign or symptom of it (in the respect in which they are of interest to us here, these different expressions work in roughly the same way and may be used interchangeably), is to imply nothing in the way of an intentional relation between the two items. We say that F's going pale at the mention of Q's name revealed or was a sign of her feelings for him, when it is implied neither that she wished to go pale, nor that she was aware of doing so. (Notice that being a sign <u>of</u> is quite distinct from "being a sign <u>for</u>", which <u>does</u> imply the presence of intention.) To say that something is revealed or reflected in some behaviour or that some behaviour is a sign or symptom of it is to imply of the relation between the two items only that, directly or indirectly, they are part of the same causally connected chain of events. Although in the case described, it is understood that mention of Q's name was a sufficient condition of F's blanching and that there was a direct causal connection between the two, often when we speak of signs and symptoms such a direct link is not present. If we say that a red sky at night is a sign of a fine day to follow, we do not necessarily mean that any but indirect causal links connect fine days with red skies: the causal conditions sufficient to produce a red sky might simply be the same ones as those which produce a fine day.

While the relations of being a sign or symptom of, revealing or reflecting, may entail the presence of a causal connection, it is of course not true that <u>all</u> causal relations between items can be described as being relations of signifying, being a symptom of, revealing or reflecting. We should not say, for example, that the door's opening was a sign or symptom of my having turned the key. Nor, to take a case nearer to the relations between mental items and outward behaviour with which we are here concerned, should we naturally, usually and in general say that my assertion that "I believe that P" was a sign or symptom of my belief that P, or that it revealed or indicated that I believed that P.

There are differences between the four distinct relations to which we have referred of being a sign of, being a symptom of,

revealing and reflecting. But what they have in common, which is of interest to us here, is an emphasis on their making apparent or showing something which is less obvious or detectable, something whose presence is, for one reason or another, in some degree of doubt. For example, since inner physical states which often comprise the core of diseases are relatively hidden and inaccessible to ordinary observation, they are spoken of as revealed in signs and symptoms. And we speak naturally of a person's assertion of a belief as a sign of his or her holding that belief, or of a person's having some trait, only when:

(i) it is a matter about which the person is likely to be dishonest or reticent;

(ii) the person whose state or trait it is, characteristically is dishonest or reticent;

(iii) the mental state or trait is as likely to be unknown to the person whose state or trait it is as it would be to another observer.

One more point about ordinary conscious motivation must be made. It is important to notice that a person might have a motive for doing something and might act in such a way as to effect the satisfaction of that motive without being motivated <u>by</u> it. I might have a motive for refraining from answering the telephone on Tuesdays and I may refrain from answering the telephone on a particular Tuesday. Yet, my doing so may be unrelated to the above-mentioned motive: I may be in the bath at the time that it rings. So, merely knowing that a person has a motive for doing X, and has done X, is not sufficient to enable us to explain the action by reference to the motive.

In addition, when we speak of some belief or desire as motivating a person to do something, or as being that person's motive, we imply that a particular <u>kind</u> of relation obtains between the belief and desire and the resultant action or behaviour - a relationship which also can be expressed by saying that the person is motivated by, or by saying that the person does the action, or behaves, <u>out of</u> the motive. It is important to specify this relationship, since it is possible for motives to affect

behaviour or action in other ways as well.

Compare the following cases:

(i) P wishes to meet X, an expert in her field of research, in order to discuss her findings with him. She persuades her friend Y to introduce her to X. In this case, we can say that her seeking to acquire the introduction is motivated by her wish to discuss her findings with X. She arranges the introduction out of her desire.

(ii) The contrasting case is this one. P greatly wishes to meet X in order to discuss her findings. She stays late after work talking to her friend Y and voicing and discussing this very wish. Her staying late means that she misses the last train home and must hitch-hike. By chance, X picks her up in his car. They introduce themselves and discuss P's research. Notice that in this case, although P's desire to discuss her findings with X may be said to have been "effective" (in the sense of an accidental consequence) in securing her introduction to X, it would be incorrect to say here either that P's wish motivated her obtaining the introduction or that she obtained the introduction out of her wish.

Notice the similarity between the latter case, in which P's wish brought about the desired introduction without motivating it, and the case described earlier in which behaviour reveals or is a sign or symptom of, some mental state or states. In each case, the behaviour may be said to have been caused or produced by the mental state, although in neither can it be said to have been motivated by it.

Motives, Reasons, Intentions and Intentional Behaviour

In introducing and working with the distinctions between "motive", "reason", "intention" and "intentional action", it must be emphasized that I shall not respect all the connections, similarities and nuances of difference between these expressions which are reflected in ordinary language. The uses of these expressions in ordinary language are at once too subtle, too loose and too elusive to be of help in fixing the distinctions which I

wish to introduce here. Even in choosing to use "motive" as the central concept rather than "reason", I am to some extent acting arbitrarily, since the restriction here placed upon "motive" sometimes has been placed upon "reason".[5] I have chosen to use motives as a sub-set of reasons. All motives are reasons, but not all reasons are motives.

I wish to adhere to the convention to which some philosophers have subscribed, for example, Davidson[6], by which intentional action is action for which there are reasons. The class of intentional actions is not co-extensive with that of motivated behaviour. Some intentional action - that is, action for which there are reasons - is unmotivated, although all motivated behaviour is intentional. The limiting case of action for which there is a reason, but no motive, is an action done merely because the agent felt like doing it, or as we might say, had a whim to do it - action done "for no reason", as it is sometimes, rather misleadingly, put. It is misleading in that "for no reason" counts, in a weak way, as the reason in these cases. The case of a whim is not an instance of motivated behaviour at all. An unmotivated killing may have been an intentional action, however.

In the case of the latter kind, the agent has no motive in acting and merely acts because of a wish or desire to do so.[7] Notice that this is distinct from the case of motivated behaviour described earlier (see page 48) where the agent did X because of a wish to do Y, when doing X (saying thank you) was doing Y (being polite). In this case (there is no motive; the agent merely acts out of a desire to do so) the desire to do X is a sufficient condition (causal or otherwise) for the action's having taken place. Thus, in the sense in which to explain an action is to cite causally or otherwise necessary and sufficient conditions for its occurrence, we may be said to explain why the person did X by citing this desire.[8]

(Notice that the desire or wish to do X must be characterized as a desire to do X when "X" is a non-teleological description of the action undertaken. The qualification is necessary because the

reason for a motivated action sometimes is given in the form of an expression of a wish to bring about that action (that is, a wish to do X), when the action is described in terms of (one of) its desired effects. For example, the same set of physical movements may be described as my waving my hand or as my attracting the woman's attention. Yet only if an explanation of my action is couched in terms of a wish to perform the action as described in the former way would it be a whim, explaining (at least, as much as one can) an ummotivated action.

For the sake of simplicity, I shall omit the latter qualification (that "X" must be a non-teleological description) from my references to the wish or want to do X when X is the action done or intended, and I shall take it henceforth as understood.)

But, to return to the point made earlier, in citing the desire to do X, alone, we explain the agent's having done X in the more usual sense of throwing new light upon it only minimally, if at all, except inasmuch as we do so indirectly by ruling out other explanations. To say that the action was done merely on a whim is at least to convey that it was not done for a motive, for instance.

We can imagine a case in which over time the considerations comprising my reason for engaging in some action (for example, my reasons for avoiding buying aerosol sprays) might be forgotten, so that although I persist in engaging in that action, after a certain time my only reason for doing so might be my wanting to. My doing so is, in a sense, a habit. Such a case fits neatly into neither the category of behaviour done on a whim, nor that of motivated behaviour. We must say that the mere wish to avoid aerosol sprays counts as a whim here after the time at which the more elaborate reason, or motive, for doing it is forgotten. From that case we can draw the general conclusion that the same want to do X, when X is some action, can in the same person at time t_1 function as a necessary but not a sufficient condition for doing X, and hence would not be said to be part of the reason for doing X, and at time t_2 can function as a sufficient condition

for doing X and would be classed as that person's reason for acting.

Summary: Motives, Reasons and Intentions

Summing up the distinctions drawn so far, we can say that motives must be distinguished from whims and motivated actions from actions done on a whim. An action done on a whim is an action for which there is no motive. But, while it is an umotivated action, it is done for a reason. The reason for an action done on a whim of this sort is simply the whim itself, the wish or desire to do X, when X is the action undertaken. However, the reason for any motivated action is just the motive with which it is undertaken.

Whims and Wants

Since the use of "whim" employed here is somewhat more precise than the looseness of ordinary usage, let us fix its technical meaning:
A whim is a wish or want whose object is an apparently unmotivated action, an action undertaken merely for the sake of doing it. It is not a state we reach by any reasoning or thought process, so it can be characterized as being unconnected with any other of the agent's desires. It will sometimes be connected with other of the agent's beliefs, as in the case in which the want itself has an obvious mental cause, when it will be connected with beliefs about the cause itself; for example, if a glimpse of the sea produces in me a sudden whim, a wish to bathe, then my wanting to bathe is at least connected with any belief(s) I might have acquired as a result of my perception of the sea.

Let us note in passing that when an action has a "mental cause", there is reason to speak more strictly of the wish to take that action as alone being the product of the mental cause, since inasmuch as it is an action, and not a reaction, the action cannot strictly be said to be done 'for' a cause; it must have been done for a reason; the wish is the reason. And we have been stressing here that the mere whim to do X can be said to be a reason for my doing X (although only when my doing X is an

unmotivated action). Clearly, in the context in which my doing X is occasioned by a mental cause, we must in strictness say that the cause (my glimpse of the sea) is the mental cause of the whim (my wanting to bathe). And my whim is the reason for my action, if it occurs - my plunging into the water.

We have made mention already of the fact that in doing X intentionally, or intending to do X, whether or not there is a motive for doing it, it is understood that, in some sense, the agent wanted to do X. However, the 'want' here reflects certain consistent ambiguities. For, while it seems in one way to be correct to present wanting to do X as a necessary condition of any person's having done X intentionally, yet there seems also to be a sense in which we can speak of the agent as having done X intentionally without having wanted to do so. I am eager to finish my novel but I believe that I shall hurt my hostess's feelings if I do not talk to her. Her conversation is boring and repetitive, but I put my book away. In one way I must be said not to wish or want to talk to my hostess at all. My only wish is to finish my novel and avoid her dull chatter.

The simplest way to account for this ambiguity is to say that the term 'want' has a narrower and broader sense:
(1) In the narrow sense, wanting to do something is contrasted with other attitudes which it is possible for us to take towards our intentional and voluntary actions, such as the various kinds of feeling of duty, obligation, etc., so that either one acts from a want or wish to do so.(I want to do X) or out of a feeling that one ought to do so (I do not, or may not, want to do so, but I believe I must).
(2) In the other broader sense 'want' is used generically to cover all the various pro attitudes which might be expected to lead us to act.[9] Having an obligation or duty to do X is subsumed under the general "wanting to do X". Now it is clearly only in this broad sense of 'want' that we can say that wanting to do X is entailed in doing X intentionally, or intending to do X.

Do we provide an explanation in citing P's wanting to do X as an answer to the question "Why did P do X?" when X is a

motivated, intentional action? This question unpacks in a number of different ways. We must first ask: is wanting to do X ever the reason for doing X when X is a motivated action? And the answer to this question is that only in the case described, of unmotivated intentional action, would this want ever be cited as the reason or part of the reason for an action.

As I shall discuss at length in chapter three, the reason which is given for an action is unlikely to include all the beliefs and desires of the agent's which were necessary for effecting that action. Rather, it will include only those conditions most likely to be informative and explanatory. And when the action is otherwise motivated the want to to X will always, as we have seen, be uninformative. When there is a motive for doing X, that motive alone is the reason which would be given for doing X. Since giving a motive establishes that an action was intentional, and since an actions being intentional implies that it was an action which the agent wanted to undertake, then, in the sense of 'want' in this discussion, it would always be redundant to add the agent's having wanted to do X to the motive for doing X in giving the agent's reason for doing X.

This issue is apparently confused by our sometimes citing having wanted to do X as a reason for doing X even when the action is motivated and not merely done on a whim. Why did I order steak? The answer might be: "Because I felt like it (I wanted to order steak), and besides, I have an iron deficiency". But such a case can be handled in the following way. The want here may be said to function as a whim, and there are actually two distinguishable reasons for the action:

(i) my whim to order steak, and

(ii) the set of beliefs and desires concerning my iron deficiency.

Together these two distinct considerations have determined my choice. But the whim (wanting to order steak, or feeling like steak - 'want' in the narrow sense) must in this case be distinguished from "wanting to order steak" as the general want (that is, 'want' in the broad sense) ascribed on the basis of the fact that my ordering steak was intentional. The former want is

an unreasoned, unmotivated appetite, whereas the latter want is one for which there are reasons, viz. those given concerning my iron deficiency. The same reasons as those for ordering the steak are also the reasons for wanting to order it, in the latter sense. And, in fact, in reply to the question asked, an equally likely answer would have made this distinction clear. Why did I order steak? "I wanted to have steak today because I felt like it (I wanted to have steak), and because I have an iron deficiency." When it is put this way, the tendency to confuse "wanting steak" as a whim, with "wanting steak" as part of the intentional context, disappears.

Wants and desires, like beliefs, sometimes are states which we reach or form on the basis of deliberation. They are 'actions' of a sort, albeit mental ones, for which reasons can be given. At other times, they are passions which assail us and whose antecedents are unconnected with the rest of our mental life. Clearly, the sort of whims which we have been discussing are cases of the latter kind of want. Yet merely citing the object of the want will not enable us to tell whether it is a reasoned want or a whim. As the above discussion showed, the very same want, in the sense of identity conferred by sameness of object - in that case, the steak - can both be an unreasoned whim and a reasoned want.

Let us again stress that the use of 'whim' here is technical. It certainly conforms to ordinary usage but is more precisely defined. In ordinary usage we allow that both what I have called unmotivated wants (merely wanting to do X) and motivated wants (wanting to do X in order to get, see, etc. Y) are whims. While according to my definition whims can only be unreasoned wants and desires of the former kind, in the broader, non-technical sense of 'whim' in common usage, some wishes which might be described as 'whims' may be reasoned wants and desires. In the non-technical use of whim, my suddenly wanting to go to the zoo to see the monkeys might be described as a whim. In contrast, in the technical terminology employed here, my wanting to see the monkeys may be a whim, but my wanting to go to the zoo to see the monkeys would be, so to speak, a 'motivated want', and my acting upon it a motivated action. Because there is a motive underlying it, viz. my wanting to see the monkeys, my wish to go

to the zoo must count, according to the usage employed here, as a reasoned, motivated want and not as a whim.

Now it is evident that the various beliefs and desires comprising the reasons for doing X are just those considerations which lead us, in the case of 'motivated wants', to want to do X. But, while they are necessary and sufficient for wanting to do X, it is those considerations with wanting to do X which only together are jointly necessary and sufficient actually for doing X.

I shall discuss in Chapter Three the question of whether or not in citing the necessary and sufficient conditions in the case of intentional action we are citing ordinary causally necessary and sufficient conditions, a special sort of causally necessary and sufficient conditions, or some kind of non-causal necessary and sufficient conditions, as some action theorists would have it. It is sufficient here to say that, in speaking of necessary conditions I am implying that a relation of efficacy obtains between the conditions and the behaviour in question. These conditions actually bring about the action.

In answer to the question raised on page 59 of whether we explain in citing P's wanting to do X when X is a motivated, intentional action, we can now offer the following reply. Only when we regard an explanation as citing all the conditions, necessary and sufficient for bringing about an action is it perhaps true to say that we explain by mentioning the want to do X in answer to the question "Why did P do X?" when X is a motivated action. In contrast, we cannot be said to explain in the sense of giving a reason for any motivated action merely by citing the agent's desire or wish to act thus.

SECTION 2: UNCONSCIOUS MOTIVATION

In this section I use examples of Freud's case studies to show how different interpretations may be given of unconscious motivation. I then try to arrive at some formal definition of 'unconscious motivation'.

Freud's Theory. Our unconscious mental states exert their influence, according to Freud, throughout all of our outward behaviour as well as inwardly, in our dreams, out thoughts, feelings and desires. They are particularly evident and easily detectable in our errors, omissions, mistakes, inconsequential actions and gestures (fiddling with objects, touching our clothing, humming tunes, doodling, and so on), dreams, wit and humour and neurotic symptoms. Of these phenomena, it is especially 'errors' of action and memory and certain neurotic symptoms which deserve our attention, since, as I shall argue, only in Freud's discussion of these do we find what might be called a theory of unconscious motivation as distinct from claims about the revelation of unconscious mental states through signs and symptoms.

The distinction between motives, one the one hand, and signs and symptoms, on the other hand, has been introduced already. It may be illustrated in the following way: an unconscious wish to avoid X may be said to be revealed or reflected in a dream about Y or, to put it differently, dreaming about Y might be said to be a sign or symptom of unconsciously wishing to avoid X. But, unless it was further hypothesised that in dreaming about Y the subject was exhibiting a further unconscious wish, perhaps a wish to express his or her unconscious wish to avoid X in order thereby symbolically to fulfil it (a theory of motivation behind dreaming which has sometimes been attributed to Freud,[10] as it happens, but one which it is doubtful that Freud consistently and intentionally espoused[11]) then, while it would be true to say that the dream revealed or was a sign of the unconscious wish, it would not follow that the dream was unconsciously motivated.

Freud seems not consistently to have recognized this distinction as such, although at times, as in his contrast between the symptom's meaning and purpose, he may be supposed to have done so:

> "(All these symptoms) would be proof enough that the obsessive act is full of meaning; it seems to be a representation,a representation of that all-important scene. But we are not bound to stop at this semblance; if we investigate more closely the relation between the two situations we shall probably find out something more, the purpose of the obsessive act."
>
> A General Introduction to Psychoanalysis (Page 274)[12]

But more often the distinction appears to be collapsed, or emphasis is placed at one time on behaviour and dreams as reflecting one kind of relation, and at another, the other. When he speaks of neurotic symptoms and the aberrations evidenced in the "psychopathology of everyday life" as "wish fulfilments" and as attempts to fulfil unconscious wishes and desires,[13] Freud seems to imply that unconscious mental states always motivate and are not merely expressed in signs and indications.

At other times, Freud seems to be appealing to a kind of semantic theory, in which what might be seen as "grammar" and "syntax" of the unconscious are to be found in the characteristic features of primary process thinking and the psychoanalytic theory of symbolism. When he does so, Freud speaks of the expression of unconscious mental states in behaviour and of the meaning of behaviour being found in the underlying unconscious states. And the latter terminology would seem better to fit the concept of unconscious states being reflected or indicated in behaviour which is sign or symptom of them than it would the more purposive concept of behaviour being motivated by unconscious wishes and desires.

Because it is not obvious that Freud meant the same thing by these apparently different ways of expressing the relation between unconscious mental states and behaviour,[14] nor yet that he was consistent in his use of such descriptions, it seems best to look at the cases offered by him and to assess the application of each kind of description to them on its merits, ignoring the way in

which Freud himself might have characterized the relation.

Recently, it has been argued[15] that all behaviour which Freud describes as giving evidence of unconscious mental states (dreams, neurotic symptoms and the slips and errors of the "psychopathology of everyday life"), is better seen as an expression or sign of those states. (Although it is a somewhat more technical sense of that expression than the one which I employed in my earlier discussion of signs, it is a sense[16] which is included in that sense.) But I would argue that only some of the behavioural phenomena which Freud describes invites the latter interpretation. There are additional cases in which we seem to want to go on and say that the behaviour also, or perhaps instead, seems to be unconsciously motivated. (It is important to notice that the two interpretations are compatible. A piece of behaviour which is a symptom might be a sign of, and might also be motivated by, an unconscious mental state.)

Some of the behaviour Freud describes may be said to look to be motivated since the relation between the unconscious state(s) and the subsequent behaviour which they occasion has what might be regarded as the observable feature characterizing motivation or, to put the same point differently, the features characterizing the appearance of motivation. That is, the behaviour may be seen as appropriate to the desire(s) comprising the unconscious motive in question. And behaviour is appropriate to a motive when it can be seen as a possible way of fulfilling that motive and is what is sometimes described as "realistic" - that is, it might actually bring about the desired state of affairs.

Now clearly no behaviour can be seen as a possible way of fulfilling a wish or aim in the absence of an acceptance of at least some beliefs about the world. However obvious it may seem, the truth of the assertion that an action A is likely to or could bring about goal G, presupposes the truth of certain additional beliefs. But in the case where behaviour is "realistic" these additional beliefs are merely generally agreed upon beliefs shared by most people.

We can distinguish, both in animals and in humans, between behaviour which seems to be motivated and behaviour which is motivated. And, of course, behaviour may seem to be motivated and not be (the cat's 'fight' may be playful). Similarly, human behaviour may be motivated and yet might not seem to be. My sitting apparently idly in the sun may be motivated by my wish to irritate P, although even knowing of my motive, an outsider could not have guessed the deliberate nature of my pose. If animals likewise have behaviour which is motivated but does not seem to be, then we will not know of it. To be described as motivated, animal behaviour must seem to be motivated. It must at least[17] have the look of being appropriate, in something like the above sense, to the animal's goal.

In somewhat the same way, I would argue, we may be guided in applying the concept of unconscious motivation to human behaviour. At least when there is the appearance of motivation, we can hypothesise an unconscious motive.

Freud's theory of unconscious motivation is best seen, as I have indicated, in his explanations of neurotic symptoms and in his accounts of certain, but not all, of the mistakes and omissions which he classes as belonging to the "psychopathology of everyday life", particularly forgetting things, losing and 'accidentally' breaking and destroying objects, and making 'mistakes' in the performance of actions.

Everyday Life

In offering an explanation of the latter class of 'actions', hitherto dismissed as haphazard and inconsequential, Freud postulates the presence of an unconscious "counter will", a current of feeling opposing and countering the conscious wish or intention to carry out the said action. When the conscious intention is to do X, the unconscious counter will or counter wish is the wish not to do X. The counter will interferes with the carrying out of the consciously intended action by inhibiting or removing from consciousness the conscious intention. For example, if my conscious wish or intention is to turn down Huntington Street, and my unconscious wish or counter

will is not to turn down Huntington Street, then my counter will acts upon my conscious wish in such a way as to negate it. The effect of this negation in my conscious mind is that on reaching Huntington Street, I fail to remember to turn down it: my conscious wish is rendered ineffective by being at least temporarily extinguished.

Notice that if Freud had stopped with the introduction of the concept of the counter will, he would have been short of a theory of unconscious motivation, according to my terminology, although he would have been postulating unconscious intentions, or "reasons" - wants or wishes which would explain behaviour in the manner of "whims". However, he also postulates unconscious beliefs and wishes which explain both the counter will and the action flowing from it.

At one level of generality, the motive for, or purpose of, the counter will is usually to be found in Freud's so-called "pleasure principle" - the wish of tendency to avoid painful feelings:

> ".....the aversion on the part of memory against recalling anything connected with painful feelings that would revive the pain if it were recalled. In this tendency towards avoidance of pain from recollection or other mental processes, this flight of the mind from that which is unpleasant, we may perceive the ultimate purpose at work behind not merely the forgetting of names, but also many other errors, omissions, and mistakes."
>
> (Page 78, A General Introduction to Psychoanalysis, 1924)

But, at a less general level, the motives or purposes in terms of which the counter will is explained are differentiated. Of the counter will which comprises unconscious wishes to lose, mislay and destroy objects, for example, Freud remarks that:

> ".....what varies (in all the cases of unconscious wishes to lose things).....is the reason for the wish and the aim of it. One loses something if it has become damaged: if one has an impulse to replace it with a better; if one has ceased to care for it; if it came from someone with whom unpleasantness has arisen; or if it was acquired in circumstances that one no longer wishes to think of. Letting things fall, spoiling, or breaking things, serves the same tendency."
>
> (Pages 80-81, op.cit.)

Notice that Freud uses explicitly purposive language here, as he does in the case, which I shall discuss presently, of neurotic symptoms. It is not merely that such 'errors' as described reveal the unconscious counter will or wish (although they may be said to do that too), but that there are reasons or motives for the counter will - a particular motive relating to the context, for example, a wish to avoid being reminded of the person who gave one the object, together with a general motive, a wish to avoid all painful or unpleasant feelings, like the unpleasant feelings which the memory of the donor might be expected to evoke. And the 'action' in question, mislaying the object, can be seen as the successful carrying out of the unconscious 'intention' or motive.

The case of verbal slips (both spoken and written), that other class of 'errors' from the psychopathology of everyday life, is not so evidently one of full-blown unconscious motivation.[18] If we consider Freud's famous example of the President of Parliament who opened a session with the words, "Gentlemen, I declare a quorum present and herewith declare the session closed"[19] - we are inclined to suppose that the slip merely revealed the speaker's unconscious wish that the session were finished. This is because the utterance could not fulfil the wish it expresses.* The relation between unconscious mental states and verbal slips better fits the category we described of the production or causation of behaviour by unconscious mental states than it does that of unconscious motivation. Thus, verbal slips would better be described as being signs or symptoms of unconscious mental states.

It is only if we ascribe to the speaker an additional and secondary unconscious wish, viz. the wish to express his primary unconscious wish (for the session to be finished), that we seem tempted to appeal to the notion of motivation of the kind which seemed to be applicable in the case of other 'errors'. Notice that this move, while it entitles us to say that the slip is motivated, and not merely produced, by an unconscious mental state, involves quite a distinct claim from the one suggested earlier. Having the unconscious wish W must not be confused with having the unconscious wish to express the unconscious wish that W. Just as having the former by no means entails having the latter, so it would seem that

---

* Please see next page for footnote

one might have the latter in the absence of the former. At the level of ordinary conscious mental states, I might have a wish to express the belief that P without believing P; for example by lying. So, it will not follow from our being able to explain some behaviour by describing it as motivated by an unconscious wish to express the (unconscious) wish W that it will also be correct to describe that behaviour as produced by unconscious wish W.

Notice that the above objection to the notion of unconscious wishes to express unconscious mental states applies only in the case of verbal slips. We necessarily attribute to a person a wish to express P in describing his or her speech act or act of writing as intentional. But clearly no such an automatic description of an unconscious wish to express the state expressed or reflected in any non-verbal action or behaviour follows from our saying that it is intentional. It may be that in engaging in some non-verbal behaviour a person is gratifying an unconscious wish to express some belief or desire, but it may not. And, since this need not be so, it will be genuinely explanatory to introduce an account in terms of such an unconscious motive, even after it has been ascertained that the action in question is an intentional one. (More detailed discussion of intentions is found in Chapter Three.)

However, there is an additional complication with the case of wishes to express, which can be shown eventually to render this interpretation unhelpful. In the same way that if one acts (does X) intentionally, then it is a necessary feature of the situation that one also wanted to do X, so if A asserts proposition P verbally intentionally, whether in speech or in writing, then it seems to follow as a necessary consequence that A must be said to have had a wish to express P. But, since it is a necessary condition of asserting P intentionally, it is not helpful to cite A's wish to

---

* This case is somewhat complicated in that declaring the parliament open is a performative utterance. For, if it is regarded as doing something, and not merely as expressing a wish, then the example does, of course, indicate a clear case of unconscious motivation, since the speaker's motive can be seen to have been the closing of the session. But, in fact, its being a performative utterance is an aspect of the action in question which, for the purposes of this discussion, we can ignore.

express P as part or all of the explanation of why A asserted P. So, if we resort to a description in terms of an unconscious wish to express in an attempt to show that verbal slips may, too, be interpreted as cases of unconsciously motivated behaviour, as distinct from behaviour in which unconscious states are revealed, then we shall be disappointed, since even if such a wish were postulated, the whole account would fail to be one of motivation, due to its lack of explanatory force.

Neurotic Symptoms

In speaking of neurotic symptoms, Freud at times employs explicitly purposive locutions and describes cases suggesting unconscious motivation, in the sense described. For example, we find him speaking of a patient with an erotic attachment to her father, who had herself perceived that her illness would prevent her from marrying, and remarking that:

> "We may suspect that she became so ill in order to be unable to marry and so to remain with her father."
>
> (Page 285, General Introduction to Psychoanalysis)

And, of that part of the same patient's elaborate bedtime ritual which included keeping the door between her bedroom and her parents' room open, Freud observed that the underlying unconscious motive was a wish to stop sexual intercourse between her parents in order to avoid a rival in the form of a younger sibling.

In other cases, however, psychoanalytic descriptions of neurotic symptoms would seem better understood as the mere expressions or signs of unconscious states than as the carrying out of explicit unconscious wishes or 'intentions'. See, for example, Freud's assertion that an hysterical patient's throat irritation and coughing signified an unconscious fantasy of fellatio.[20] Because we cannot even imagine what behaviour would count as being apparently motivated by a fantasy, it is impossible to classify this case as either apparently motivated or not according to our earlier analysis of apparent motivation. And, if we reinterpret Freud's description and posit an unconscious wish on the patient's part to engage in the act of fellatio, then we must conclude that since suffering throat irritation and coughing could not bring about that

action's occurrence, the behaviour has not even the feature of being apparently motivated.

A puzzling case, though so common as to constitute the normal one in clinical psychology is that in which fulfilment of the unconscious wish is symbolic rather than real. Although, in the case described earlier, the patient's keeping the door open might be likely, in fact, to have ensured the fulfilment of her wish (that her parents refrain from sexual intercourse), most of her obsessional, bedtime rituals (and most neurotic symptoms in general, according to Freud)only involve wish fulfilment at a symbolic level. For example, Freud explains his patient's insistence that her bolster not touch the back of the bed by saying that the bolster symbolizes a woman, the bedstead a man, and:

> "She wished.....by magic ceremony, as it were, to keep man and woman apart; that is to say, to separate the parents and prevent intercourse from occurring."
>
> (Page 278, op. cit.)

Such symbolic carrying out of unconscious wishes seems to fall midway between the two categories we have been employing, that of unconsciously motivated behaviour and that of behaviour which merely reflects or is a sign of unconscious states. We can regard the bolster and bedstead ritual described as nothing more than the symbolic revelation or representation of the unconscious conflict, much the same as we would normally regard the symbolic expression of such a conflict in the work of a painter or writer[21] or in a dream (as long as we exclude the theory that we dream and create art for the purpose of symbolically fulfilling our unconscious wishes).

If, on the other hand, we treat the bolster ritual as unconsciously motivated, in the full sense, then we have two alternative interpretations available to us. First, we may admit the action as the fulfilment of an additional unconscious wish, concerning the separation of mother and father for the avoidance of sexual intercourse. The secondary wish is an unconscious wish to express the conflict. Or, secondly, we may go further and treat the separation of bolster and bedstead as an action which constitutes the explicit carrying out of the primary unconscious wish, an action

which in its own way (that is, symbolically), constitutes just as successful a fulfilment of the primary unconscious wish as did the more 'realistic' action of opening the bedroom door (when the degree of 'reality' accorded to a belief is a measure of the extent to which a belief concurs with generally agreed-upon beliefs about the world).

On the assumption implicit in the symbolic identification that the bolster and bedstead stand for and thus in a sense are the patient's mother and father, the ritual can be seen as an appropriate one. For an appropriate realization of a wish to avoid sexual union between a real flesh and blood man and woman would be to keep them physically apart. This case differs from a more 'realistic' action whose appropriateness would not rest on any assumptions of symbolic identification, only in respect of this 'unrealistic' belief that the bolster and bed are the particular man and wife.

(Notice that Freud merely states that the identification is between the bedstead and bolster and man and woman in general. While this identification may be one link in the chain of identification made by the patient, unless the further step is made to identify the symbolic figures (whether bedstead and bolster, or "men" and "women" in general) with the particular man and woman in question, viz. the patient's own parents, then this ascription of appropriateness cannot be said to apply. Attempting to separate any other man and woman than the two in question would not have the requisite likelihood of effecting the unconscious goal.)

Summary of Unconscious Motivation in Freud

We can sum up our analysis of the theory of unconscious motivation as it is reflected in Freudian writing by saying that there are strictly speaking two kinds of phenomena described by Freud which might be thought to indicate a theory of unconscious motivation, as distinct from an account of the way in which behaviour and mental states are signs and symptoms of, and may thus be said to be produced by, unconscious mental states. In the first place, there are cases where identifiable unconscious

motives or purposes to which we can refer in explaining the unconscious wish or counter will are present, and where the action which represents the carrying out or fulfilment of the underlying 'intention' or motivation is appropriate to, in the sense explicated earlier, and perhaps even likely to be successful in carrying out the unconscious wish. The action in these cases has at least the appearance of having been selected in the light of a 'realistic' comprehension of the world.

There are also the sort of cases (for example, the bolster case) described previously, in which on the assumption of a symbolic identification between the objects of the desire(s) and belief(s) comprising the motive or motives, and some other objects in the world, it is possible to see the behaviour in question as apparently motivated. On the assumption of that identification, then the behaviour is an appropriate way of achieving the goal or aim expressed in the motive.

In addition to these two kinds of case, the possibility was mentioned of an additional way in which behaviour may be seen to show unconscious motivation - not only that behaviour which we have already spoken of as unconsciously motivated, but also that behaviour so far dismissed as being merely a sign or symptom of unconscious mental states. We can postulate that the unconscious motive explaining a person's behaviour is an unconscious wish to express his or her various unconscious wishes, feelings and beliefs. In the light of this sort of unconscious motivation, not only seemingly or apparently motivated behaviour like that described above counts as unconsciously motivated - so too does behaviour, like a verbal slip, which does not otherwise have the earmark of apparently being motivated; behaviour which until now we have described merely as being a sign or indication produced by unconscious mental states.

The traditional Freudian conception of unconscious motivation has often been supposed more closely to resemble the model of motivation described on pages 47 and 48 under (1) in our earlier discussion of ordinary conscious motivation, where the behaviour

is explained in terms of some emotion or attitude like guilt, jealousy, remorse, etc., than the more purposive model described in (2) which implies that in addition to the presence of wants and desires there are certain beliefs about the relation between the behaviour in question and the objects of the desire(s) or wishe(es) involved in the motive. Freud himself more often spoke, moreover, in identifying the sources of their neurotic symptoms, of his patient's unconscious wishes than their unconscious 'intentions' or purposes. However, it is certainly possible, as the previous discussion of the notion of unconscious motivation indicates, to extend the notion of unconscious motivation to cover the purposive model described in (2) on page 48, in which motivated behaviour is viewed as directed towards goals or ends. In fact, it can be shown that even among Freud's own examples there is a whole spectrum, ranging from cases best understood on the model of (1) to those which do seem to presuppose the presence of those additional beliefs (whether conscious or unconscious) about the relation between what is desired, the goal or object of the wish, on the one hand, and what is done, on the other, which characterize the model of motivation described in (2). In the case of Freud's patient's bedtime ritual of opening the door between her own and her parents' room, for example, we seem entitled to suppose that either consciously or unconsciously the girl had the belief that doing so would be one way, and a likely way, of bringing about her wish that her parents not have sexual intercourse.

Kinds of Unconscious Motivation

In the most commonly described cases of unconsciously motivated behaviour, the person is said to be unconscious of at least one of the beliefs and desires or wishes which may be said to constitute the motive for his or her behaviour. Interestingly, there is an additional sort of case, the case of behaviour described as unconsciously motivated when neither the primary wish or desire comprising the motive, nor the primary beliefs about it are unconscious. Rather, what is unconscious in this case is a belief about the connection between the particular belief(s) and wish(es), want(s), etc., and the action in question - the motivation. One might have a general, consciously-held belief that doing D tends

to bring about X, and a general consciously held wish to bring about X, together with an unconsciously held belief that one's consciously held belief and desire or wish had produced one's behaviour in doing D. For example, I might be aware of my wish to attract X's attention and might consciously believe that doing Z, entering the room alone and after a pause, would be a good way of doing so. Yet even after contriving to do so by busying myself with something outside the door until the right moment, I might fail to realize that my entering the room at the time at which I did was motivated by my consciously held wish and belief. I might suppose it to have been motivated by the matter outside the door which I had allowed to take my attention and distract me from entering immediately and with a group. In this case, while I may be said to be unconsciously motivated, the only thing which is unconscious is my belief about the relation between my consciously held belief(s) and desire(s) and my behaviour. Strictly speaking, I am perhaps unconscious of my motivation rather than my motives: that is, I am unconscious of the connection between my belief(s) and my subsequent behaviour.

Freud's patient's efforts to prevent sexual intercourse between her parents by keeping the door adjoining the bedrooms open may be taken to count as an instance of such a case. Originally, as he points out, his patient had acted upon her then explicit and conscious intention in a straightforward way:

> "Years before the institution of her ritual, she had attempted to achieve this end (of preventing intercourse from occurring) by a more direct method. She had simulated fear, or had exploited a tendency to fear, so that the door between her bedroom and that of her parents should not be closed."
>
> (Page 278, op. cit.)

By the time of the analysis, her methods were more veiled. It is not made clear whether her fear was totally unconscious, or whether she was merely unconscious of the connection between her fear and her behaviour. But even if the former were true, then we can nevertheless postulate that at an earlier time, the latter state obtained. At some time between when she had employed the "more direct method", as Freud puts it, and was fully aware of her desire, and when she was thoroughly unconscious of it, we may conclude that there was a point at which she was still aware of her wish to

prevent her parents from having intercourse, and aware of her belief in the efficacy of her ploy, without being aware that she acted out of the motive comprising these two.

In my earlier discussion on the concept of motivation it was pointed out that a necessary condition for some motive's being said to motivate some behaviour is that the said behaviour be done from or out of the motive X, it is necessary not merely:
(i) that there exist a relation of efficacy between the two, but also
(ii) that the relation itself be of the required kind, that the action be undertaken out of the motive, and not merely produced or caused by it.

In the context of motivation which is unconscious the former qualification applies just as it was shown to do in the case of ordinary conscious motivation. A person might have an unconscious wish to do X, and might do X, without it being true that that person's unconscious wish was effective in any way in bringing about the behaviour, or actually motivated him or her. However, in the case of behaviour which is unconsciously motivated, applying the distinction described in (ii) above - between doing X because of some motive and doing X out of that motive - is different from applying it in the case of behaviour which is motivated in the ordinary way. For, as it is applied in the ordinary case, the distinction is a purely introspective, intuitive one, often supposed to be unfailingly evident to the agent. Where a person has a conscious motive, this is always evident to him or her. However, since in most of the cases of unconscious motivation described, the agent is conveyed as unaware of either or both of the belief(s) and desire(s) comprising the motive and of what we have been calling the motivation - that is, the connection between the motive and the subsequent behaviour - it is not a distinction which the agent is in a position to make. We must simply stipulate that in order for the concept of motivation correctly to be ascribed in such cases, it must be that this relationship does obtain, that is, that the person acts out of his or her unconscious motive(s).

Unconscious Motivation - A Definition

Taking into account the two different models of motivation discussed earlier in connection with ordinary, conscious motivation, we can sum up the various kinds of unconscious motivation by defining unconscious motivation in the following way:

Either (1) P has some motive M; that is, some emotion or attitude the very concept of which has built into it directedness towards certain sorts of behaviour, behaviour of which B is an instance. Consciously or unconsciously P believes B to be an instance of such behaviour; P does B out of motive M; M is unconscious, or P is unconscious of the connection between M and B.

Or (2) P has motive M, that is, some desire or wish D, and believes that doing B would be, or would be at least likely to be, a way of achieving the state of affairs S which is the object of D. At least one of P's desire(s) and belief(s) is unconscious or P is unconscious of the connection between M and B. P does B out of D.

One further distinction must be stressed here. As has often been remarked, it is possible to view what is in one sense the same action under different descriptions, some of which are descriptions in terms of intentions or motives which (that, is the intentions or motives) are not themselves correctly ascribed to the agent whose action it was. Thus, while engaged in some ordinary intentional action, such as conversing with P, I may also be correctly described as hurting Q's feelings. Yet, it may not be my intention to hurt Q's feelings. I may not even be aware of doing so.

Clearly, the situation is further complicated by the introdcution of the notion of intentions or motives which are unconscious. My hurting Q's feelings in conversing with P may be seen to be compatible with a number of different accounts. I may neither consciously nor unconsciously intend to hurt Q's feelings, and yet I may do so nevertheless. Now my doing so might merely be unwitting, or 'unconscious' in the non-Freudian sense introduced in Chapter One, that is, I may hurt Q's feelings unintentionally and I may be unaware of doing so. On the other hand, I may do so unintentionally while being aware of what I do. I may be indifferent to the fact that I do so. But I may also hurt Q's feelings intentionally. In this case, my intentions either might be conscious

or they might be unconscious, in the full, Freudian sense. I may have an unconscious wish to hurt Q's feelings of which I am not aware.

The class of behaviour which can be described as unconsciously motivated covers, of course, not only bodily actions of all kinds but also decisions, choices, even beliefs and desires inasmuch as they are actions or things done, rather than mere passions or things which happen to us.

To say that some behaviour was unconsciously motivated is not to say that its sole motivation was unconscious. Among the various kinds of explanation in terms of unconscious motivation, the one in which a particular unconscious motive, or unconscious motives, are presented as a complete explanation of the phenomenon is perhaps relatively rare. Nevertheless, there are some fairly well defined situations in which we are regarded as having been actuated solely by unconscious motives. Let us consider these first before turning to the more common case in which behaviour is explained by what we shall call mixed motivation, of which some is conscious and some unconscious.

Solely Unconsciously Motivated Behaviour

Clearly, we are solely moved by unconscious mental states in the case of what might be called acts of omission, in which we fail to remember something or fail to (remember to) do certain things. We could not have had any consciously held motive in failing to remember or to do something for the simple reason that had we had one the forgetting itself would to that extent have been unsuccessful. Thus, if I fail to (remember to) keep a dentist's appointment, then given that it is plausible to explain my behaviour in such a way at all, it can be said of me that my motivation was completely and solely unconscious. In addition, in the class of solely unconsciously motivated 'action' there are the kinds of 'mistakes', 'errors', 'accidents' and apparently chance and playful occurrences which Freud describes in The Psychopathology of Everyday Life.[22] These are things which we are said to do, which have not hitherto been regarded as

intentional actions. For example, Freud describes the case of a woman (Page 148, op. cit.) whose stumbling and falling on a heap of stones in the street be explained as an action unconsciously intended to administer punishment for guilt felt over having had an abortion. In the case of such 'accidents' a simple causal explanation would hitherto have been supposed appropriate. There are other cases, thinking up an apparently random number, playing and fiddling with objects about one's person, scribbling apparently aimlessly, etc. (see pages 151 and 195 in Freud, op. cit.) which have not hitherto been supposed to have required explanation at all; typically, they have been relegated to the class of chance occurrence.

Apart from the various unconsciously motivated things people do which I have mentioned so far, there is another class of phenomena often explained by reference to unconscious mental states. I refer to the class comprising the states which may be ascribed to us: our passions, feelings, moods, certain of our emotions, beliefs and desires, and our bodily states and conditions. Examples of the unconscious motivation of states of the kind I have just mentioned are common in Freudian theory, with its emphasis upon the effects of unconscious mental states on ordinary bodily states like those Freud investigated in his "hysterical" patients, on the one hand, and upon disturbances of feeling and emotion, on the other.

In proposing that the latter class may be said to be unconsciously motivated and not merely produced, caused or affected by unconscious mental states, Freud may be seen to be introducing a concept which is doubly at variance with common intuitions. For the class of mental and bodily states, while resembling certain 'actions' of the above kind ('mistakes', 'errors', 'omissions') in ordinarily being considered to be the sorts of things which are happenings, themselves beyond our voluntary control, have the added feature that they are not usually characterized, gramatically speaking, in the active voice. Thus there might be said to be some justification for distinguishing such states from omissions, errors and actions done or left undone due to failures of memory, since

the latter, but not the former cases are spoken of as things we do, even before any hypotheses are entertained about their being motivated actions.

A third class covers those things we do - not themselves actions, perhaps, although they may comprise some actions - unknowingly, even when we may have known or been aware of doing the thing(s) comprising doing them. For example, in engaging in an ordinary series of intentional actions, in conversing with P, X might be said to be doing a number of different things - A, B, C, D - or to be acting in a number of different ways. While aware of what she was doing, in one sense, or under one (or several) description(s), X might fail to be aware that in doing A-D she was also hurting P's feelings. In this case it would be said that her hurting P's feelings, although her doing none of A-D, was solely unconsciously motivated. Her conversing with P would presumably be said to be an action with mixed motivation, partly conscious and partly unconscious. Whether or not we choose to regard hurting P's feelings as something extra which X does, or as another way of describing what she does in conversing with P, that is, the questions of the counting rules to be applied in describing things done or actions, is unimportant for this discussion. Although in postulating that there is an unconscious motive for hurting P's feelings actually we may be supposed to have justified describing hurting P's feelings as a distinct thing which is done.*

What might be classified as a subgroup of this last class of phenomena is the class of cases comprising those where the agent can be said not to have known the motive for his or her having acted or behaved as he or she did at all. Within this class can be distinguished:

(i) those cases where the agent neither knew nor believed himself

---

* While it is true that in such cases our behaviour may be described as being entirely unconsciously motivated, as it was remarked earlier, we very often do things unknowingly or unwittingly in this way in the non-Freudian sense - hurting P's feelings may be done unconsciously without it also being true that unconsciously the person wanted to do so.

or herself to know the motive behind the action, appearing to act on a whim - wanting to do X but having no notion why; and (ii) those cases where the phenomenon of 'rationalization' is present. A rationalization, as I remarked earlier, is a false account of his or her motivation which the agent believes to be true and offers to explain some behaviour. For example, philosopher P's competitive and uncompromising position in a philosophical debate might stem solely from unconscious sibling rivalry. However, P might believe that her desire for the correct answer and the strongest argument dictates her behaviour. Although she may love the truth, this love may not actually be effective in bringing about her behaviour. So her analysis of the reason for her behaviour may merely be a rationalization. Her unconscious wish to outdo her hated sibling may have been, in fact, the sole condition effective in producing her behaviour.

In each of the types of case described above in (i) and (ii) - behaviour apparently done on a whim and behaviour accompanied by a rationalization - the agent has, of course, some conscious desires or wants which truly are ascribable to him or her. In the former case, there was the (apparent) whim itself. In the case involving the rationalization described above, P wanted to insist, demand, interrupt, shout, etc. But, in neither the case of action apparently undertaken on a whim, nor of action for which there was a rationalization (which may be seen to reduce to a case of action apparently undertaken on a whim), could the agent be said to have a conscious motive in acting; in both kinds of case the agent's behaviour was completely and solely unconsciously motivated.

Such cases deserve particular attention in the light of my discussion in an earlier section on wants, and the difference between behaviour done on a whim and hence unmotivated, and behaviour which is motivated. In that discussion it was concluded that the whim itself is the sole sufficient (mental) condition for an action undertaken on a whim. And the difference between wanting to do X when the want functions as a whim in an instance of doing X without a motive, and wanting to do X in an instance of doing X when X is a motivated action, was seen to be expressed in the kind

of role played by the want in each case. In the former case, the want was treated as a sufficient (mental) condition for the action's taking place and described as the reason which would be given for the action; in the latter case it was regarded as a necessary but not a sufficient condition of its taking place and it was not accorded the status of being even part of what would be cited as the reason for the action.

To summarise here, we must distinguish between:

1. Wanting to do X - when the want functions as a whim - as an instance of doing X without a motive.
2. Wanting to do X - as an instance of doing X - when X is a motivated action.

The role played by the want is different in (1) and (2):

In (1) the want is a sufficient (mental) condition for the action's taking place and is described as the reason for the action.

In (2) the want is a necessary, but not a sufficient, condition of the action's taking place and is not be be accorded the status of being even part of what would be cited as the reason for the action.

With the introduction of the concept of unconscious motivation, it would also appear necessary to re-examine my definition of a whim. It will be remembered from my earlier discussion that a whim was characterized in several ways in addition to the way just mentioned - that is, in terms of its functioning as the sole sufficient (mental) condition for an action's taking place. Whims: (i) were described as passive states in which we find ourselves, and (ii) were said to be unconnected with any other of the agent's desires.

Although whims are characterized by being unconnected with other consciously held desires, to admit that they spring from unconscious motives is to acknowledge their link with other mental states of the agent's, albeit that they are unconscious ones. Similarly, while it may be the case that all desires to act which are unconnected with any other consciously held desires present themselves, phenomenologically speaking as passive states; nevertheless in the cases in which there actually are unconscious motives from which

they stem, it is at least misleading to say that these desires to act are merely passions.

Now a theory of unconscious motivation such as the one I have been discussing encourages us to believe that much behaviour, hitherto dismissed as done on a whim, actually is motivated. But not all whims might be expected to be unconsciously motivated and so, even if we accept a theory of unconscious motivation, there still may be supposed to be cases of actions which are completely explained by reference to the whim on which they were done. The theory merely suggests that in fewer cases than was previously supposed does an apparent whim count as a sufficient condition or as a sufficient and complete reason for some action's having taken place. Or, to put it differently, in fewer cases than was previously supposed is an apparent whim really a whim.

Before leaving my discussion of the ways in which behaviour can be described as being solely unconsciously motivated, let us return briefly to the cases I discussed in Chapter One, in which the person became aware of his or her unconscious motive for acting at the time of the action and yet continued to act upon it, and, we must add, continued to be (unconsciously) motivated by it. May we say that cases such as these count as solely unconsciously motivated actions, or only that they are partially unconsciously motivated actions?

There is some temptation to insist upon the latter, for if the agent continued to pursue his or her activity after learning of its unconscious springs, we might suppose that alone to be evidence of some degree of conscious acceptance of the motive. If in learning of the unconscious reason for his wish to skip the cracks between the paving stones, the man in the earlier example continued to skip them, we might be inclined to attribute to him some conscious motive over and above his (conscious) wish to do so. Perhaps the same motive held now at a conscious level as well - which in the context of this particular example, and assuming the man otherwise to be reasonable, seems unlikely - or some wish

based upon a general policy of accepting and acting upon one's conscious wishes provided they were harmless and satisfying, or of "humouring one's id", or some such. However, such a demand seems to be implausible. There may sometimes be reason to count such cases as examples of solely unconsciously motivated behaviour. While accompanying conscious motives of the kind described might be expected sometimes, and perhaps even usually to enter the picture when an unconscious motive becomes conscious in the way described - there also seems to remain the possibility that the agent might be aware of the unconscious motive and yet might be unmoved by it (at a conscious level) and might continue to experience the wish to act as if it were an unmotivated whim. So it seems at least possible that cases such as these might also count as solely unconsciously motivated and not merely as partially so.

Partially Unconsciously Motivated Behaviour

Frequently, when we describe unconscious motivation, we do not take ourselves to be offering a complete explanation. Rather, we regard ourselves as giving some (perhaps the more salient) of a number of conditions, some of which will be ordinary, consciously held beliefs and desires, which were jointly sufficient to produce the behaviour being explained. The standard case of an explanation citing both conscious and unconscious motivation is familiar enough. Let us turn again to the example of the philosopher used earlier. Contrary to the explanation given when the example was introduced on page 81 we can say that both P's unconscious sibling rivalry and her conscious love of the truth and insistence upon the right answer together explain her competitive and uncompromising behaviour in a philosophical debate. Without her consciously held values P's unconscious motive would have been insufficient to produce her competitive behaviour, and without the unconscious motive her conscious wishes and desires would likewise have been insufficient to do so.

In addition, there is reference in the psychoanalytic literature to another way in which a number of different factors together produce behaviour, a way expressed in the notion of the "overdetermination" of behaviour. The term is Freud's own. His

attempts at a definition of it are compatible with the thesis that he used it merely to refer to the sort of causation described above, where several conditions, none of which alone would be sufficient to do so, operate together to bring about the effect.[23] However, at least in the interpretations of the concept made by later psychoanalysts, it has been used to indicate a phenomenon whereby although individually sufficient to do so, a number of different conditions together produce some behaviour. Thus, in contrast to the example just described, if the behaviour were overdetermined, either P's conscious wish or her unconscious attitude would alone have been sufficient conditions for her competitive and uncompromising behaviour, although in fact, they both contributed in bringing the behaviour about. The concept of overdetermination is taken to apply both to the case where the several determinants of behaviour are conscious and unconscious mental states, as in the example just described, and to the case where they are all or both unconscious ones.

The Thesis of Overdetermination

What can be made of this concept of overdetermination? There is an obvious and simple interpretation of the notion which allows that it is distinct from the one described earlier where several conditions, none of which alone would be sufficient to do so, jointly produce some effect, without suggesting that it introduces the puzzling and paradoxical idea that two different sets of sufficient conditions can actually be referred to to explain the occurrence of the same phenomenon. The force of this interpretation rests on a certain ambiguity or unclarity in the notion of some condition or set of conditions being sufficient to bring about some phenomenon. For "X was a sufficient condition for P" might be taken to mean either:

(i) that X would have been sufficient to produce P, or

(ii) that X actually produced P.

Because of such an ambiguity it is possible to emphasize (i) rather than (ii) and to interpret the thesis of overdeterminism of P by sufficient conditions X and Y as indicating that while either of the conditions X and Y were sufficient to produce P and might have done so, X actually produced P and Y did not; although Y would have done so had X not done so. This account at least differs from the earlier one in which X and Y were only jointly

sufficient to produce the phenomenon.

In ordinary explanations of non-psychological phenomena we do, in fact, resort to explanations of each of thetwo sorts mentioned. We might explain P's odd behaviour by saying that she had taken a glass of wine on an empty stomach, when it was understood that neither taking the glass of wine, nor having an empty stomach would alone have been sufficient to produce her state of intoxication, but that jointly they had been sufficient to do so. On the other hand, I might account for a shrivelled plant by saying that it gets no light in that corner and I forgot to water it, when it was understood that one or the other of these conditions actually must have been the cause of death, the lack of water drying up the roots or the lack of light stopping the process of photosynthesis. Although I might not be in a position to say which was the cause of death, nevertheless, if the plant had not died due to lack of water, it would have died from lack of light.

Can we also countenance the more radical interpretation of the thesis of overdetermination in the case of ordinary physical phenomena? Let us alter the latter example to fit the radical interpretation by saying that both the lack of water and the lack of light together killed the plant, although each would have been, and actually was, sufficient to do so. What do we mean? Such a proposition can most plausibly be taken to suggest that the plant was weakened by lack of proper photosynthesis so that a dryness of the roots which would not ordinarily have been sufficient to do so at that time was in this case sufficient to kill it. But as soon as we mention the temporal clause in a description of the effect it becomes apparent that this is just to say that the sort of causation first described explains what occurred; it is to say that while either condition would not have been sufficient alone to bring about the effect, viz. the death of the plant at time t_1, together they were sufficient to do so. So it is not a case of overdetermination at all.

However, it would be possible to describe a case so as to exclude the above interpretation and to make a description in terms of overdetermination the only one applicable. This can be

done simply by building the temporal qualification into the description, that is, by insisting that each of the two (sets of) conditions was sufficient to bring about the effect at time t_1. Such a description remains sounding awkward and unlikely; there seems to be a strong presumption in favour of the alternative interpretation. Yet, I would suggest, it is impossible to establish a good reason for the presumption and while that is so, it would seem that the possibility of complete overdetermination of the kind described must be allowed.

But, interestingly, it is only in ordinary explanations of non-psychological phenomena that any such oddity or implausibility attaches to the radical version of the thesis of overdetermination which allows that several conditions or sets of conditions sufficient to do so actually were operative in bringing about the effect. And, since we are interested in psychological rather than non-psychological explanations, we may ignore the peculiar features of the overdetermination of non-psychological phenomena discussed above. As soon as we turn to psychological explanations of ordinary conscious motivation the more radical version of the overdetermination thesis seems to lose any implausibility which may be thought to characterize it in the context of non-psychological explanations. We often cite several distinct reasons for choices we make and actions we take, with the implication both that any of them would have been sufficient to bring about the choice or action, and that all of them actually were influential in doing so. My reasons for going abroad in January might both be a wish to see my long lost friend and a desire to avoid the cold of a London winter. And it seems perfectly acceptable to say that, while either reason would have been sufficient to make me go, I may be going because of each.

Now it may be argued that the modality of intensity enters here and saves us from the conclusion that mental phenomena can be said to be overdetermined in the strong sense in cases like this. For although it is not always made explicity, every want or desire has a specific intensity, which can be established roughly by ranking ordinally the objects of all of our wants and desires. When X is the goal or object of one want and Y of another, and I

would prefer X over Y if presented with the possibility of having to choose between the two, then it can be said that my wanting X is more intense than my wanting Y. And it may be said that even if either one of the two different reasons cited in the example above is sufficient to make me want to go abroad in January, I want to go more if both reasons influence my decision than I should do if either one alone did. All that is necessary, it may be said, is to refine the question asked in order to receive a non-paradoxical answer about my motivation. If I am asked what made me want to do X as much as I did (when this is some quantifiable amount) then in citing both my reasons for going, I should merely be giving an explanation (of the type first described) in terms of conditions only jointly sufficient to explain the particular want or intention in question.

However, the latter answer will not suffice for two reasons. In the first place, since intensity comes in with the concept of wants and desires but has no place in the concept of actions, the above answer may be sufficient to dissolve the claim that wants are sometimes overdetermined in the radical sense, but it does nothing to show that actions, which are undertaken for the same reasons, are not. My going abroad because of the two reasons given still would seem to be overdetermined, even if my wanting to go were not. In the second place, at least on an introspectively gauged scale of intensity (and the complications and hazards involved in introducing any other kind at this stage of our knowledge of brain science are obvious) it would seem that a person might want to do X because of some reason, and want to do it with maximum intensity (which might be measured as: not being able to imagine wanting anything more than it), might subsequently discover an additional compelling reason for doing X, and might then want to do X with the same intensity but for both reasons.

Whether the difference between the way we regard explanations of mental or psychological phenomena and the way we regard explanations of non-psychological phenomena is merely a contingent one, resting on our present relative ignorance with respect to the nature and constituents of mental phenomena and the complex connections between beliefs and desires and human action - or whether

it reflects something deeper about the nature of psychological phenomena,[24] is not a question with which we can here concern ourselves. It is sufficient to say that, for whatever reason, this difference is apparent: not even the presumption that for any particular occurrence only one sufficient condition is present prevails in the realm of the psychological. And since - other than their being concerned with unconscious rather than conscious states - explanations in terms of unconscious motivation and unconscious reasons for action function in a way which is parallel to the way that explanations in terms of ordinary conscious motivation and ordinary conscious reasons for actions and beliefs do, we seem free to admit the possibility that explanations in terms of unconscious motivation can be of the form expressed in what we have called the radical form of the overdetermination thesis. It is possible to say of some behaviour that it is over-determined in the sense that several conditions or sets of conditions, some or all of which are unconscious mental states, together brought it about, when each was sufficient to do so.

Thus, in summary, we must allow that there are three different kinds of explanation, corresponding to the three ways in which unconscious mental states can be said to motivate. There are:

(i) explanations citing one condition or set of conditions (some of which are unconscious mental states) which was sufficient to bring about some behaviour, and did, in fact, do so;*

(ii) explanations citing several conditions or sets of conditions, some or all of which are unconscious motives, which while individually insufficient to do so, jointly were sufficient to bring about some behaviour, and did bring it about;

(iii) explanations in terms of overdetermination, where several conditions or sets of conditions, some or all of which were unconscious mental states, individually were sufficient to bring about the said motivation, and jointly did so.

---

* Notice that (i) includes the kind of explanation described on page 85, where there are several conditions or sets of conditions each one of which would have been sufficient to bring about the effect, and one of which did so.

Notes and References

1 Anscombe makes what I take to be the same point in her discussion on what she called backward-looking motives, (Intention, Section 13):

> ".....if I kill a man as an act of revenge.... revenge is my object; but revenge is not some further thing obtained by killing him, it is rather that killing him is revenge."
>
> (Page 20)

But, Anscombe's account is incomplete in its suggestion that this characteristic is limited to the class of backward-looking motives (revenge, gratitude, remorse and pity for something specific), as my example about politeness illustrates. As well as backward-looking motives, some of what Anscombe calls motives-in-general have as their object merely the behaviour in question as seen under some different description.

2 Kenny (Action, Emotion and Will, Routledge and Kegan Paul, London, 1963, Chapter IV) makes a move somewhat akin to this in claiming that whenever we ascribe a motive, a certain pattern is present. There is (1) a state of affairs of which the subject disapproves; (2) an action, and (3) a state of affairs of which the subject disapproves, such that A is P (when P = the state of affairs of which the subject disapproves), then A acts, then A is Q (when Q = a state of affairs of which the subject approves). And, whenever someone's action has a motive:

> ".....if we are to understand his action, we must know how he is better off (or thinks he is), or how the world is a better place (or is thought by him to be) as a result of what he does. He may explain this either by showing the badness of the preceding state of affairs or by showing the goodness of the (expected) succeeding state of affairs."
>
> (Page 91, op. cit.)

However, Kenny's "disapproves of" seems rather too strong. If generosity is my motive for giving to a friend it may not be that I disapprove of my prior fiscal state, (I may look upon it with approval as allowing me my act of generosity) but merely that I approve of the state in which I am able to place my friend. It would perhaps be better to say that there is a state of affairs which is perceived by the subject as having some advantage over the prevailing one.

3 It seems that we can say this even when, as is the case in some cultures, the act of revenge merely reflects a formalized reaction. Although not perhaps so strongly felt and well-considered, such a wish is equally attributable to the Sicilian taking ritual revenge as to the avenger in our society.

4 The more the thing (object) is envisaged as a likelihood, as Anscombe puts it, "the more wishing turns into wanting."

(Intention, Page 67)

5 Others have restricted the expression 'reason' in the way in which I here restrict 'motive'.

6 Davidson, 'Actions, Reasons, and Causes', J.Phil., 1963.

7 It sometimes has been insisted that, as Kenny puts it (op. cit., Page 71):

> ".....with most things it doesn't make sense to say I just wanted to'. We have to specify a desirability characterization."

However, the latter claim would seem to be misleading. We certainly accept "I just wanted to" in the sense of "I just felt like doing it", as a sensible account of why we indulged the better known appetites associated with food, drink and sex. But these, Kenny would argue, may be said to be a special case, since there is a well-known teleological explanation which is tacitly assumed when these cases are given and which serves to make explanations in terms of them intelligible. However, it seems to me that we also may sensibly cite just having felt like flying kites, or taking a walk or just having wanted to talk to someone after a day of solitude (not because of such a day in the sense which indicates the presence of a reason rather than a brute cause). Rational these wants may not be, but they are surely possible.

8 Cf. Anscombe -"'I want to' is not an explanation of what a man is doing". (Page 90, op. cit.)

9 Although the details of this discussion need not concern us here, it is interesting to note in passing that the contrast which has seemed in philosophical discussion to invite the use of the broader sense of the term 'want', is roughly that drawn between being free and being unfree or determined. Being free is often characterized as being free to do what one wants to do - from which it is taken to follow that acting freely involves doing X out of a wish to do it. Thus, for every action freely undertaken there will be a wish or want, in the broad sense of that expression, to do it.

10 On the basis of the following remarks in The Interpretation Of Dreams (1900) (Page references are to the Avon Books edition, N.Y., 1965):

> "(Dreams).....are psychical phenomena of complete validity the fulfilment of wishes; they can be inserted into the chain of intelligible waking mental acts; they are constructed by a highly complicated activity of the mind."

(Page 155)

and ".....in the meantime, the meaning of the dream was born in upon me. I became aware of an intention which was carried into effect by the dream and which must have been my motive for dreaming it."

(Page 151)

See also the case referred to by Peter Alexander in the symposium: "Wishes, Symptoms and Actions" by Frank Cioffi and Peter Alexander.

11 As is argued in a detailed, thorough account of the meaning of 'meaning' (Sinn) in Freud's dream theory by Shope ("Freud's Concepts of Meaning", in Psychoanalysis and Contemporary Science, Vol. 2, 1972).

12 All page references to Pocket Books edition, New York, 1971.

13 See for example his remarks in The Interpretation of Dreams, that:

> "The theory governing all psychoneurotic symptoms culminates in a single proposition which asserts that they too are to be regarded as fulfilments of unconscious wishes."
>
> (Page 608)

and "The pathological actions of everyday life (also) involve the carrying out of an unconscious intention."

(Page 610)

14 Which makes attempts to ferret out the view correctly ascribed to Freud, like Shope's, op.cit., somewhat misguided, although, of course, interpretations of Freud in the light of one position or the other, like that of Lacan, described, are not.

15 For example, by Lacan. The most accessible account of this interpretation is that given by Ricoeur, in Freud and Philosophy, translated by Savage, (Yale University Press, 1970).

16 Ricoeur, for instance, emphasizes that although he is indebted to Cassirer for the general notion of a sign or symbol, his own definition is much narrower than Cassirer's: he wishes to restrict the notion of a "symbol" to those signs whose:

> "intentional texture calls for a reading of another meaning in the first, literal, and immediate meaning.... (Signs which) share the peculiarity of designating an indirect meaning in and through a direct meaning, and thus, call for something like deciphering."
>
> (Page 12, op. cit.)

17 Since the number of motives we are prepared ever to ascribe to animals is severely limited, the latter condition's being met would not alone free us to describe animal behaviour as motivated. Although the dog's behaviour may give the appearance of being motivated by, for example, a spirit of detached scientific interest, we should not allow that it was correctly described as motivated thus.

18 Although it pertains only indirectly to our discussion, since verbal slips are not central cases of behaviour which is unconsciously motivated, it is interesting to note that Freud analyses the mechanism producing verbal slips differently from that producing other 'errors'. Unlike the case of error he postulates not two but four opposing forces, or 'intentions' in explaining slips of tongue and pen, the initial conscious intention to say or write X, together with an intention not to say or write not-X, and the "interfering tendency" or counter will which actually may be momentarily conscious in some cases, as he points out, a wish to say or write not-X, and one not to say X. An attempt is thus always made to stop the interfering tendency, it is 'forced back' and:

> "The speaker has determined not to convert the idea into speech, and then it happens that he makes a slip of the tongue; that is to say, the tendency which is debarred from expression exerts itself against his will and gains utterance, either by altering the expression of the intention permitted by him, or by mingling with it, or actually by getting itself in place of it."
>
> (Pages 68-69, op. cit.)

19 Page 44-45, op. cit.

20 "This hysterical girl.....who had heard of the occurrence of such a method of sexual intercourse.....developed an unconscious fantasy of this sort and (gave) it expression by an irritation in her throat and coughing."

> "Analysis of a Case of Hysteria"
> Page 63, Vol. III, Collected Papers of Sigmund Freud.

21 It is perhaps important here to forestall a possible misinterpretation. It has sometimes been suggested that there is a way of ascribing intentions to behaviour and to art, which enables us at least to remain neutral as to (i) the actual presence of an intention, and (ii) any relation between such a postulated intention and the phenomenon (behaviour or art). It is sometimes put more strongly as a categorical denial of the presence of any such intention. Gustafson ("Unconscious Intentions", Philosophy, 1973) argues this on the basis of a putative distinction between "intentions in" and "intentions with which" - claiming that unconscious intentions, like intentions in art, are to be identified with the former, rather than the latter; unconscious

intentions, as "intentions in", are not actual intentions present in the person whose behaviour they are introduced to explain, any more than the so-called "artist's intention's" are actual intentions in the artist's mind. I wish firmly to dissociate my position from any such as Gustafson's. Apart from the fact that in the cases of what I have called unconscious motivation, as distinct from the revelation of unconscious states, I believe that the motives or "intentions" are present and are effective in bringing about the behaviour, I have an added difference with the proponents of the above view. I do not see that the analogy between intentions in art and unconscious intentions holds even in the cases I am describing, not as unconscious motivation, but as unconscious revelation. For, to say that an unconscious mental state is revealed in certain behaviour, or that the behaviour is a sign of some unconscious mental state, is to say something much more like: "Her going pale at the mention of his name revealed that she still cares for him", or "His mentioning the cash straight after that shows that he has guessed our secret" - than it is to say "The purpose of that line was to draw our eyes back towards the centre of the picture" or "The simile is intended to heighten the starkness of the initial description".

In the case of the artist's intentions, the question of the relation between any putative intention in the artist's mind and the effect achieved; that is, the line or the simile, in the above examples, is one which rightly does not, and certainly should not enter; (the "fallacy of the artist's intentions" is just the fallacy of supposing it to be a legitimate question). But, in the case of the unconscious or conscious states like those described being revealed in behaviour, since there are actual states postulated it is quite legitimate to suppose a relation of efficacy to obtain between them and the subsequent behaviour: a simple causal relation. The significance of the concept of behaviour revealing states, whether mental or physical, lies in the fact that the revelation itself is quite unintended by the agent, and whether or not the agent is aware of the significance of his or her behaviour in this respect at all, is beside the point. Thus, we may equally say that in this case the unconscious mental states "produce" rather than motivate the behaviour - to anticipate a distinction later drawn. But, it does so in these cases in a particularly revealing way: the behaviour is redolent with meaning.

So, in the case of the unconscious states of artists is a special one; not all "artist's intentions" are unconscious states revealing themselves in the work and whether or not any particular intention was of this kind would be uninteresting, aesthetically.

22 I have distinguished (1) and (2), unconsciously motivated errors of omission and comission, but Freud does not: he has a general term "Fehlleistungent", usually translated as "parapraxis", to cover both kinds of error.

23 For example, from Studies in Hysteria (1893):

> ".....the principal feature in the aeteology of neuroses - that their genesis is as a rule overdetermined, that several factors must come together to produce this result."
>
> (Page 263)

and, from "A Reply to Criticisms of my Paper on Anxiety Neurosis" (1895) in the Standard Edition of The Complete Psychological Works, London, Hogarth Press and Institute of Psychoanalysis, 1955-1964:

> "As a rule, the neuroses are overdetermined; that is to say, several factors operate together in their aeteology."
>
> (Page 131)

24 As some psychoanalysts have wished to suggest. See especially Guntrip, (Personality Structure and Human Interaction, 1961, Hogarth Press, London); for example, Page 155:

> "Psychoanalysis provides a new type of model for personality as a complex of various psychic levels and structures that enables the phenomena of personal living - i.e., those of conscious and unconscious conflict - to be explained on the basis of overdeterminism and plurality of causes, 'cause' being no longer understood in the physical sense."

CHAPTER THREE

Summary

The logical connection often thought to obtain between the notion of having a reason or intention, on the one hand, and being able to become aware or conscious of it, on the other, is now examined. I argue that the notion of having a reason or an intention is not logically tied to being able to become aware or conscious of that reason or intention by a simple act of introspection. I deal largely with reasons and consciousness, but provide some discussion specific to intentions at the end of the chapter. I show that there might be an indeterminate, rather than finite, number of beliefs and desires which might be reasons for a particular action and that, therefore, the agent will always be unable to give all his reasons for his actions.

I then distinguish between reasons and causes. I allow that while reasons may be causes, they are a distinguishable and a philosophically interesting subset of causes. I argue that, unlike behaviour which is merely caused, an action done for a reason is one where a wish or desire on the part of the agent is one of the causally necessary conditions for action. I further distinguish between 'his or her reason' and 'the reason' for an action. It is then argued that there can be no conceptual impossibility in suggesting that beliefs and desires which are unconscious - in either the sense of unknown to the agent or that of known only in a qualitatively distinguishable way - are reasons. It is argued that: (a) in the absence of a way of distinguishing so called 'primary reasons' we must allow the possibility that no person knows all of those reasons which are his or her reasons for an action, and (b) in certain special but everyday situations, which differ from ordinary ones only in the special respect concerned, we allow that reasons are present of which the agent is unaware and unable to become aware. The inconclusive nature of conceptual arguments of this kind is discussed, but I conclude ultimately that there is no reason to doubt the conceptual possibility of unconscious reasons and intentions.

Finally, the further possibility of unconscious rationality is then introduced and is shown to arise out of the suggestion of Toulmin and Flew that unconscious mental states function in the same way that ordinary conscious motives and reasons do. Toulmin and Flew were concerned to show that while Freud was significantly right when he introduced the concepts of unconscious motives, purposes and intentions for neurotic behaviour; he was profoundly wrong when, at a theoretical level, he insisted that these unconscious motives were actually inferred efficient causes. I claim, with them, that Freud should not be interpreted as introducing efficient causes, but as introducing motive-type explanations. I conclude this chapter with an examination of the claim that unconsciously motivated behaviour can be regarded as rational inasmuch as it is a rational response to the situation as perceived from the point of view of the agent's unconscious state of mind. I argue that there are two quite different types of case: those where behaviour is produced by unconscious mental states and those where behaviour is unconsciously motivated. In the latter case behaviour is sometimes rational and sometimes irrational. This leads us on to Chapter Four's discussion of rationality.

* * * * * * * *

In discussing the concept of unconscious motivation I drew attention to a distinction between behaviour which is merely produced or caused by unconscious beliefs and desires and behaviour which can be said to be motivated by them. However, it has been argued by some philosophers that the very notion of behaviour which is at once occasioned by unconscious mental states and, in some broad sense, intentional, is contradictory or, as it has sometimes been put, that the very notion of an unconscious reason or an unconscious intention is one which can be shown to be incoherent.

This is not, it will be noted, a claim about the incoherence of the notion of unconscious mental states in general. The traditional objection to the idea of something's being at once mental or psychological and unconscious, to which Freud addressed himself[1] is no longer encountered. But, while the mind may not now be regarded as coextensive with consciousness, the idea of having a reason or an intention is still often thought to be firmly tied to that of being able to become aware or conscious of that reason or intention by a simple act of introspection. For example, we find Hampshire[2] describing intention as:

> ".....the one concept that ought to be preserved free from any taintof the less than conscious."

Clearly, as long as this view were to prevail, the category which I have distinguished of unconsciously motivated behaviour would have be be subsumed under the category, against which it was there juxtaposed, of behaviour which is merely produced by unconscious mental states.

Now while intentional behaviour is behaviour done for reasons, it may be argued that reasons and intentions are distinguishable. (In particular, while all intentions are perhaps reasons, not all reasons may be intentions.) Thus I wish to maintain a separation between:

1. The claim that there cannot be unconscious reasons, and
2. The claim that the concept of an unconscious intention is incoherent.

And I shall concentrate first on the former claim, returning briefly to the question of the coherence of unconscious intentions only at the end of this chapter.

Unconsciousness and Reasons

It was observed in Chapter One that the notion of immediacy or non-observationality has been introduced to mark, among others, the distinction between two supposedly distinct sorts of explanatory entity, both of which are appealed to in explanations of human behaviour: "reasons" and "causes". And it was shown that, however it is construed, "immediate" or "non-observational" knowledge, which usually is associated with our knowledge of the reasons for our behaviour, cannot be mapped onto the distinction between conscious and unconscious states. It cannot be shown that our knowledge of our unconscious mental states is always in the required sense mediate, or observational, while knowledge of our ordinary conscious mental states is immediate or non-observational.

However, there is a second way in which the distinction between reasons and causes may be supposed to correspond with that between the sensse of 'conscious' and 'unconscious' described in Chapter One. It follows from our definition of unconscious mental states that ordinary conscious mental states are definable as those of our mental states of which we are aware or of which we can become aware, merely by an act of attention or introspection. And, being items of which we either are, or necessarily can become aware or conscious,attention or introspection has been introduced[3] as a criterion by which the reasons for actions might be distinguished from their causes. Because of this, it has sometimes been assumed that only what we have been calling ordinary conscious mental states could ever function in explanations as reasons for actions and that a criterion of the sort we used in Chapter One to distinguish conscious mental states could also be used to distinguish reasons.

A corollary of the latter view, of course, is that unconscious mental states could never be said to function as reasons for actions and that the accounts we give of human behaviour in terms of unconscious mental states are best regarded as part of a causal explanatory framework. It is roughly this view which is proposed by philosophers attempting to dismiss the possibility

of describing unconsciously motivated behaviour as in any way rational.

My main aim in this chapter is to investigate the supposed distinction between reasons and causes and the distinction between reasons explanations and causal ones. In particular, I shall investigate what I shall from henceforth call the awareness condition for reasons - the assertion that our being aware or conscious of something might be said to be a necessary condition of its being a reason for an action as distinct from its cause - with a view to evaluating and defending (with certain qualifications) the conceptual possibility of unconscious reasons. I centre the discussion initially on the purported distinction between reasons and causes and reasons explanations and causal ones but, first, some general observations about the nature of reasons for actions are called for.

Reasons

A reason for an action is a particular kind of explanation. In baldly making this claim, no reification of reasons as mental entities is implied. To offer a reason is to offer certain beliefs and desires of the agent's (or other of the agent's propositional attitudes - fears, expectations, hopes conjectures, etc. - themselves reducible to beliefs and desires) which pertain to the way the action in question is conceived. How can the beliefs and desires comprising reasons foractions be characterized? In answering this question I shall make use of some of the terminology, although not the analysis of reasons, introduced by other philosophers.

Giving a reason for an action involves giving what has been called (originally by Anscombe) a "desirability characterization",[4] that is, citing some way(s) in which the action undertaken was seen as desirable, appealing or necessary. And, on this view, the role played by the desirability characterization in having reasons might be put thus: having a reason for doing X entails having a pro attitude,[5] when having a pro attitude entails wanting to do the action towards which the attitude is directed in the broad sense of the term 'want' which covers having an inclination

based upon any kind of positive evaluation (wishes, whims, feelings of obligation, etc.), either towards the particular action in question or towards some class of actions of which the particular action is (believed to be) a member.

Notice that this characterization allows the inclusion of reasons varying in their degree of explanatory power. Compare two different kinds of answer which might be offered in response to a demand for a reason for making a certain remark to X: such as, "I thought you called yourself a pianist. I would have thought that all but novice pianists would have found this piece easy to play."

Reason 1: Because I felt like doing it.
Reason 2: I like deflating braggards and X has been boasting all evening.

A reason of the kind given in the first example would never be more than minimally explanatory of the action in question, while Reason 2 may be highly explanatory. Nevertheless, an answer of the form given in Reason 1 does fulfil the requirements expressed above: "feeling like doing X" is a case of a pro attitude directed towards a particular action. The significance of the latter point, and of the kind of reason expressed in Reason 1, was explored in the last chapter where I discussed the relation between reasons and motives.

Notice that there may be more than one reason for any particular action. Reasons 1 and 2 above might each function as reasons for the same action. We may make the same point by saying that it is possible for the reason for an action to comprise more than one desirability characterization. Notice also that the requirements for something's counting as a reason for an action designate the minimal conditions which must be met by a person to be said to have a reason in acting, although not necessarily those conditions required in order for an adequate answer to the question "Why did X do that?" to count as a reason. The reason-giving context, especially the relative knowledge enjoyed by the inquirer and agent, greatly affects what is offered as the reason for an action. It is rarely necessary in giving a reason for an action to cite both the general pro attitude and the belief that

the action in question is a member of the class towards which that attitude is directed, for example. According to the context one or the other usually is understood.

On the other hand, sometimes the context demands that at least in giving a reason for an action we supply additional beliefs and desires or desirability characterizations other than those comprising what we have so far described as the minimal conditions for having a reason. These include:

(i) additional beliefs about the particular action which show it to be an instance of the class of actions towards which the pro attitude is directed;

For example, in the case of Reason 2 above, it may be necessary to add certain facts Y showing that the particular remark had sarcastic import for him. For instance, if the remark was made about pianists, an added feature of the reason for acting may be the belief that

Reason 3: X is a bad pianist.

There are also:

(ii) additional beliefs and pro attitudes directed towards the general pro attitude expressed in the primary reason;

For example:

Reason 4: Boasting and bombast deserve withering sarcasm because they ruin parties.

We appear to have some choice in classifying statements such as those in Reasons 3 and 4 above. They might be treated as additional reasons for the action or as secondary aspects of the reason for it. Alternatively, however, they may be dismissed as not strictly part of the reason for the action at all. This decision involves the general one of how reasons are to be counted, and what constitutes the having of a complete reason. In choosing to allow Reasons 3 and 4 as parts, albeit secondary ones, of the reason for the action in question, we are employing a fairly generous counting principle. In choosing instead to restrict the reason for the action to what is encompassed in a reason in the sense defined above we are appealing to a more economical one.

Now there are certain obvious considerations which might be supposed to be pertinent to a decision of this nature. Both (i) and (ii) may be said to be reasons not for the action (making the remark, in this case), but instead for (parts of) the reason (for the action) itself. Reason 4 above might be said strictly to be the reason why the sarcasm was deserved, rather than the reason why the remark was made. And similarly, Reason 3 may be seen as the reason why the sarcastic remark had the import which it did, rather than the reason for the action. Thus, it is arguable that our additional reasons for the action in question are part of the reasons for other actions or beliefs.[6]

And, in the interests of avoiding a regress of reasons, for instance, we might be expected to prefer the more economical counting which limits the reason(s) for any particular action to the reason for it and relegates considerations such as those expressed in (i) and (ii) to the status of reasons for associated beliefs, actions and evaluations.

I shall argue for a generous counting principle, but one consequence of a counting method dictated by a less economical principle would, of course, be that such a method would allow that reasons for actions might comprise an indeterminate number of beliefs and desires instead of a finite one, and thus that the agent might be unable to give all of his or her reasons for any of his or her actions. Such a consequence may well be thought to be an undesirable one (although as I shall argue, it is one which we are obliged to accept).

However, the fact that Reasons 3 and 4 would more commonly be cited as reasons for other beliefs and actions is not alone an argument for saying that they are not part also of the reason for this action. And, although it might seem in certain cases intuitively obvious where the cut-off line would be drawn, if we would employ an economical counting method and limit the reason for any action to the primary reason(s) for it, we must first establish the basis on which one desirability characterization would be designated as the reason for an action and another as merely a secondary reason or as a reason for that reason.

It might, for instance, be supposed that we can delimit the primary reason(s) for an action by pursuing a progression of questions demanding reasons of the kind described above until reaching a reason for that action which, while expressing a belief or attitude held by the agent, was, for example, so general or remote that it could not be said to be a reason because of which the agent acted. The reasons 'standing between' this general or remote reason and the action would then be deemed to be the primary reasons. The latter point can be illustrated with an example like than one in Reason 4. Pressed for her reason for making the remark, the agent might reach some very general principle like:

Reason 5: Unnecessary suffering ought to be avoided.

It seems possible that the agent might at the same time deny having acted "because of" her belief that unnecessary suffering ought to be avoided. In somewhat the same way additional reasons of the kind given in Reason 3 above appear to be regressive. Depending upon inquirer ignorance, "giving the reason" might involve explication of such things as the meaning of the term 'pianist'. Thus, if the questioner were a child or a foreigner, we might have:

Reason 6: 'Pianist' means one who plays the piano.

It seems on the face of it strange to suggest that the agent undertook the action in question, even in part, because of her belief that the assertion expressed in Reason 6 is true.

However, to appeal to our use of "because" here is not actually to advance us further in the decision to use the more limited concept of a reason inasmuch as "because", like "reason", is a flexible expression. "Because" is not always, although it is perhaps usually, restricted, in accordance with my restriction on "reason", to certain beliefs and desires which somehow are primary in influencing the action.

That "because" is used too freely to handle the task can be shown if we turn to the concept of the cause of an event. In everyday usage we count only some of the various conditions necessary for its production as "the cause(s)" of an event, describing it or them as the factor(s) because of which the event occurred. Nevertheless, the decision to thus limit the description of a cause

is in one way artificial, since all the conditions necessary to bring about the event in question are describable as causes of that event, as factors because of which that event occurred. Clearly, we cannot make use of limits on the use of "because" to establish the boundaries of reasons proper, since "because" would be flexible enough to handle whichever counting principle were to be adopted.

In certain respects the "decision" I have been discussing runs interestingly parallel to the decision to count some factor or factors as the "cause" of an event and to relegate other factors which were equally necessary in causing the event in question to the status of "mere conditions", and it may be worth examining that decision further.

Although the latter distinction is not one which has a place in scientific conceptions of causation, where all the conditions required for the production of the effect are causal conditions and might thus be said to form part of the "cause" of that effect, it has been observed by some philosophers that informal discourse does allow of the distinction and, indeed, legal theory, embodying as it is said to do commonsensical distinctions, relies upon it. Further, just as it was shown above that at least what is given as a reason is relative to states of the inquirer and the context of the inquiry, it has sometimes been argued that the distinction between the cause of an event and the mere conditions involved in its production is an entirely relative one - relative, that is, to the ignorance (Mill) or the purposes (Collingwood) of the inquirer. But attempts have also been made to establish the distinction as something more than that and to show it to be based upon principles (albeit vague ones), and this corresponds to my effort to emphasize that while conditions of context may be introduced to explain any variations in the practice of reason giving, nevertheless, there still might be expected to be general principles which regulate how reasons for actions ought to be delimited.

One attempt of the kind described to establish principles underlying the distinction between "the cause" and "mere conditions"

is that of Hart and Honoré.[7] In their discussion it is argued that what decides that any condition necessary for an action is treated not as a "mere condition" but as the or a cause is primarily its abnormality - its not being present as part of the usual state or mode of operation of the sequence in which the effect occurs. This is so, Hart and Honoré emphasize, because demands for the cause of an event in everyday and legal contexts of inquiry, as distinct from scientific ones, are associated with a concern with abnormal functioning, with "....the puzzling, particular contingency" (page 31).

But although reason giving and justification may be seen to some extent to be affected by the kind of consideration which Hart and Honoré show to affect the limitations we place on the expression "the cause", as well as by the more obvious considerations mentioned earlier (the relative purposes and ignorance of the inquirer) - having reasons clearly is not so limited. Its obvious and everyday nature may mean that I am never called upon to give my wish to go to work and to get there on time as my reason for getting up each morning, yet however normal and obvious it may be that is my reason nevertheless.

Thus, we seem to be left without a way of distinguishing what we have called the reason(s) for an action from the additional desirability characterizations and beliefs necessary for the action's being carried out. That we often rely on some such intuitive distinction is evident. We speak of the "main" reason(s), of the "primary" reason(s) and of the "most important" reason(s) for undertaking actions. And, just as we distinguish the cause(s) of an event from mere conditions involved in its occurrence, so often we carefully seem to refrain from allowing that certain beliefs and desires, which (it is admitted) were necessary conditions for our action's taking place, played an important or central role in the decision. But these expressions and qualifications appear to be used too loosely to be helpful in providing a justification for distinguishing primary reasons. Sometimes what is regarded as the main reason is the desire which the agent regards himself or herself as feeling most strongly or intensely.

At other times it would appear to be distinguished by being the more unusual or abnormal consideration in the sense in which Hart and Honoré show "the cause" to be, while at yet other times it appears merely to be the one which the agent regards as best grounded or justified. And yet, confusingly, we do not always appear to be inclined to recognize such a distinction at all: we speak rather as if any consideration, however remotely it affected the decision to act, counts as part of the reason for so acting.

It seems, then, that we cannot distinguish primary reasons. Without a way of delimiting some particular desirability characterization(s) as the primary reason we must not see ourselves as presenting here anything other than the conditions which must be met in order for something to be said to be sufficient to count as (part of) the reason why we act. Let us now turn our attention to the general distinction between reasons and causes.

Reasons and Causes

It has been argued by some philosophers that the explanations which are offered for the events and phenomena comprising human behaviour are to be assimilated under two distinct models: some are expressed in terms of reasons and some are expressed in terms of causes.[8] Causal explanations are given of human behaviour when that behaviour is characterized in terms of movements made by bodies and bodily parts. Reasons are only offered for human behaviour when that behaviour consists of and is characterized as being, intentional action. There is a difference, it is observed, between my arm rising and my raising my arm.[9] The difference may be expressed by saying that only in the latter case is the phenomenon an intentional action which invites explanation in terms of reasons; in the former case the phenomenon is merely a happening for which a causal explanation is appropriate.

Since what is in one sense the same event or phenomenon can under one description be an action and under another a movement (although my arm's rising does not always entail my having raised my arm, nevertheless my raising my arm always entails that my arm rises), it follows that what is the very same phenomenon, in one sense, may sometimes both be described and explained according to

the two different kinds of schema. Event X may be regarded as an action and explained by a reason, or it might be regarded as a collection of movements and explained by a cause. This much is admitted.[10] But the proponents of the reasons/causes distinction (such as R.S. Peters) would deny that any event which under one description counted as a true action could ever completely or sufficiently be explained merely by causes. Only reasons completely explain intentional actions. In addition, there is thought to be a class of cases comprising mere movements; that is, the class of things which happen to us, which we suffer rather than do, for which reasons explanations are never applicable. For example, if my arm rises in a reflex jerk or because it is pushed by an external force, then an explanation of why it does so in terms of reasons would never be appropriate.

The strong version of the theory proposing a dichotomy between reasons and causes asserts that anything which is a reason in no sense also is a cause. A more moderate position allows that while reasons may be causes, they are at least a distinguishable subset of causes.

It is this latter position which I espouse. I would assert that, although reasons are a subset of the class of causes, actions done for reasons or intentional actions, are distinguishable in philosphically interesting and important ways from 'actions' which are not. My doing X because I believe that by doing so I can achieve Y may in one fairly uninteresting sense be on a par with my doing X because I was pushed from behind. But in significant ways it is quite different from it. The difference between the two kinds of case may be distinguished, moreover, in the following simple way. Unlike behaviour which is merely caused, an action done for a reason, or what is sometimes described as an intentional action, is one the conditions causally necessary for whose occurrence include a wish or desire on the part of the agent to undertake the action in question.

At risk of inviting confusion, I shall from henceforth use "reason" as contrasted with "cause" when the above considerations are taken as understood.[11]

"His/Her Reasons" and "The Reasons"

Linguistic considerations sometimes have been adduced in support of the distinction between reasons and causes and reasons explanations and causal ones. It is said[12] that our usage reflects and upholds the distinction between reasons and causes. But, in fact, the situation is considerably more complicated than such a claim suggests. If ordinary usage can be appealed to at all in support of this distinction, it cannot in the simple way just mentioned. Rather, ordinary usage allows of an additional distinction among what generally are called "reasons" for phenomena, a distinction which is reflected in the difference between "the" reason and a person's own reason, or "his or her" reason. In certain cases, when we speak of the reason why some even X occurred, even when X is some human behaviour, we should not naturally substitute the possessive form. In other cases we should; only in certain cases the reason why P did X naturally may be expressed as P's reason for doing X.

Underlying the distinction which we do appear roughly to observe in language between "the reason" and "a person's reason", is the fact that reasons are not always contrasted with causes. Not only is there a perfectly good sense of 'cause' in which it appears to be used to cover what we usually mean by 'reason' when the latter term is being contrasted with 'cause' - as when we say "She has cause for her alarm", meaning that there is a reason for her to be alarmed - so too there is a perfectly good use of the term 'reason' in which it is treated as synonymous with 'cause'. Sometimes we speak of causes as reasons. For example, we say "The reason why I tripped was that the floor was slippery", when in citing the role played by the slippery floor we are offering nothing more than a simple causal explanation. In addition, even in non-sentient behaviour, we speak of "the reason"": we say "The reason the flowers grow in that direction is to get the sun".

Now, the distinction mentioned, which ordinary usage supports, between "the reason" and "a person's reason", overcomes the previously mentioned ambiguity in 'reason', since it is actually only the looser expression 'the reason' which admits of these

ambiguous uses. The expression 'his or her reason' does appear to be restricted to those assertions in which 'reason' is contrasted with 'cause', rather than being used as synonymous with it. The slippery floor may be described as being the reason why I tripped, but it would not naturally or correctly be described as being "my reason" for tripping. And this presumably reflects the fact that the tripping, not itself being an intentional action over which I had voluntary control, was merely caused.

The latter distinction between "the reason" why a person acts and "his or her reason" for so acting was first emphasized by Peters (1958).[13]

> "We.....often say of a man that his reasons may have been X but the reason why he acted like that was Y.....whether the explanation in question is correct or incorrect does not much matter: the point is that to speak of the reason why a person does something is different in that it is a way of calling attention to the law or assumed law that a given case actually falls under. His reason may coincide with the reason....But, whereas his reason - whether real or not - entails that a man is conscious of his objective, the reason why he did it does not. The reason why he did it (walked across the street) might well be sex or aversion to work; yet the individual might be quite unaware of pursuing or avoiding the relevant goals. And whereas to say that he had a reason for something is more or less to rule out a causal explanation, to give the reason why he did it is sometimes to subsume it under a law-like proposition of a causal kind. This is not necessarily so. For we can say that sex or aversion to work was the reason why he did it and simply be insisting that a different directive disposition is being exercised."
>
> (Page 8, The Concept of Motivation)

We can sum up that part of Peters's distinction which is important to us in the following way:

1. The person's reason for an action may ormay not actually have been operative in bringing about the behaviour in question; its characteristics are merely that:
   (i) the person was conscious of it, and
   (ii) reference to it indicates that what Peters calls a a reasons type of explanatory framework is appropriate, rather than a causal one.

(The distinction in (1)(ii) is a common one.) Peters's idea was that a reasons explanatory framework differs from a causal one in being based on what he calls the purposive, rule-

following model: explanations in terms of it actually explain only on the assumption of certain norms of efficiency and of social appropriateness.[14] So "his or her reason" is a reason but not a cause. On the other hand:

2. When we speak of "the reason" why a person acted in a particular way Peters claims that we refer to the phenomenon, whether reason or cause, which actually was operative in bringing about the said behaviour or action (or else to the habit or tendency which explains the behaviour by showing in what light it is to be regarded; what Peters elsewhere describes[15] as a trait, like tactlessness, or a directed disposition, like aggressiveness). We may be speaking of a cause of the behaviour, or introducing a causal explanatory framework. But we need not be: "the reason" may be a reason or a cause.

(The context of Peters's introduction of the latter distinction, it is interesting to note, is closely akin to the present one. Peters introduced his distinction in a discussion about unconscious mental states and the explanatory status of the concept of unconscious motivation. His purpose in introducing it was to argue that only in the (above) second sense of "the reasons" could it be said that unconsciously motivated behaviour is behaviour done for reasons.)

While in agreement with Peters, as was indicated earlier, in maintaining that a distinction is to be recognized between an action undertaken for a reason and behaviour or movements which are merely caused or, to use the terminology more closely reflected in everyday language, that between:

(1) The agent's having a reason of his or her own for acting, and

(2) There being a cause (or reason) of his or her behaviour or movement -

I now wish to argue that no kind of consciousness or awareness distinguishes and characterizes the former class of reasons. Before I do so, however, let me clarify my purpose in selecting Peters's account of the distinction between reasons and causes.

Apart from appeals to purported facts about the way in which we speak, a number of different considerations, of which Peters's

awareness condition for distinguishing "his or her reasons" is but one, have been introduced to establish that there is indeed a genuine difference in kind between reasons and causes, and between reasons explanations and causal ones. Although several of these are important, I do not intend here exhaustively to discuss the distinction. In trying to argue for the possibility of unconscious reasons, I wish to deal solely with the more limited question of the value of the previously mentioned awareness condition for distinguishing between reasons and causes, the claim that the reasons for our actions as distinct from the causes of our movements are such that we are always aware of conscious of them. The rationale for my restriction is the following one.

I shall argue that the awareness condition is not a necessary one for something's being a reason for a person's action. From the fact that the awareness condition is not a necessary one for something's being a reason, it may not be thought immediately to follow that unconscious mental states may be described as reasons for actions. For, unless it could be shown that the awareness condition expressed the only feature enabling us to distinguish reasons from causes, which it does not, then it would seem that we must first establish that the class of unconscious "reasons" meets any additional criteria for being a reason. It might thus be thought necessary, in order to establish the possibility of unconscious reasons at all, to examine all the other ways, mentioned above, in which it has been proposed that reasons are thought to be distinguished. This, however, is not so. In fact, none of the other criteria introduced to distinguish reasons will actually be sufficient to exclude unconscious beliefs and desires from being classed as reasons. This is so because the only way in which conscious mental states are taken to differ from unconscious ones, according to my analysis in Chapter One (which differs here, it will be remembered, from Freud's, whose notion of unconscious mental states varies in several respects from that of ordinary conscious mental states), is in virtue of their not having the quality of consciousness at all or, if they do, presenting themselves phenomenologically in a distinct manner. It follows that any feature which characterizes ordinary conscious reasons will apply equally to reasons which are "unconscious" according to my

definition.

The above remarks apply, of course, not only to other accounts of the distinction between reasons and causes but also to Peters's claim that explanations of human behaviour, as distinct from causal accounts of mere movements, are characterized by explaining only if we assume certain norms of efficiency and social appropriateness.[16] This claim of Peters has no particular force in the context of a defence of unconscious reasons.

The Awareness Condition

My aim in examining the awareness condition, it will be remembered, is to establish whether such a condition represents a necessary condition for something's being a reason in the sense in which 'reason' is contrasted with 'cause'; that is, where "the agent has a reason" is contrasted with "there is a cause/reason". I argue that the awareness condition is not a necessary condition.

Two different versions of the awareness condition now must be distinguished. In the discussion in Chapter One concerning a definition of unconsciousness, a distinction was drawn among those of our mental states of which we are conscious or aware, between those which count as our own reasons and those 'reasons' which seem alien and unfamiliar. It is possible that the distinction between reasons and causes which we are now considering actually maps onto this finer distinction rather than onto the one between mental states of which we are simply aware or conscious, in any way, on the one hand, and those of which we are not, on the other. If this were so, only those of our reasons of which we are aware in the requisite intimate way would count as reasons proper while the rest of our so-called 'reasons' would be relegated to the class of non-reasons, or causes. Not even those of our unconscious 'reasons' of which we happened to become aware would count as real reasons. This stricter interpretation of the awareness condition sometimes has been espoused by philosphers.[17] However, to undermine the validity of the view sketched - that awareness is a necessary condition for something's being a reason - I need only show the implausibility of the simpler version of the awareness condition. I shall do this by showing that sometimes we have reasons proper

of which we are not aware in any way. I return, briefly, to the stricter interpretation of the awareness condition on page 120.

In order to show the implausibility of the simpler version of the awareness condition let us begin by observing that it is hazardously ambiguous to characterize this condition, as Peters does, in terms merely of "being conscious". Not all reasons are such that we actually are conscious of them, as long as we are merely speaking, in Freud's terminology, "descriptively". We engage in many (probably most) of our actions while not being conscious of our reasons for doing so, in the sense that we do not attend to them or think about them at the time of acting (or, often, ever).[18] For instance, we do so when we act distractedly, or when we engage in habitual or routinized actions. I usually lock the door when leaving the house without thinking of my reason for doing so, my wish to deter intruders. Yet it seems just as true to say of me that I have my reason for so acting in this case as it does in the cases where I consciously think about and attend to my reason for undertaking the action; for example, on the occasions when I purposely leave the door unlocked, thinking as I do that I will be out only in the front garden and will be returning indoors soon.

Thus, in order for an awareness or consciousness condition for reasons to have any plausibility at all "conscious of" and "consciously" must be characterized "dynamically" rather than "descriptively", to use Freud's terminology. Though my being conscious of my reason may in fact involve my actually attending to it, all that is required is that I am able to attend to it. Although I lock the door without thinking about why I do so, were my attention drawn to it at any stage during my performing the action of locking the door I could readily become aware of and give my reason for doing so.

It is perhaps important to point out that it is not in the spirit of a wish to improve upon ordinary language that the distinction between the dynamic and descriptive senses of "unconscious" is introduced here; it is merely a wish to capture a distinction which usually is conveyed by the context in ordinary discourse.

At worst, it might perhaps be said against ordinary language that it is ambiguous with regard to this matter. But there is certainly a common sense of the expressions "conscious of", "aware of", "having in mind" and suchlike, which implies the kind of capacity for attention and awareness which Freud's concept of preconsciousness connotes. Depending upon the contrast intended, "being conscious of X", or "keeping" or "having 'X' in mind", convey actually attending to X in one context, and having the capacity to attend to X, in another.

Since certain extra conditions must be met in order for a person to be able to state or express his or her reasons (a certain verbal and conceptual facility, and the necessary physical wherewithall, such as muscular control, etc.), it has been emphasized rightly, for example by Peter Alexander,[19] that the actual capacity in which we are interested here is best described in terms of recognition. The person must be able, at least inwardly, to recognize it as his or her reason when that reason is presented to him or her. The inarticulate person without much verbal facility or the mute, for example, might thus still be said to have reasons, even though he or she cannot give them. And this seems to be correct. The concept of being aware or conscious of a reason does not appear to be so closely tied with that of giving or expressing it, in verbal form or otherwise, that we would judge that it should be otherwise. Notice, however, that the notion of recognition does presuppose a certain degree of conceptual and verbal ability. The paralyzed mute who can merely recognize his or her reason when it is drawn to his or her attention understands language and uses concepts, at least inwardly.

I now wish to argue that even in the light of the above refinements to the awareness condition, it can be shown that there are cases of behaviour in which the person may be said to have a reason and is, yet, unaware of the reason. Some such position would seem to be suggested, in the first place, by the conclusions which I drew in my earlier discussion about the open-ended nature of reasons. If, as was there argued, there is no clear and consistent set of rules to allow us to distinguish those beliefs and

desires which played a more central role in effecting an action (what are roughly designated as the agent's primary reasons), from amongst all the beliefs and desires comprising the necessary conditions for that action, then there seems little plausibility in supposing that the agent might be expected to know or recognize all of the beliefs and desires comprising any of his or her reasons for acting.

However, it might be insisted that to demand awareness of all the beliefs and desires comprising each of one's reasons for acting would be to interpret the awareness condition unnecessarily stringently. While the agent is never aware in the requisite sense of all of his or her reason(s) for any particular action, it may be said, nevertheless he or she is always aware of some of the beliefs and desires that, in part, constitute the reason(s) for acting.

Even with the latter qualification, however, the awareness condition is, I believe, still in difficulties. In particular, there are a number of cases of adult human behaviour which seem to be best described as cases of a person's having a reason, in the fullest sense of "his or her reason", of which he or she is not aware in the requisite way. These cases fall into two classes:

(i) There are cases where the presence of awareness or consciousness may itself be doubted, and

(ii) Cases where, while there may be said to be awareness or consciousness of the reason for an action, the awareness condition may only be said incompletely to be met since the quality or nature of the awareness or consciousness makes the condition's being fully satisfied questionable, for it is non-conceptual awareness.

Consider cases of type (i): those split-second reactions, of what are also called reflex actions, for which an explanation in terms of habit is not applicable, sometimes are cited as presenting difficulties for an awareness condition inasmuch as they appear to involve acting upon what look to be reasons which the person would not have had time consciously to have arrived at. For example, the apparently calculated moves of the tennis player, or of the car driver, reacting in an unusual situation. However, against my case the following could be put. The fact that the person in these

cases was not actually conscious or aware, descriptively speaking, of the reason for acting while doing so, nor yet of the reasoning process which formed the reason, is not in itself a problem for the awareness condition, since it requires only that the person could become aware or conscious of the reason, and the presence of the capacity is not precluded in these cases. It is sufficient that were the driver X stopped at the instant at which he slammed on the brakes, he would be able to give his reason: "I saw the child dart out in front of me". And this account certainly correctly describes a number of split-second reactions of the sort I am considering.

In support of my case, however, it would seem that there are also more recalcitrant and extreme cases in which the perception on which the reaction is based, the basis of the reason, seems itself to occur at a subliminal or unconscious level. In the previous case, the driver's report sounds rather like a description of a speeded up version of ordinary conscious reasoning and reason formation in which the reason may be seen as leading to the action: "I saw the child and thought 'I must pull over to the right, fast!'"; in the latter case it sounds more like an inference leading from the action to the reason: "I must have noticed the child because I swung over so quickly.....". In more extreme cases the driver could not say why he or she acted at all; yet the facts are such that the obvious reason is that he or she noticed the particular threat in the situation to which the reaction can be seen as being a response - without perhaps noticing himself or herself noticing it. While in one sense we want to say that the person must have been aware or conscious of it, in the sense of consciousness or awareness which is demanded in an awareness requirement, he or she was not.[20]

Do we say that in such a case a driver's noticing the danger, albeit that she does so unconsciously, or does not notice herself doing so, cannot be said to be her reason for reacting as she does? If, against my argument, we follow the awareness condition for reasons then, at least in the case last described, the woman has no reason (of her own) for her action (although there may be a reason why she did it). But such a conclusion is implausible. What makes

it so is precisely that the latter action seems to differ from the ones described earlier in no other detail than the one of the woman's awareness of her reason. In all other important ways the actions seem to be exactly the same: the perception of the danger must have occurred, even if it did so unconsciously, and so must the reasoning. While this is so, it seems to be arbitrary and captricious to say that merely because the reason was noted at a conscious level in the one case and not in this one, it is at best "the reason" in the latter case, while it is "his or her" reason in the former one. I suggest, therefore, that in the latter case it also makes sense to speak of the woman <u>having her reason</u>: despite the lack of requisite consciousness.

I shall return to the force of this objection to the awareness or consciousness condition after first looking at the other class of cases described as presenting difficulties for that condition, those in which a person learns to make classifications between items on the basis of perceptual discriminations. One case of type (ii) might be as follows: the trained sexer of day-old chickens, for example, makes unerring distinctions on the basis of visual cues which are, at least in one sense, unknown to him or her. In this case, the "reasons" why the chicken sexer picks one of the chicks as female and another as male remain unknown to him or her, since the skill is not taught discursively: the components of the act of discrimination at no point are analyzed conceptually. A similar skill is that of recognizing a composer from listening to music. Here the skill may be taught by non-discursive trial and error techniques in the way that chicken sexing is. Notice that if it is, the picker might not only be able to distinguish "Scarlatti", "Chopin" or "Bach" unerringly while unable to cite reasons for doing so; further, he or she might not even be able to recognize a certain factor as his or her reason even if presented with it.

There are two possible explanations for an inability of the latter sort. One of these is relatively trivial and the other fundamental. The composer picker may simply be unfamiliar with the technical vocabulary of discursive musical "reasons". He or she may fail to recognize his or her own reason for discerning

unidentified works as Scarlatti's in: "The characteristic half-way modulation to the relative major when the piece begins in the minor key", even though in fact this was the reason. However, such a difficulty can in principle be overcome, since with the use of ostensive definition the composer picker could be taught the technical vocabulary so that he or she could come to recognize and state his or her reason. A more serious problem arises with the possibility that discriminations are not made on the basis of distinct, identifiable elements at all, but on the basis of an unanalysable perceptual gestalt which cannot be described or taught discursively.

In contrast to the kind of cases described are other cases of making distinctions on the basis of perceptual discriminations where the basis for the judgement originally are learned discursively and only come, with practice, to be exercised intuitively. The "reasons" on the basis of which the person was first taught to make the distinctions may themselves be forgotten over time, so that even the composer picker with a technical knowledge of music or the art historian judging fake from real antiques, for instance, might eventually come to proceed in a manner comparable with that of the chicken sexer. There are reasons on the basis of which the judgements are made, but they are not such that the person can state or perhaps even recognize them. Are we to say that they are not real cases of reasons, or are not the person's own reasons, in these cases? I suggest that they are indeed real cases of reasons and are the person's own reasons.

The latter cases have some bearing on the question raised earlier in relation to the suggested distinction between having a general conceptual and linguistic ability and being able to verbalize or conceptualize about the particular reason for a particular action or choice. It would seem that we do want to say that there are reasons for the composer picker's choices and that these are his or her own reasons for naming one composer rather than another. Just as I remarked that it seemed arbitrary to suggest of the driver who swerved on account of subliminally or unconsciously noticing the child dart out, that her noticing the child was not her reason for swerving, so it seems arbitrary to suggest of the composer picker who, for instance, originally learned his or her skill discursively and merely came to forget this (so that he or she made accurate discriminations discursively

at time t_1 and intuitively at time t_2), that at time t_1 he or she had reasons in the full sense of for the choice, and yet at time t_2 he or she had not. And if we do accept this conclusion, then we are able to say that the conceptual and linguistic capacity which has to be built into the awareness or consciousness condition is a general capacity, not a specific one. Being aware of one's reason entails having general conceptual and linguistic skills, but it does not further entail being able to conceptualize about one's reason or express it, even inwardly, in words.

In conclusion, I would suggest that the cases described would appear to be sufficient to cast doubt upon the awareness condition itself. If there can be reasons for apparently ordinary actions, no part of which the agent is aware, then the awareness condition cannot be presented as one which is necessary for the correct ascription of (his or her own) reasons to the agent.

Unconscious Beliefs and Desires as Reasons

I can now relate these conclusions to the question of whether unconscious "reasons" can be regarded as true reasons. Since it would seem that consciousness or awareness need not characterize all that we should wish to call genuine reasons, we can conclude that even the case of unconscious beliefs and desires affecting our behaviour and action may rightly be described as our reasons for what we do.

Let us now return (long enough to dismiss it) to the more sophisticated version of the awareness or consciousness condition which states that we must not only be aware of our reasons, but be aware of them in an intimate way as constituting our own reasons. Remember that, even if the simpler thesis I have been discussing had been established and it had been shown that "reasons" proper must be such that we are conscious or aware of them, it is apparent from the discussion in Chapter One that this characteristic of reasons would have failed to exclude all cases of unconscious mental states from the class of reasons. For we can and do become aware of some of the unconscious 'reasons' for our actions as the result of self-analysis and therapy. However, all unconscious

reasons would be excluded by the more sophisticated hypothesis that real reasons or reasons proper (his or her reasons) are those with the characteristically familiar feeling associated with our awareness of them. This hypothesis would allow that unconscious 'reasons' could not be reasons since even when we happen to be aware of them, they do not have this characteristic familiar feeling or aspect.

However, I would suggest that since it has been shown that not all reasons need be such that we are aware or conscious of them, the sophisticated and stricter awareness condition described above has nothing left to recommend it. We are entitled to speak of unconscious reasons, both in the case where the reasons in question comprise beliefs and desires of which we are in no sense aware, and in the case in which, although aware of them, we are so in a particular, qualitatively distinguishable way.

Despite this conclusion, there remains one context in which our awareness of and capacity to give our reasons does become critical. It is the context in which we evaluate an action as "reasonable" and we shall return to it when we discuss the questions of the rationality and reasonableness of behaviour in Chapter Five.

Reasons, Infallibility and Rationalization

My conclusion that there can be unconscious reasons runs counter, of course, to any suggestion that the infallibility or incorrigibility we enjoy with regard to them is what characterizes real reasons. Although such a claim was already raised and dismissed in Chapter One, it perhaps requires further comment. Notice first that since fallibility or corrigibility involves ignorance as well as error, the cases described of having reasons of which one was completely unaware serve to refute the latter theory as much as do cases of having reasons of which one was aware, but in less than the requisite, intimate way.

The theory that reasons proper must be known infallibly appears to be cast into doubt by a person's being able not only

to be wrong about his or her deductions of unconscious reasons but also, despite the alleged special access that person seems to enjoy in relation to them, to be wrong about the consciously held beliefs and wishes, and consciously felt emotions and attitudes which are cited as reasons. Sometimes people believe that they act for certain reasons, which they are aware of in the requisite, intimate way, when in fact it seems that these 'reasons' are merely "rationalizations" which have played no part in bringing about the said actions at all. And, while they may still be credited with being reasons, in the weakest sense of being the reasons for which the person believed himself or herself to have acted, it seems to me that on the standard meaning of the term this makes the reasons cited merely "rationalizations" and not really reasons proper at all.[21] Indeed, without the concept of rationalization, Freudianism, Marxism and most of modern social science would be impossible.

The evidence from rationalization, however, while at first sight lending support to my simple dismissal of the infallibility or incorrigibility criterion for reasons, in fact may be regarded as showing, rather, that the question of whether or not we are infallible or incorrigible in our knowledge of our reasons (whether or not, that is, reasons can be characterized by the infallibility or incorrigibility we enjoy in relation to them) is one which is interestingly moot. For someone wishing to maintain such incorrigibility or infallibility has always the recourse of denying that the cases cited are properly described, and insisting instead that as long as the person was sincere in making the assertion, the 'reasons' given by a person must be the person's real reasons. That is, someone could simply stipulate that incorrigibility or infallibility must be a characteristic feature of all reasons. And because this move remains possible, it seems to be question-begging to cite cases of rationalizations, since the very concept (that is, the concept of 'a reason' which the agent sincerely, but falsely, believes to have been the reason for his or her action) presupposes that we are not always infallible or incorrigible in our knowledge of the real reasons for our behaviour. So a criterion for reasons based on the infallibility or incorrigibility we may or may not be supposed to enjoy in relation to reasons, must be philosophically inconclusive.

The above-mentioned inconclusiveness invites a more general observation. Any dispute about the coherence of the concept of unconscious motivation (or "unconscious reasons" or "intentions which are unconscious") presents the same peculiar difficulties (difficulties which seem to be endemic to any dispute of this kind, including the one mentioned on page 98 concerning the coherence of the concept of a mental item which is not conscious), resulting in the inevitably inconclusive nature of any finding reached. The difficulty to which I refer may be put by saying that whatever is adduced in support of one or the other of the two possible positions (viz. the position that the concept of an "unconscious intention" or "unconscious reason" is conceptually incoherent and the position that it is not) may be shown to have the form of a petitio principii: it begs the question in favour of the position it would support, and is only convincing given a previous acceptance of that position. Thus, Hampshire acted correctly in phrasing his urge for the unintelligibility of unconscious intentions, quoted at the beginning of this chapter, as an admonition or prescription, rather than as an assertion of a proven, or provable, fact.

Even more than the arguments proposed in support of the conceptual incoherence of unconscious reasons, those presented to show the impossibility of speaking of an unconscious intention illustrate the nature of the impasse which I have just described. Let us then turn briefly to an examination of such arguments in order further to understand it.

Unconscious Intentions

The strategy of arguments put forward[22] in support of the claim that the notion of unconscious intentions is a conceptually incoherent one consists in showing, or attempting to show, that the way in which we speak of "unconscious intentions" is incompatible with key features of the concept of ordinary, conscious intentions, features which are consistently reflected in the ways in which we speak about them. In particular:

(1) It has often been argued that one's intentions are the sort of things about which one cannot be mistaken. If a person

sincerely asserts that he or she has an intention then that intention is truely attributable to that person; and
(2) It is asserted that when a person has an intention it is true that that person acknowledges to himself or herself the intention to do that which it is his or her intention to do.

In support of the first claim, that we cannot be mistaken about our intentions, the following argument is proposed. Being mistaken, it is said, entails knowing by observation that one is ignorant or in error; but since we do not know about our own intentions by observation, we cannot be said to be ignorant or in error concerning them. But both tenets of this argument, premiss and conclusion, seem to be question-begging. For surely we would want to say that *if* it is *assumed* that there are unconscious intentions (and that the concept of an unconscious intention is coherent), then, in the first place, there are intentions (of ours) which we know observationally, viz. those of our unconscious intentions which, with the help of therapy or special skills of self-analysis, we come to deduce from our knowledge of our behaviour and psychological tendencies. And, in the second place, we can be said to be *mistaken* about our intentions, for just the same reasons and in the same way as it was shown earlier that we may be said to be mistaken about our *reasons*.

So, *if* we were to accept the conceptual coherence of the notion of an unconscious intention, then it couldnot be said that the feature expressed in (1) *did* unfailingly characterize all intentions.[23]

Turning now to the second feature of ordinary, conscious intentions which, it is argued, precludes the possibility of having intentions which are unconscious (the requirement expressed in (2) that the intender always *acknowledges* the intention), we find that it would seem to be question-begging in somewhat the same way as were the assertions made in the requirement expressed in (1). And, again, as was shown with the requirement expressed in (1), there is no neutral ground. Those who refuse to accept the coherence of the notion of an unconscious intention refuse to admit cases of unconscious intentions as instances of intentions -

when intentions are described as having particular essential features - even when, if unconscious intentions were possible, they would rightly be said to have these features. But those (such as myself) admitting the possibility of unconscious intentions equally can be said to err in assuming the coherence of the notion of unconscious intentions by introducing the fact that unconscious intentions have the essential feature in question: in this case, the feature expressed in (2), that the agent acknowledges his or her intentions.

On the assumption of the coherence of the notion of an unconscious intention, my first move is to say that the post-therapeutic assertions we make about our unconscious intentions can be said to be assertions about intentions and, ipso facto, that the concept of an unconscious intention is one which is coherent. Now those arguing against me, for the incoherence of the concept of an unconscious intention, acknowledge the class of such supposed assertions of 'intention', but make the move which exactly parallels that made above, claiming that such post-therapeutic assertions of 'intentions' are not instances of intentions at all. They cannot be, it is said, since the concept of an unconscious intention is conceptually incoherent.

Thus, the whole question of whether or not there can be unconscious intentions or reasons, the question of whether the concept of an unconscious intention or reason can be said to be coherent, is, in an important sense, moot. In order to prove the conceptual incoherence of the notion of an unconscious intention, the standard distinguishing features of ordinary, conscious intentions or reasons are introduced. But, if there were such things as unconscious intentions or reasons, then the standard criteria would collapse, or at least be stretched and weakened to cover them, just as Descartes's criterion for mental substance has given way in the face of the apparent coherence of the concept of unconscious mental states. To appeal to the standard criteria to argue against the coherence of the notion of unconscious intentions or reasons would be to beg the question. Yet, in the same way, to introduce the fact that we speak of unconscious intentions or reasons in order to prove the criteria to be inadequate, seems

equally question-begging.

(Similarly, attempts to prove that Freud himself did not propose the concept of an unconscious intention which opponents of the notion of unconscious intentions attribute to him,[24] can have little other than historical value. Whatever concept Freud actually employed (if he was consistent in his use of one at all, which can be questioned), his work has given rise to the interpretation in question, and it has thus taken on a significant life of its own and deserves our attention for its intrinsic interest and plausbility.)

Needless to say, the above-mentioned inconclusiveness leaves us free to disregard such a theory and to accept my conclusion that unconscious beliefs and desires may be said to function as the agent's reasons or intentions for his or her actions. For, if the conceptual impossibility of unconscious reasons or intentions cannot conclusively be established, then the conceptual possibility of them remains open. Moreover, given the substantive work of Freud and neo-Freudians, there is every reason to posit the existence of unconscious reasons or intentions.

Unconscious Rationality

In admitting the possibility of unconsciously motivated behaviour, are we further committed to the claim that the person whose behaviour is wholly or partly unconsciously motivated may be rational in doing what he or she does? Some have said, mistakenly as I shall argue, that the answer to this question would appear to be a resoundingly negative one. For unconsciously motivated behaviour is often cited as one of the paradyms of irrational behaviour and the neurotic as the very opposite of what we mean when we speak of a rational person. Nor do we have far to go to find what look to be substantial reasons why this should be so.

First, it may be said, calling an action rational implies that it stems from, or is based upon, beliefs and decisions arising out of a sequence of logical reasoning (about one's priorities, the means at hand, one's goals, the facts of the case, etc.). Yet

Freud has described the very nature of primary process or unconscious 'thinking' as illogical and unreasonable, as thought processes which patently fail to come to grips with or to take into account any features of external reality at all.[25] (And, although we dismissed the concept of primary process thinking as a criterion for distinguishing unconscious mental states, we cannot so readily ignore the evidence which prompted Freud to introduce the concept.)

Secondly, it may be supposed that the very notion of rational action presupposes an awareness and understanding of the reasons why we act, since it is understood to be based upon a cool appraisal of alternative courses of action. Yet, by its nature, behaviour which is unconsciously motivated implies, at least for the most part, the absence of such awareness. Thirdly, it may be said that most unconsciously motivated behaviour, or at least most of that which is described by Freud, is, seemingly, pointless and futile in the extreme. Even when regarded as a means of achieving unconsciously perceived objectives, it often appears singularly ill-designed for that end. And such behaviour is in that respect essentially irrational. What could be less rational, it might be asked, than the repetitive and apparently stupid actions of the typical obsessive ritual described in Freud's case studies? And, lastly, it may be claimed that only behaviour which counts categorically as voluntary may be a candidate for being described as rational, and since the status of unconsciously motivated behaviour as voluntary is, at best, in question, it cannot be described as rational behaviour at all.

Peters's distinction between "the reason" and "his or her reason", to which I referred earlier, has important ramifications for this discussion. Peters, it will be remembered, argued that only in the sense of "the reason" (see page 110) - where the person was conscious of a reason, but it may or may not have actually been operative - could it be said that unconsciously motivated behaviour is behaviour for which there are reasons.[26] In arguing thus, Peters aimed to defuse the line or argument which might suppose that since being explained in terms of reasons is what

constitutes rationality, unconsciously motivated behaviour, for which there can be said to be reasons, might be said to be rational behaviour. Even though we do not generally speak of the reasons for unconsciously motivated behaviour, Peters insisted, we do not, or at least we should not, speak in the possessive form of a person's own ("his or her") unconscious reasons.

Since the publication of The Concept of Motivation, there have been intermittent attempts to support Peters's move and to show that the phenomena described by Freud and the psychoanalysts is an instance of irrational, or at least non-rational, behaviour: attempts which sometimes have emphasized and elaborated upon one, and sometimes upon another, of the different approaches to, or conceptions of, rationality already mentioned.[27]

Peters's and subsequent efforts to repudiate the notion that unconsciously motivated behaviour can be said to be rational may all be traced to two significant articles written during the late 1940's and early 1950's, Toulmin's "The Logical Status of Psychoanalysis"[28] and Flew's "Psychoanalytic Explanation".[29] Toulmin argued that:

> "The kernel of Freud's discovery is the introduction of a technique in which the psychotherapist begins by studying the motives for rather than the causes of neurotic behaviour."
>
> (Toulmin, Page 218)

Toulmin and Flew were concerned to show that while Freud was significantly right when, in his clinical discussions, he introduced the concept of unconscious motives, purposes and intentions for neurotic behaviour, he was profoundly wrong when, at a theoretical level, he insisted that these unconscious motives were actually inferred efficient causes.

> "Freud as a working psychoanalyst is primarily concerned with discovering and making patients realize and admit the motives, purposes and intentions of their neurotic behaviour, which motives, etc., are called 'unconscious' because, until he has done his work, his patients have no knowledge of them. But Freud as a theoretician seems to think.....that he has inferred the existence of unconscious mental processes which produce real and palpable obsessive actions."
>
> (Flew, Page 10)

They advocated that Freud be interpreted not as introducing explanations in terms of "recherché, not to say disreputable" (Flew, Page 12) efficient causes, but as introducing motive-type explanations. And, while admitting that there are differences between the explanations of our ordinary conscious motives and explanations in terms of unconscious motives, Toulmin and Flew insisted that they are merely different in degree - while the difference between motive-type explanations and causal-type explanations is a difference of kind.

What did Toulmin and Flew regard as the major differences between explanations in terms of motives, purposes and intentions, on the one hand, and those in terms of causes, on the other? They were concerned with the same kind of distinction as that discussed earlier, between causes and reasons and causal explanations and reasons explanations. Toulmin and Flew subscribed to the view that there is a radical difference in kind between causes and causal explanations, on the one hand, and motives or reasons and motive explanations, on the other hand. Further, that each kind of explanation was appropriate to different phenomena: motives and reasons explain human action, causes explain movements - so that explanations of the one sort are simply not reducible to explanations of the other.[30]

The emphasis placed by Toulmin and Flew upon the analogy between unconscious motives and the ordinary, conscious reasons and motives which we offer in explanation of our ordinary, rational behaviour, suggests the possibility of viewing unconsciously motivated behaviour itself as rational, inasmuch as it is a rational response to the situation as perceived from the point of view of the agent's unconscious state of mind. For Toulmin and Flew, arguing that unconscious mental states were motives meant that ipso facto they could not also be causes. While not entirely accepting this position (see page 108), I wish to examine the suggestive notion that since unconscious mental states seem to function in the way that motives and reasons do, it may yet be possible to assess the rationality of unconsciously motivated behaviour.

To return to the example used in the previous section, given that Freud's patient wanted to prevent her parents from having sexual intercourse by opening the adjoining door, then albeit that the wish may not have been one of which she was entirely conscious, nor one which was based upon solid grounds (her fear that her mother's giving birth to another child would be to her own disadvantage may have been ungrounded), nevertheless, as was observed previously, there does seem to be a way in which it is true to say that her action was appropriate to her wish. And, thus, there might be said to be a sense in which Freud's patient acted rationally in doing what she did. More generally, what looks to be a bizarre, inappropriate and thoroughly irrational action is often so only relative to the agent's consciously held beliefs and wishes. In the light of her consciously-held beliefs about cause and effect, the nature of pillows and the ways in which two people can be prevented from having sexual intercourse, and her consciously held desires and wishes, Freud's patient's ritual with the bed and bolster was clearly irrational. Yet, if we can temporarily forget that her unconsciously held beliefs may themselves be peculiar and unjustified, and if we regard her behaviour in the light of those beliefs, then it does take on the appearance of having a certain rationality. It makes sense and seems appropriate in a way that it fails to do if we regard it as a response to her consciously held beliefs and desires. It is not so obvious, as was also remarked earlier, that all, or perhaps even many at all, of Freud's other cases of neurotic behaviour with unconscious origins are rational, even in this sense. Thus, the assumption which seems to be shared alike both by those urging that unconsciously motivated behaviour be classed as rational, and those who would regard it as irrational, or non-rational - the assumption that all neurotic behaviour and all unconsciously motivated behaviour in general, has the same characteristics, and is either rational or not - is one which may be questioned. It seems likely that just as it was shown that the concept of the unconscious motivation of behaviour (as distinct from behaviour's being produced by unconscious mental states, or its being a sign or symptom of them) fitted only some of the cases which Freud cites, and not others, so it might be that only in some cases of behaviour for whose explanations we introduce talk of unconscious mental states can it plausibly be said that it is

rational.

However, I wish to defer any further discussion on this point in order to deal first with the prior question of the nature of rationality itself. For the most obvious explanation of the disagreement between those describing unconsciously motivated behaviour as rational and those describing it as irrational, is simply that the term 'rational' is subject to different interpretations, and that its very ambiguity has served to cloud discussions about its application in this area.

Notes and References

1 In the essay on the Unconscious (1915) Freud makes the following remarks about such an objection:

> ".....our most intimate daily experience introduces us to sudden ideas of the source of which we are ignorant, and to results of mentation arrived at we know not how. All these conscious acts remain disconnected and unintelligible if we are determined to hold fast to the claim that every single mental act performed within us must be consciously experienced; on the other hand, they fall into demonstrable connection if we interpolate the unconscious acts we infer. A gain in meaning and connection, however, is a perfectly justifiable motive, one which may well carry us beyond the limitations of direct experience. When, after this, it appears that the assumption of the unconscious helps us to construct a highly successful practical method, by which we are enabled to exert a useful influence upon the course of conscious processes, this success will have won us the position that it is both untenable and presumptuous to claim that whatever goes on in the mind must be known to consciousness."
> (My underlining)
>
> (Page 117, op. cit.)

2 "Disposition and Memory", Pages 174-5. Reprinted in Hampshire, ed., Freedom of the Mind, Princeton University Press, Princeton, N.J., 1971.

3 By Peters in The Concept of Motivation, Routledge and Kegan Paul, London, 1959.

4 By Anscombe, Intention, op. cit.

5 From Davidson, "Actions, Reasons and Causes", op. cit.

6 See, for example, Shwayder's claim (The Stratification of Behaviour, Routledge and Kegan Paul, 1965) that reasons for doing an action can be distinguished from reasons for doing it in a certain way:

> "One may kill a dog for the reason that it is rabid; but knowing other things about dogs and drugs, he may do the act by injecting the creature with strychnine. What the agent knows shows itself not as a reason for doing the act, but, if you wish, as a reason for doing it in a certain way. But, his knowledge that dogs are creatures mortally allergic to strychnine may not figure at all, for he may act solely from the belief (for the reason) that the animal, whatever its kind, is rabid."
>
> (Page 88)

This, at least, seems to me clearly too restrictive. While in this case, the reason for doing the action in a certain way may have been no part of the reason for doing it, it seems clear that the reason for doing it must be counted as part of the reason for doing it in a certain way.

7 Hart and Honoré, Causation and the Law, Oxford University Press, 1958.

8 Anscombe, op. cit.; Melden, Free Action, London, Routledge and Kegan Paul, 1961 and "Willing" in Philosophical Review, 1960; Hart and Honoré, op. cit; and Peters, op. cit.

9 Melden, "Willing", op. cit.

10 See Peters:

> "..... if we are, in fact, confronted with a case of a genuine action.....then causal explanations are ipso facto inappropriate as sufficient explanations. Indeed, they may rule out rule following purposive explanations. To ask what made Jones do something is at least to suggest that he had no good reasons for doing it. Similarly, to ascribe a point to his action is ipso facto to deny that it can be sufficiently explained in terms of causes, though, of course, there will be causes in the sense of necessary conditions. A story can always be told about the underlying mechanisms; but this does not add up to a sufficient explanation, if it is an action that has to be explained."
>
> (Page 12, op. cit.)

11 See L.R.V. Burwood and C.A. Brady's paper: "Philosophical Models of Man", Educ. Review, Feb. 1981, for further explication of the reason-cause distinction in relation to the free will-determinism problem.

12 For example, by Hart and Honoré:

> "It would be somewhat unnatural in the informal discourse of ordinary life to describe any of this range of cases by saying that one person caused another to act, and in some cases this description would be positively misleading. 'He caused me to act' would be merely unnatural (and 'He made me do it' natural) in those cases where one person merely advised, or tempted or requested another to act, or procured his action by offering a reward."
>
> (Pages 48-49, op. cit.)

13 Peters, op. cit., especially Pages 9-16.

14 Peters, op. cit., Pages 1-8.

15 Peters, op. cit., Page 63

16 Peters argues that Jones's walking across the street in order to get some tobacco explains what Jones does only on the assumption that walking across the street is an efficient way of getting to the tobacconist's. And, because this norm is an assumption which we must make in order to see Jones's wish to buy tobacco as completely explaining why Jones acts as he does, Peters concludes that we must view such an explanation as fundamentally different from an ordinary causal one in which giving an explanation involves citing causally necessary and sufficient conditions for the phenomenon explained.

But no amount of assumed norms will enable us to explain Jones's behaviour unless we are able to ascribe a belief in those very norms to Jones himself. And, if Jones's subscribing to the norms in question is thus necessary in order for us completely to explain his action, then there seems no objection to describing his belief that walking across the street is an efficient way of getting to the tobacconist's as a necessary condition of Jones's action. So, Peters's distinction between two fundamentally different types of explanation appears to lose significance.

I would suggest that the error which Peters made stems from precisely the fault, described earlier in our discussion of reasons, of presupposing some method of delimiting those beliefs and desires comprising the main or primary reason(s) for an action. As was shown earlier, in the absence of any way of so delimiting "reasons", it must be allowed that any beliefs and desires of the agent's which were necessary for the occurrence of the action in question, count as part of the agent's reason(s) for so acting. As long as this open-ended feature of reasons is recognized, there would be no temptation to claim, as Peters does, that Jones's wish for tobacco alone was Jones's reason for his action.

Let us notice too that Peters's characterization of the distinctive nature of "his or her reasons" explanations appears to exclude from the class of "reasons" those explanations described in Chapter Two in which a "whim" is cited. A desire to act which is (subject to the qualification made in that chapter) unconnected with any other of the agent's desires, is likewise independent of any norms of efficiency or social appropriateness. To cite my feeling like steak may fully explain my ordering the tournados: we need introduce no adherence to norms to do so. (We must, of course, include that I see that tournados are on the menu and that I know that they are steak.) Now Peters may be prepared to describe actions taken on whims of the kind described as better fitting a causal explanatory framework (certainly actions taken on a whim are in some cases awkwardly indistinguishable from actions for which there are said to be so-called 'mental causes' which must, presumably, be placed in the category of causal explanation). But linguistic considerations would appear to favour our allowing even whims as "his or her reasons". My sudden yen for steak, unreasonable as it may be, is still surely my reason for ordering steak, and not merely the reason why I do so, or the cause of my doing so.

We may conclude that little plausibility attaches to Peters's first way of characterizing the distinction between "his or her reasons", on the one hand, and the cause of a movement or of some behaviour, on the other.

17 For example, by Peter Alexander, "Rationality and Psychoanalytic Explanation", op. cit.

18 The point has been made often. See for instance Anscombe's remark in "Intention", reprinted in White, ed., Philosophy of Action, Oxford, 1969, that:

> ".....it is not in all cases that 'I did so in order to.....' can be backed up by 'I felt a desire that....'. I may simply hear a knock on the door and go down stairs to open it without experiencing any such desire."

(Page 146)

19 By Alexander, op. cit. See Page 331:

> "I may have a reason in mind without attending to it.... however, it is a necessary condition for my acting for a reason that I should be able to become aware of my reason if I think about my behaviour, although I need not be able to state it.....(I must be able to) recognize it as my reason."

20 Dennett (Content and Consciousness, Routledge and Kegan Paul, London, 1969) makes interesting use of this dual role played by the notion of awareness with his distinction between 'awareness$_1$' and 'awareness$_2$'.

21 I differ from Peters on this point, since his characterization of "his or her reasons" includes all reasons cited by the person, whether or not such reasons actually were effective in bringing about the behaviour in question.

22 For example, by Seigler ("Unconscious Intentions", Inquiry, 1967). At the same time as denying the coherence of the concept of an unconscious intention, Seigler actually admits the notion of unconscious desires, wants and wishes, concluding that:

> "Perhaps Freud's insight is not that he can see slips as or discover them to be, intentional, but rather, that we can discover slips to be very revealing about the wishes, wants, desires, hopes, etc., of the speaker....."

(Page 263)

But, despite its superficial similarity to my previous conclusions concerning unconscious motives (that is, that there are unconscious desires, wishes, wants and purposes), Seigler's claim, in the context of this argument, is actually tantemount to a denial of the category which I have

explicated as that of unconsciously motivated behaviour, as distinct from behaviour which is merely produced by unconscious mental states, and to an assertion that the only way in which unconscious mental states can be said to occasion behaviour and be introduced to explain it, is in this latter way - as producers of it.

Seigler's particular strategy is to show (i) that our ordinary concept of an intention is inconsistent with the uses which are made of "unconscious intention", and (ii) that certain characteristics of the concept of a desire correspond to the characteristics which we attribute to unconscious mental states when we introduce them into explanations of behaviour, in particular, to those characteristics associated with desires as passions. He then perfunctorily bundles together wishes and wants with desires in order to conclude that there are only unconscious desires, wishes and wants (in his sense of these terms), and there are not unconscious intentions. According to his analysis, unconscious wishes, wants and desires produce behaviour, in my sense, but they do not motivate it.

In addition to arguing (1) that we cannot be mistaken with regard to our intentions; and (2) that having an intention entails acknowledging to oneself that one intends to do Z, Seigler argues that having an intention also entails (3) that A desires that a particular state of affairs obtain which does not now obtain; (4) that A often does, or makes an effort to do, what he says he intends to do; and (5) that A has other desires connected with desiring to do X.

The defects with the claims made in (1) and (2) are dea with at length on pages 123-6. In addition, a difficulty arises in that there is an incompatibility between the requirement cited in (1) and that in (4). If (1) is true then (4) cannot also be true, for the truth of (1), it seems, implies at least the possibility that A might never do what he claimed to intend to do, nor yet even make an effort to do so. If (1) is true, and a person cannot be mistaken in asserting "I intend", then intending must consist solely in sincerely making that assertion, and Seigler must be wrong in claiming that:

> ".....it must be the case that very often when A intends to do X, subsequently, he does, in fact, do X, for it seems that if A never did what he said he intended to do, unless there were very special accounts of his circumstances we should hesitate to say that he intended to do those things, but perhaps, depending on the particular details, that he wished or hoped, that he would X, or thought about the prospect of X-ing."

(Page 260)

Now it may be replied that Seigler is proposing not what must be the case here, but merely what is, contingently, the case. He might be prepared to acknowledge the possibility in principle of this lack of correspondence between

intending to X and doing X or making some effort to do X, while yet insisting that as a contingent fact about the way we are, the said disparity never arises. And this is indeed how he seems to be arguing a little later, when he remarks that:

> "If the entire language community only on rare occasions did what they said they intended to do, they would thereby not have our concept of intention."

(Page 260)

But even this claim is false, as it will be shown, either an outer, behavioural criterion must be chosen, or an inner one: what is important is that one must be chosen as the final arbiter. They cannot, as Seigler seems to imply, function together as joint final criteria, because of the possibility described of their leading us to contradictory conclusions.

So, it would seem that if he wishes to retain the claim about intentions which he makes in (1), then Seigler cannot also assert (4), and if he wishes to assert (4), then he must relinquish (1).

(3), the claim that having an intention involves that A desires that a state of affairs obtain which does not now obtain, and (5), the claim that having an intention involves that A has other desires connected with his desiring X, may both be said to be obvious and incontestable truths about intentions - although it is difficult to understand why Seigler chooses to introduce them in this particular discussion, since there is no reason why they should not also be true of A when A was unconsciously motivated; nor, for that matter, why they should not also be true of A if A merely had desires, wishes or wants, as distinct from intentions. Many desires are characterized by both of the requirements expressed in (3) and (5): and, if they are possible at all, unconscious intentions or desires might also be expected to be so characterized.

Assuming that (1) represents a more important feature of intentions than the requirement expressed in (4), and that given a choice between (1) and (4), the requirement expressed in (4) would be dispensed with in favour of the requirement expressed in (1), it would seem that (3), (4) and (5) may all be dismissed, at this point, as having no bearing upon the argument in question. This leaves us with the requirements expressed in (1) and (2), both of which are shown on page 124 to be question-begging.

And the positive side of Seigler's thesis is actually no more helpful than the negative one. Having established what unconscious mental states are not, viz. intentions, Seigler proceeds to give some arguments towards his conclusion concerning what they are - desires, wants and wishes. But these, too, prove inconclusive. Rather than considering these three concepts separately, Seigler deals only with desires,

on what must be taken as the assumption that in the respects mentioned, the three are not distinguishable. And what he cites as the characteristics of desires which make them closest to our idea of an unconscious mental state, are all features of desires which they have by virtue of being passions, or states which we are in:

(1) we can be unaware of them;
(2) they can overcome us;
(3) they are things which we try to resist, and are ashamed of.

While all the characteristics which Seigler cites in (1)-(3) are true of some desires, it is neither true that all desires have these characteristics, nor that wishes and wants which, it will be remembered, are lumped together with desires in his conclusions, do.

The question of whether we can be said to be unaware of the desires, wishes and wants which we have is one the answer to which could only come with a decision on the question at issue concerning the coherence of the expressions "unconscious desire", "unconscious wish", "unconscious want". To assume an answer would appear to beg the questions, since if we conclude that there can be unconscious desires, wishes or wants, then presumably we must admit that these states are such that we can be said to be unaware of having them. With regard to (2) and (3) above, desires are not always passions at all - they are sometimes reached by rational processes. And there is no reason to suppose that such a rational desire should overcome us, or even be a candidate for resistance, or an attitude of shame. In addition, we do not often, certainly we do not always, try to resist our wishes; on the contrary, their very impotence, the fact that they are mere wishes, seems to obviate that. Nor, for roughly the same reason, do we speak of their overcoming us. Similarly, we do not naturally speak of resisting a want, nor of our wants overcoming us. Both the resistance location and that to do with overcoming, are restricted to pure passions, and only desires (and not all desires at that, as we have seen), ever count as these.

23 In fact, this issue is somewhat complicated because we may even question whether these features can be said unfailingly to characterize all ordinary, conscious intentions. It does seem to be true that we do not appeal to observation or go through a process of finding out what our own conscious intentions are. And, we are usually treated as being unable to be mistaken with regard to them. In fact, "I intend" seems often to work like a performative utterance, the very saying of which constitutes the doing of it. However, sometimes, although less commonly, 'intend' seems to be used differently. It seems to me that we do ask "Did she really intend to do that?", at times, when the person's honesty is not in question, but her future conduct is.

Because of this ambiguity, and the fact that common usage has become somewhat tainted by psychoanalytic thought

and no longer gives us entirely consistent guidelines, philosophers have no choice but to arbitrate. Either they must decide to treat intention as a pure performative, in which case it will be correct to continue to ascribe intentions to a person even if that person never moves towards doing what he or she claims to intend to do, or else 'intend' will have to be given some sort of behavioural criterion for its correct ascription. The final decision as to whether a person has intended to do some action will have to come, so to say, from the outside.

The latter alternative need only commit us to the introduction of a behavioural criterion in extreme cases, of course. Normally, if a person expressed an intention then we should simply take that person's word for it, treating it as having the force of a performative. If, for instance, someone claimed to have been intending to write a letter, yet delayed for days, we should probably concede that that person intended to write it. But if a person were to claim to have been intending to write a particular letter for some number of years, and without evident impediment had failed to do so, then we might be tempted to introduce the behavioural criterion and rule against the correctness of the ascription of intention, insisting that while the person may have entertained a hope or wish to write the letter, he or she could not really be said to have intended any such thing. And, if the latter decision is made, it will be true that the person is corrigible, even with regard to those of his or her own intentions which are ordinary, conscious ones.

24 As some writers have done, for example, Shope, "Freud on Conscious and Unconscious Intentions", Inquiry, 1973.

25 a) Freud speaks, for instance, of the:

> ".....irrational character possessed by everything that is unconscious when we translate it into consciousness....."

(Page 226, "A Metapsychological Supplement to the Theory of Dreams", 1917, Collected Papers).

b) Peters, op. cit., has suggested that this is actually the strongest argument in support of the irrationality of unconsciously motivated behaviour that there is. He quotes Ernest Jones:

> ".....Careful students have perceived that Freud's revolutionary contribution to psychology.....was his proposition that there are two fundamentally different kinds of mental processes, which he termed primary and secondary respectively, together with his description of them. The laws applicable to the two groups are so widely different that any description of the earlier one must call up a picture of the more bizarre types of insanity. There reigns in it a quite uninhibited flow towards the imaginary fulfillment of the wish that stirs it - the only thing that can. It is unchecked by any logical contradiction, any causal associations: it has no sense of either time or external reality.

and comments that:

> "This surely amounts to saying that insofar as unconscious processes are involved, the thinking about getting to the goal cannot be described as either correct or incorrect, efficient or inefficient, intelligent or unintelligent."
>
> (Page 64)

26 although Peters's own conclusion was actually somewhat less categorical. See Page 94, op. cit, where he remarks that:

> "Neither the rule following purposive model nor the mechanical model of explanation are really adequate for conceptualizing his (Freud's) revolutionary insights."

27 For instance, Peter Alexander, "Rationality and Psychoanalytic Explanation", Mind, 1962; Mullane, "Psychoanalytic Explanation and Rationality", J. Phil., 1971.

28 Analysis, 1948-49.

29 Analysis, 1949-50.

30 Toulmin and Flew may also have been influenced, in stressing the difference in kind between motive type explanations and causal type explanations by certain considerations thought to flow from facts about the ontological status of motives, wants and wishes. In his analysis of motives, Ryle (The Concept of Mind, 1949) rejected once and for all the notion of motives, wants and "acts of will" as ghostly, introspectable mental occurrences which antecede every intentional action. Instead, Ryle insisted that, ontologically speaking, motives are much more like behavioural dispositions or tendencies. Now this analysis of the ontological status of motives sometimes has been thought to be linked with the distinction between the two different kinds of explanatory model in the following way. It is said that since causes are independently specifiable events, motives cannot be causes, for they are not always events. This argument has at least two flaws. Motives might be said to be events of some sort, but events of which we do not always have access through introspection. The fact that they are not always events of which we are aware, does not prove that they cannot be events. But a more fundamental difficulty arises with the stipulation that dispositions cannot be cited as causes. We do cite dispositions as causes, and as long as dispositions are themselves analyzed as states this seems to be a perfectly plausible move.

CHAPTER FOUR

Summary

The notion of rationality is explored and three different ways in which 'rational' qualifies particular actions are distinguished:

(1) 'Acting rationally' as acting in such a way that one furthers one's interests (or maximises) one's expected utility, minimizes one's expected disutility, minimizes regret, etc.;

(2) 'Acting rationally' as having a rationale or purpose in acting; and

(3) 'Acting rationally' as acting reasonably, that is, having and being moved by good reasons in acting.

The point is made that the assessment of a person's interests is, in part, evaluative. Moreover, a distinction is made between "being in a person's interests" and "being in a person's best interests". Reference to a person's best interests seems to imply an explicit comparison between the course of action undertaken and all other possible courses of action available to the agent.

It is argued that, minimally, an action is rational to the extent that it is in the interests of the agent as much as (or to a greater degree than) any other course of action which is known to the agent, or which might have been expected to have been known to the agent, and, maximally, that that course of action which of all possible courses of action is the one which is (in fact) in the agent's best interests is the most rationalone which the agent might have undertaken.

The notion of acting rationally, as related to having a rationale or purpose in acting, is then explored. Some sociological literature is drawn on and it is shown how an (initially) odd and unintelligible action nevertheless can be deemed rational as a means, given that it is done with some purpose or aim in terms of which it becomes intelligible. In this sense of acting rationally, no qualitative restrictions appear to be placed on the kinds of beliefs which can count as reasons. We do not have to hold that it is reasonable to pray for a good harvest, for example. All the

anthropologist says is that given the (unreasonable) assumption that praying can bring about a good harvest, we can regard the action as a rational way of achieving the desired end.

This leads us to the notion of acting rationally in the third sense: that of acting reasonably and being moved by good reasons. Several different requirements for something's counting as a good reason for an action are identified and the question of the justification of beliefs is tackled. I argue that for an action to be described as reasonable, it must be the case that:
(a) the agent is capable of citing good reasons for having so acted (and is sincere), and that:
(b) the agent was actually moved by those reasons. However, although perhaps paradigmatically the case, the agent does not necessarily have to be conscious of them.
This (third) sense of rationality as acting reasonably includes the notion of being justified in believing one's action to further one's interests.

Finally, having analyzed these three main accounts of what it is to act rationally, I make certain points about irrationality. I show that the term 'irrational' is used not only to describe a failure to meet each of the requirements expressed in the rationality characterizations in (1)-(3) above, but that it is also sometimes used to described:
(i) behaviour which is merely abnormal, and
(ii) behaviour which is involuntary.

* * * * * * * *

What does it mean to describe, as we do, behaviour, people, decisions and courses of action (both those taken and those merely proposed) as "rational" or "reasonable"? What does rationality consist of?

The aim of this discussion, that of establishing what is generally meant by expressions exploying the terms 'rational' and 'reasonable', is one which deserves a word of justification. For, in introducing the concepts of rationality and reasonableness, we risk erring by attempting to treat as precise or to make precise something which is by its nature imprecise. In particular, accounts of the meaning of 'rational' and 'reasonable' suffer, as has been observed,[1] from "inflated definition". And the ascription of the terms 'rational' and 'reasonable' encountered in everyday conversation seem sometimes in their function to hover uncomfortably between representing actual properties, on the one hand, and alluding to ideals and abstractions on the other.

In the light of such peculiarities and complications, we might be expected to approach even the simple question of how the terms are ordinarily used with some trepidation. My justification for attempting this task is a belief that beneath the outer layers exist a solid core of intuitively agreed upon distinctions which, while they sometimes are disguised, nevertheless are roughly discernable in a careful look at ordinary usage.

The Breadth of Rationality Ascriptions

Before turning to the apparently numerous ways in which the above issues might be addressed, let us note that, though the class of things described as "rational" is, as the above list would indicate, an extensive one, it nevertheless has limits. In being affected, that is, in being in certain mental states (of emotion, mood and desire) we are rarely described as being "rational" or "irrational", although we are in the actions we do out of our afffections. (My feeling sad is neither rational nor irrational, but my refusing companionship out of my grief may well be.) Although some beliefs are themselves in the nature of affections, nevertheless, the having or holding of beliefs, in contrast with

being affected in the various other ways described, is sometimes said to be rational or reasonable, or otherwise. For believing has been regarded traditionally as a voluntary act, as something we do rather than something which happens to us.

Yet, while few actions, apart from ordinary straightforward voluntary actions, are described as rational or irrational, not all voluntary actions naturally are so described. There are certain voluntary actions, in particular those which have no likelihood of significant long-term consequences and which are done without much prior thought (actions like that of turning to someone in a conversation and smiling) which are treated neither as rational nor as irrational. I shall return later to the significance of these qualifications on the application of the concept.

As has been remarked already, there seem to be a number of different ways in which the questions raised on page 143 might be answered; the term 'rational' takes on different colours in different contexts.

Rational Persons

Sometimes by being rational, we mean simply having what Kant called the faculty of reason. In this sense all human beings, including those whom we might also describe as "irrational", are rational, inasmuch as they were all born with the capacity to think and reason. When described in this way, our human rationality is being contrasted with the capabilities of animals.[2]

In a somewhat more refined sense, our rational capacities are contrasted with those of children, on the one hand, and the insane and mentally retarded, on the other, all of whom are treated as lacking rationality due to
(a) underdevelopment and/or
(b) malfunction or underdevelopment,
respectively, of the reasoning powers.[3] It is in this sense that the child is said to acquire rationality on reaching the "age of reason", that is, the time at which mental capacities reach something approaching a full state of development.

There is a slightly different sense in which the term 'rational' sometimes is treated entirely normatively, so that the insane, for instance, are regarded as lacking rationality through no particular incapacity of reasoning powers, but merely by dint of the unusual nature of their beliefs, inclinations, reactions or reasoning processes.[4] Notice that the three uses described are all uses of the term 'rational' which apply primarily to persons and their capacities rather than to the characteristics of particular actions.

(1) Rationality as Furthering Interests

Of interest to us here are two common notions of rationality which are applicable particularly to:
(a) the idea of rational action, or of rational behaviour, and
(b) to rational persons, only inasmuch as their rationality is a function of the likelihood of their engaging in rational behaviour.
The first of these introduces the concept of interests. Hobbes's thesis on the relation between self-interested behaviour, on the one hand, and rationality or reasonableness, on the other, which derives from his psychological egoism, is that behaviour which serves our own interests would never be "against reason",[5] or, more strongly, that to act in one's own self-interest always would be to act rationally.[6] We can convert this into the following definition which, if not an exact statement of Hobbes's view, nevertheless has been interpreted as his position:

> If and to the extent that a person's action is in his or her interests, when being in a person's interests means benefitting that person in some way, then it is to be regarded as rational.

For reasons which I shall explain, I suggest that the definition be amended to read:

> An action is rational to the extent that it is in the interests of the agent as much as or to a greater degree than any other course of action which is known to the agent, or which the agent might have been expected to have known, and
>
> That action which of all possible courses of action available is the one which is in the agent's best interests is the most rational action which the agent might have undertaken.

Let us begin by observing that this notion of rationality allows of rationality's being a matter of degree. Notice also that, unlike Hobbes's, this definition of rationality takes into account the agent's state of knowledge. I shall return to this feature on page 149.

When can the consequences of an action be seen as benefitting the agent? It might be supposed that we could define that which benefits an agent as that which the agent would view as beneficial or satisfying. If we are able to do so then we could treat rationality simply as a function of the extent to which an action is likely to satisfy an agent's desires or wishes. Unfortunately, as I shall argue, such an analysis is not open to us in accounting for rationality in terms of interests. For the concept of interests has an evaluative element.[7] Not everything which a person would desire or wish or view as satisfying or beneficial would be regarded as in that person's interests. Some courses of action are seen as being intrinsically harmful, either because they have intrinsic demerit as actions (one view) or because (another view) they are seen as instrumentally harmful in unfailingly producing in the agent states of mind themselves judged to be intrinsically harmful. And, because the concept of interests is a partially evaluative one, what would be beneficial to the agent cannot be equated with what the agent would judge to be beneficial.

Anything which will benefit a person at all may be said, as in my initial definition, to be in that person's interests. But there are different ways in which we may be benefitted by a particular course of action (for example, spiritually, materially, etc.) or, as it may be put, we have differing interests which may differently be served in any situation. Because of this, there is the possibility that some course of action might be in our interests in some way while not being in our interests in another way. As the definition was originally stated, an action might be said to be rational as long as it furthered any of the agent's interests and to be more rational in proportion to the number of different interests of the agent's it satisfied at once. But,

while the latter use is plausible, the former one is not. Although accepting a promotion may be in a person's financial interests, the suffering produced by her guilty conscience over other duties neglected might so outweigh the financial benefit as to make it a course of action which could not be said to be rational for her in the least degree.

So, we must define a course of action as being in a person's interests, simpliciter, as not merely its being in some particular one(s) of that person's interests and thus being beneficial in some way(s), but, in addition, being unlikely to harm that person in any other way to such an extent that the said harm might outweigh the expected benefit. Let us consider an example. In assessing whether P ought to travel to Ripon by bus, we might weigh up the likely disadvantages and advantages of P's doing so in the following way. Because she can read on the bus, what might be termed her intellectual interests will be furthered by travelling this way. Because she is prone to bus sickness, it is a mode of travel which in the interests of physical comfort, she would better avoid. We might judge that her intellectual interests outweigh in importance the mere interests of physical comfort and might, thus, judge it to be in her interest simpliciter to take the bus trip. Should we say such an action was rational?

It is certainly true that P could be said to have acted less irrationally in taking the bus, given our assessment of the weight of the interests involved, than had she not done so. But, we should nevertheless wish to know more about the possible alternatives available to her before conceding that her action was rational. So, judging an action to be in a person's interests simpliciter cannot merely be a case of weighing up the consequences of the course of action in the manner described.

It will be helpful at this point to contrast the notion of being in a person's interests with that implied in the phrase "being in one's best interests". To say of a proposed course of action that it is in the agent's best interests obviously includes the assumption that it will benefit the agent in some way and will

be unlikely to result in consequences whose ill effects are likely to outweigh the good expected from the aforementioned benefit. But, in addition, reference to best interests seems to imply an explicit comparison between the course of action undertaken and all other possible courses of action available to the agent.

(We may define a possible course of action available to the agent in the following way. Course of action X is possible for agent P just when X is causally possible for P and, were P aware of X, she would, without further reflection, find no overwhelming objection to adding X to her list of serious alternatives and assessing with an open mind its advantages and disadvantages. For example, it may be physically possible for P to reach her destination by getting out her hunting rifle, highjacking a passing car and compelling the motorist to act as chauffeur. However, such a course still may not be a possible one for P in the above sense since, were P aware of the possibility, she would without reflection preclude it from her final list on account of her moral scruples.)

Reference to the agent's best interests implies that the course of action undertaken is the one of all possible courses of action available which would be most likely to benefit the agent. When X, Y and Z are the only possible courses of action available, in the sense defined, my doing any of X, Y and Z would be likely to be in my interests, in the way described earlier. But my doing Y might be in my interests more than would my doing X or Z. (This may be on account of (a) a greater likelihood of Y's having the consequences expected, or (b) its likely consequences themselves being of greater benefit, or both.) Were this the case, then Y might be said to be in my best interests, although it could be said to be in my interests to do any of X, Y or Z.

Now the distinction I have drawn between acting in one's best interests and merely acting in one's interests has an artificial ring inasmuch as failing to do X, or failing to effect the consequences of doing X, may be regarded as another of the consequences of doing Y. If (the consequences of) not taking one alternative course of action counts as a consequence of taking

another course of action, then it will always be the case that the course of action which is in one's best interests and that which is in one's interests would be one and the same.

However, there would seem to be some advantage in limiting the notion of a course of action's consequences in order to retain the distinction between:
(1) a course of action's being in the agent's interests, and
(2) its being in the agent's best interests.
I shall explore the way in which this limitation might be established in the following pages. First, let us look at intuitive support which might encourage us to maintain such a distinction. Consider our case again. P's town has no railway, and limitations of money and time preclude transport by air or water. P's options appear to her to comprise her driving to Ripon (during which she could not read, although neither would she feel unwell, not being prone to car sickness) or taking the bus. But her possible options are greater. Unbeknownst to P, her neighbour, a reticent and private person who never ventures such information, is driving to Ripon in an empty car and is wishing for company. In this case, we should surely want to say both that it is in P's interests to take the bus, and thus that she would act rationally in doing so, and that it would be in P's best interests to ride with her neighbour, since doing so she could both read and feel well, and so it would be more rational to travel with the neighbour than by bus.

The distinction between "being in one's interests" and "being in one's best interests" must be maintained in order to allow us to say, as I think that our intuitions would encourage us to do, that P acted rationally in taking the bus despite her ignorance of her best interests. How, then, can we mark the distinction? It will be remembered that we must do so in order to avoid the objection raised earlier that any action satisfying our present definition for being in an agent's interests will necessarily also meet the requirements for being in that agent's best interests.

With the concept of an action's being in an agent's best

interests we have, as was indicated earlier, a criterion enabling us to ascribe maximum rationality to a course of action in a situation. That is, maximum rationality is to be accorded to any course of action which (in comparison with all possible courses of action available to the agent in that situation) is the one in the agent's best interests. It is important to notice that this would include courses of action which the agent may not have considered because he or she was not aware of them, including those of which he or she could not have been expected to have been aware.

However, while it concurs with our intuitive concept of a course of action's being the most rational one for the agent to undertake in a given situation, actions which will be in the agent's best interests, in the above sense, cannot be supposed to be the only ones which, on a purely intuitive basis, we should judge to be rational. Our intuitive concept of rationality is, after all, one of degree, and thus being the most rational course of action possible is not the only way of being rational. A characterization of rationality in terms of best interests and maximum rationality alone fails to take into account the additional courses of action to which our intuitions encourage us also to ascribe some lesser degree of rationality.

So, we must also specify the minimum conditions which must be met in order for rationality ascriptions to be made. One possible way of doing so immediately suggests itself. Corresponding to our definition of maximum rationality, we can say that the minimum conditions for rationality are met as long as the agent selects that action of all courses of action known to him or her which will most further his or her interests, or which will be in his or her interests at least as much as any other course of action known to him or her.

While this characterization provides us with a clear-cut definition, it is at variance with our intuitions. Ignorance of other possible alternative courses of action does not always work in the simple way it was shown to do in my last example. For we should not always wish to describe as rational the action of a

person who, as we should say, "ought to have" or "might have been expected to have" known of another course of action which would better have furthered his or her interests. If, for instance, P's neighbour were of a chatty turn and had mentioned her trip, her empty car, her wish for company to P, and yet P failed to be aware of the possibility of travelling with her neighbour (having forgotten the remarks, or simply having failed to connect them with her own plans), then we should hesitate to describe her action in taking the bus as even minimumally rational. Our intuitions about the rationality of pursuing a course of action are affected by purely normative considerations about the extent to which particular agents in particular situations might have been expected to have been aware of alternative courses of action. But, giving a satisfactory definition or characterization of a normative clause such as the one just mentioned is, as I shall show, impossible.

We could attempt to give an account of what the agent might be expected to be aware of by delimiting the class of alternatives of which a normal adult of average capabilities, perceptual powers, etc., would have been aware in that situation. But the difficulty with that account is that the situation must be so specified as to include the beliefs of the agent in question at the time, and thus so tailored to fit the particular case as to be unhelpful. To state what a normal person would have known given the particular situation is in no way to advance on what the agent "might have been expected to" have known. Similar difficulties beset an account in terms of something like "All those alternatives which, after a reasonable perusal of the particular situation, that person would be aware of", for we must appeal again to the normal observer to give any account of the otherwise question-begging "reasonable".

Notice that there is an additional difficulty with such normative phrases. They are ambiguous. I shall illustrate this with reference to the phrase "what a person might have been expected to have known", but it will be seen that a similar ambiguity is present in all phrases of this kind, including the one mentioned earlier, "what a person ought to have known".

"What a person might have been expected to have known" might, on the one hand:

(i) be taken to imply an epistemic judgement which could only be made by a speaker in ignorance of whether or not the person judged actually was aware of, or knew, the item in question. But, on the other hand:

(ii) the phrase may also be used to imply a normative judgement made regardless of such epistemological considerations - the judgement that the person is to be condemned if he or she does not know some item.

The latter judgement is made indifferently whether or not the speaker is ignorant of the actual situation. I may know that A is not aware of P, and yet I may still judge that A "might have been expected to have known that P", in the latter sense. Clearly, it is only in the latter way that the notion "what a person might be expected to have known" is used in the context we have been discussing.

Let me recapitulate the arguments of the last pages. Applying one or both of the definitions established earlier, we have two alternative criteria for the ascription of rationality:

(1) An action is only correctly described as rational when it is in the agent's best interests.

(2) (a) Minimally, an action is rational to the extent that it is in the interests of the agent as much as, or to a greater degree than, any other course of action known to the agent, and

(b) Maximally, that course of action which of all possible courses of action available is, in fact, the one which is in the agent's best interests is the most rational one which the agent might have undertaken.

But definition (1) is at variance with our intuitions in being over inclusive, and definition (2), more particularly (2)(a), is at variance with our intuitions in being over exclusive.

If we respect our intuitions with regard to the concept of rationality we cannot define the minimum conditions for rationality ascriptions and we must rely on undefinable normative clauses like that concerning what the agent might be expected to know. We seem thus to be faced with a dilemma. And, while concise definitions are available, as I have shown above, making use of them would mean riding roughshod over our intuitions, for they do not

completely concur with our intuitive conception of rationality.

While realizing the limitations inherent on doing so, I intend to plump for the former horn of this dilemma and to embrace the undefinable normative clause. I do so on the assumption that there is at least sufficient concurrence between the intuitions of different people as to the extent of "what the agent might have been expected to have known" for the notion not to have lost all usefulness, despite the absence of a definition for it. Thus, I shall make use of the second definition above, inserting the normative clause as follows:

(1) (a) Minimally, an action is rational to the extent that it is in the interests of the agent as much as (or to a greater degree than) any other course of action which is known to the agent, or which might have been expected to have been known to the agent, and

(b) Maximally, that course of action which of all possible courses of action is the one which is, in fact, in the agent's best interests is the most rational one which the agent might have undertaken.

There are a number of additional remarks to be made about a conception of rationality of the kind I have been discussing.

It is interesting to notice that in approaching the question of the rationality of a particular action, as judged by criteria of the sort we have been considering, when that action is one stemming from more than one distinguishable motive, two possible strategies are available to us. Let us say that person A gives person Z a large sum of money, when that action was motivated in two distinct ways:

(i) A was touched by Z's hard-luck story and wanted to help him, and
(ii) A wanted to be thought generous and liked.

Let us assume that the gift merely "pauperises" Z and leaves impulsive A short of money, so that none of Z's interests (and none of A's fiscal ones) are furthered by the action. However, at the same time, the action has the effect of gaining popularity for A, as he wished, and let us add that in this case the desired popularity is of great benefit for A. Now, as long as no alternative course of action of which A could reasonably have been expected to know would better have advanced A's interests while satisfying his desires, in order to decide whether A's interests were served

by the action of giving away the money, we need merely decide whether the good gained by making A popular could be expected to outweigh the harm caused by 'helping' Z. And, it might well be the case that because the likelihood of benefit would outweigh the likelihood of harm, the action might be said to be a rational one, despite its only being rational, strictly speaking, when it is regarded in the light of one of the motives occasioning it. It is still true to say X acted rationally, although the action was not very rational.

However, as I have said, an alternative strategy presents itself here: that of assessing the rationality of the action when regarded in the light of each of the different motives and, rather than summing the two, assessing the rationality of the action under each separate description. In accordance with such a strategy, we should have to say that while it was rational to give money away to gain popularity, it was not rational to attempt to help Z by giving him money. And, indeed, this would be a natural and common way of answering a question about the rationality or otherwise of the action in question.

The latter strategy would lead to a more stringent rationality requirement than the former one, for it may be stipulated that only when under each separate description, considered separately, the action is in the agent's interests can we describe the action as a rational one. I shall return to the significance of the difference between these two strategies.

From the way we speak it may be supposed that an action's rationality is sometimes regarded as a function, in part at least, of the efficiency with which it achieves the goal for which it was undertaken, as distinct from its likelihood of doing so.[8] Thus, if I want X and have an equal likelihood of achieving it by method A and by method B, when method A is a more efficient way of achieving X, then, all other things being equal, it would seem to be more rational to choose method A than to choose method B. But this is not, I think, because efficiency is an additional criterion by which we can judge the rationality of some behaviour; rather, its being so follows as a consequence from what has already been

said concerning the concept of being in a person's interests. For one course of action to be judged likely to be more efficient than another, it must be the case that it is more economical of effort or expenditure in some respect. And any unnecessary expenditure of effort would, of course, be seen as an undesirable consequence of a proposed course of action. Thus, ceteris paribus, a relatively inefficient course of action will never be in any of a person's interests to the extent that an efficient course of action will be - so it will not be in that person's interests simpliciter.

What is the limiting case of an action which is likely to benefit the person and hence to be in that person's interests? Does anything which is not not in one's interests count, albeit weakly, as benefitting one and hence as being, ceteris paribus, in one's interests? On the surface of things, we might judge not: there would appear to be a neutral category of actions which neither substantially harm nor substantially help us. The example given earlier of smiling during a conversation sometimes would count as such an action. It may be said neither to be likely to harm nor to benefit. But smiling perhaps works here on account of, in the kind of case proposed, lacking any important consequences at all. For it would seem that if an action is likely to have any relevant substantial consequences (that is, is likely to affect either subsequent events in the agent's or any other person's life, or their mental states, other than their simplest beliefs about the action's occurrence), then it must either be in the agent's interests or not.

In addition, certain actions often are considered to be intrinsically beneficial or harmful, regardless of the likelihood of their having substantially beneficial or harmful consequences. This, as was remarked earlier, is either because they are seen as having intrinsic merit or demerit as actions or, on another theory, because they are seen as instrumentally good or harmful in unfailingly producing states of mind which are intrinsically good or bad. (There is, as might be expected, little agreement on the limits of the classes of actions considered to have this intrinsic merit or demerit, although there is perhaps sufficient agreement over the

central cases to justify reference to them in such an account as this.)

So, we can describe the class of actions which are neither in a person's interests nor not in his or her interests, as that class of actions which:
(a) are neither intrinsically beneficial nor harmful, and which
(b) are unlikely to have substantial consequences.

Only slightly different from the characterization in terms of interests are some of the suggestions of modern decision theory. For example, according to one such account, an action is regarded as rational or optimal in the precise degree to which it maximizes the agent's expected utility - when the expected utility which a course of action has is a function of the degree of desirability of its consequences, that is, the extent to which they would satisfy the agent's desires, on the one hand, and the likelihood of those consequences occurring, on the other. (Thus, the course of action which would maximize the agent's expected utility would be that one which was calculated to provide the greatest likelihood of desired consequences occurring.) Alternatively, an action is sometimes said to be rational if and only if it has <u>at least as much expected utility</u> as has any alternative course of action.[9] Another suggestion,[10] embodying the cautious intuitive principle of expecting the worst, is the so-called <u>minimax strategy</u>, by which the most rational course of action would be said to be that one whose worst consequences are likely to be the <u>least</u> bad possible. And, according to a fourth account propounded by decision theorists,[11] a person may be said to act rationally in pursuing a course of action just when it is that course which will minimize the <u>regret</u> arising from making one decision rather than another.[12]

Notice that the decision theorist's formulations all rest on the assumption that it has first been estimated what the agent <u>most</u> wants, desires, abhors, would regret, etc. My analysis of the concept of some action's being in a person's interests shows that <u>what would benefit the agent</u> cannot as such be taken to refer to what would most benefit the agent or what would be in the agent's

interests simpliciter (nor, of course, what would be in his or her best interests). And, a similar qualification must be made to the decision theorists' formulations. Merely what the agent desires, or what would satisfy the agent must be interpreted to read "what the agent most desires" or "what would most satisfy" - for, as it stands, that expression allows of the possibility that the agent might desire X but also desire Y when X and Y and incompatible and he or she desired Y more than X. And, in certain cases, we should be loath to ascribe even the minimum degree of rationality to a person acting upon a desire which was not, all things considered, what he or she most desired, where the action was one from the range of possible actions open to the agent.

Now, in one way, the decision theorists present not so much a concept of rationality different from the one introduced in my version of Hobbes's thesis as an accurate method of calculating rationality according to that concept. It is possible, for instance, to substitute "benefitting the agent" for "satisfying the agent's desires" and, with certain qualifications described below, to have in the first two decision theorists' formulation described previously a method for calculating which among a number of strategies would most likely further the agent's interests. Similarly, it might be said that the minimax strategy involves choosing that action whose worst consequences least hindered the furtherance of the agent's interests, and that the concept of the agent's regret is to be interpreted as a function of the failure of an outcome to further the agent's interests.

Yet, between the decision theorists' approaches and the one I have been discussing there is one significant difference. It must again be stressed that there is a strongly evaluative element in the characterization of rationality in terms of interests, and one which the decision theorists' accounts more easily avoid. Because the concept of interests is a partially evaluative one some outcomes of our actions, regardless of whether they satisfy the agent in being the outcomes which were desired, would not be seen as having been beneficial to the agent, and hence as having been in the agent's interests. Since we can have and act upon desires which are contrary to our interests, there may be actions which, on

an analysis of rational action in terms of interests, cannot be judged to be rational. A person's accepting a promotion was shown to be likely to further her monetary interests, yet we concluded that, nevertheless, it would not be judged to be in that person's interests simpliciter in any degree. While accepting that the promotion would be "in her interests", in one sense, we must admit that in a more serious way, it would not be at all: it would not be in her overall interests.

In contrast, the decision theory account of rationality is most naturally expressed in a value-neutral way. On it an action's rationality may be evaluated entirely on the basis of the relation between aim and means of achieving it, so that however it may fail to further the agent's interests, an action still may be judged to be rational if it satisfies the agent's wishes. Thus, on the decision theory analysis, if the agent desired greater financial benefits, then to the extent that he or she used a likely means in acting upon that desire (by seeking promotion), she might be considered to have acted rationally.

The decision theorists' account need not be expressed in a value-neutral way. A limitation can be put on the expressions 'desire', 'satisfaction', 'regret', etc. It may be said to be impossible truely to desire or be satisfied by certain ends. But the decision theorists' formulations lend themselves more readily to a value-neutral interpretation, nevertheless.

Although the formulation in terms of interests, expected utility and regret, respectively, capture one characterization of the notion of rationality, there is a common refinement of our intuitive notion which none of the above kinds of formulation adequately covers.

The disparity between at least one of our intuitive notions of rationality and the notion expressed in the characterizations discussed so far may best be illustrated by an example. To return to my earlier case, let us suppose that P began a conversation with her reticent and private neighbour and, not knowing of her neighbour's intention to drive there, mentioned her wish to travel

to Ripon. Because of the neighbour's desire for company on her trip, we may further suppose that P's doing so would prove to have been in P's interests by resulting in her being offered a ride. But, while we may admit that it turned out to have been a rational thing for P to do, we should not perhaps unqualifiedly wish to describe P as having acted rationally in speaking to her neighbour unless it were also true of her that she knew of her neighbour's intentions and wishes. The addition to the characterizations of rationality expressed earlier which this suggests is one to the effect that sometimes, to be able to say that the person "acted rationally", the action must not merely further a person's interests, it must also be believed to do so by the agent at the time of acting.

Nor, it might be argued, is the above requirement quite sufficient. P might hold the true belief that speaking to her neighbour would be in her interests and yet her belief might not be one which she can be said to be justified in holding. Even if she were able correctly to discern the manner in which her interests would be furthered by her action and yet was unjustified in holding this belief - doing so, let us say, on the basis of a false assessment of her own ability to charm her neighbour into lending her car - then we should hesitate to accord rationality to her action. Not only must the agent believe his or her action to be in his or her interests; that belief must be justified.

We can say, roughly, what it is that makes a belief "justified". It is:
(i) held on the basis of what would generally be regarded as good grounds, and
(ii) when it is inferentially reached; it is reached on the basis of what would be regarded as sound inference - inference which is in accord with the laws of logic[13] and is such that its premisses are drawn from all the available evidence.

Notice the form of "on the basis of" in (i) and (ii) above. It is possible to have good grounds for a particular belief without holding that belief because of those grounds. I might

believe that doing X would be in my interests and I might have good grounds for my belief in the sense that I might be aware of strong evidence which could be adduced in favour of it, and yet I might not believe it because of that evidence. Clearly, "on the basis of" must, in the above account, be read as "because of".

But, although it is possible to cite examples of justified and unjustified beliefs, difficulties arise paralleling those encountered already, when we attempt to tighten up these specifications or to rid them of their normative clause. One part of requirement (ii) above is quite exact. We have rules for the soundness of an inference, viz. the laws of logic. We have no such formal rules to guide us in specifying what counts as "all the evidence" upon which the inference is supposed to be based, except something rough and normative (and hence, circular) like the following: all the evidence which a moderately intelligent person with normally functioning perceptual equipment and an average amount of general background information would be likely to ascertain as relevant. It is at once obvious that the phrase "would be likely to" suffers from all the drawbacks earlier noted in "might be expected to". With (i) there is even less to go upon. The innumerable ways in which beliefs, hunches, doubts, conjectures, etc., can stand in relation to further beliefs in rendering the latter "justified", simply defy classification. We are thrust back upon an account which is both vague and, again, normative. And, although some most surely are, not all non-inferential beliefs are well grounded either. (Beliefs derived from extra-sensory perception, for example, are not usually regarded as such.) What makes a non-inferential belief count as well grounded is simply that most perceivers would be inclined to hold the same non-inferential belief under the same conditions. So, a non-normative and thus a non-circular account of what makes a belief "justified" is not available.[14] And, in defence of a characterization of rationality which employs the concept of "justification", I can do no more than point, as I did in my earlier discussion of the normative element in an account of rationality in terms of interests or expected utility, to the possibility of widespread agreements in judgement between users of the intuitive concept.

Even justified beliefs, notice, may be false. And, it is not necessary for the agent to know that his or her action is in his or her interests for it to be described as a rational one. What it is rational to do is not necessarily what eventually turns out to have been in anyone's interests. An unforeseen (and unforesseeable) event may mean that the course of action apparently least likely to, actually would most have benefitted the agent, but the agent, in choosing that action still would have exhibited a lack of rationality according to the characterization under discussion as long as its likelihood of bringing benefit was to him or her unknown, or was believed without justification.

It is possible for a person to do X knowing it to be in his or her interests without doing so, in any sense, because it was in those interests. That is, it is possible knowingly to act in one's interests without one's knowledge being any part of one's reason for so acting. In addition, we seem to want to make a rough intuitive distinction between:

(a) doing X when X is in one's interests and one knows it to be, where one's knowledge is merely a necessary condition of one's acting, in the sense that unless one knew it to have been in one's interests, one would not have taken the action, and

(b) doing X, as we sometimes say, 'because' it is in those interests,* when the knowledge that the action is in one's interests may be seen as the main reason or as one of the important reasons for the action.

---

* An ambiguity in my formal use of 'because'in the phrase 'doing X because it is in those interests' confuses the issue here. I have discussed already (page 104) limitations on the use of the expression 'reason' and 'because'. As was remarked in that discussion, there is a certain ambiguity in 'because'. If the expression is interpreted to cover the relation between an action and any condition necessary for its occurrence, then the stricter formulation would be more correct, and it rightly would be said that in order to act rationally the agent must not only believe that it would be in his or her interests to do X, also he or she must do X because of that belief. Whereas, if we restrict 'because' in accordance with our common, although ultimately unjustified restriction of 'reason', then the consideration because of which a person acts is just that person's main reason for acting.

It would seem that the former case is sufficient for an ascription of rationality. To require that cases of the kind described in (b) alone be classed as exhibiting rationality above would be to introduce a rationality requirement altogether too stringent. A person may act rationally for reasons other than his or her belief that the likelihood of the proposed course of action's being in his or her interests. Person P might enter medical school primarily to gratify her father's dying wish that she do so; she may also hold the justified belief that entering medical school would be in her interests. And we should not regard her as less than rational, it would seem, if this belief did not enter her decision in the sense of being an important factor in it. And yet, if it did not play some part - if it were not at least a necessary condition of her choice, in the manner described in (a) above - then we should, I think, be inclined to describe her action as less than rational.

Clearly, in order to meet the intuitive notion of what counts as "acting rationally" - which I have been discussing - the same kind of qualifications made to the account of rationality in terms of interests would equally apply to a definition of rationality of the kind given by decision theorists. For example, in order to say that the agent acted rationally in doing X, we should need to say not only that X maximized the agent's expected utility, but also that the agent believed that X was likely to maximize his or her expected utility, and was "justified" in holding that belief, in the requisite sense. And, similarly, it would seem unnecessary in this case, as in the case of rationality criteria in terms of the agent's interests, to postulate that the said belief played an important part in bringing about the action. Without these qualifications, in speaking of the agent who unwittingly happened upon the course of action which maximized his or her expected utility, we should not, I think, unqualifiedly describe him or her as having acted rationally.

(2) Rationality as Having Reasons

The preceding discussion has served to introduce a further way of characterizing "rational" which I have not thus far considered. Sometimes "rational" conveys "having reasons". This characterization may be seen as a way of formulating the notion of rational action as action which is goal-directed, or purposive - action or behaviour for which there is a "rationale". To have had reasons in acting in this sense is merely to have acted with some aim, goal or purpose, or (subject to a qualification mentioned below) to have been 'motivated' in the sense discussed earlier, in so acting.

Notice that having a reason in this sense does seem to suggest having a motive, rather than having a mere whim. Although, according to the terminology introduced in Chapter Two, merely wanting to do X, or feeling like doing X, would count as having a reason for doing X, in the way "having a reason" is here introduced, the reason must be more substantial than that. Notice, too, that having a reason or having been motivated in acting here entails believing one's action to be a way of bringing about the end or goal desired. In discussing the concept of motivation in Chapter Two (Page 47) two models of motivation were distinguished. It was remarked that to describe behaviour as motivated sometimes is to imply, in addition to the presence of some want or desire, a certain belief about the relation between the behaviour in question and the object(s) of the aforementioned want or desire which make up the motive for the action: viz. the belief that the action or behaviour undertaken is a way of producing the state of affairs wanted or desired. It was argued that not all cases of motivated behaviour naturally fit the purposive model. However, when we describe behaviour as "rational" on the grounds that it is motivated or purposive, it is strictly this model which is presupposed. To say that its rationality rests on some behaviour's being motivated is to imply in every case the presence of beliefs such as the one above.

In their discussions of the supposed rationality of primitive behaviour, ritual and religious rite, it is often to this second sense or use of "rational" which sociologists and anthropologists

refer.[15] To say that some initially odd and unintelligible action nevertheless was rational is to say that it was done with some purpose or aim, and that when it is seen in the light of that purpose or aim, it becomes intelligible to us.

Notice that in the general sense of having a reason described above, there seems to be no particular implication as to the agent's awareness of, or ability to give, the reason in question. This, of course, concurs with the conclusion which we reached in Chapter Three: there are reasons for actions such as rituals, of all or some of which reasons the participants are not aware.

It is also interesting that in the characterization of rationality under discussion no qualitative restriction appears to be placed on the kinds of beliefs and desires which can count as reasons. In decrying the myth of the "savage mind" and insisting upon the "rationality" of praying for a good harvest,[16] the anthropologist is not to be regarded as committed to the belief that it is reasonable to hold that prayers affect the crop. All that the anthropologist asserts is that given this (unreasonable) assumption, we can regard the action as a rational way of achieving the desired end.

(3) Rationality as Reasonableness and Being Justified

In a refinement of the notion of rationality described above, proceeding rationally sometimes is taken to mean not merely having some reason(s) for acting or proceeding but doing so on the basis of good reasons. Sometimes we would allow that a person's action, belief or choice was irrational or unreasonable, despite their having been able to give a reason for it, and this would be when we regarded the reason cited as less than adequate as a reason.

Appeal was made to the latter conception of rationality when I was discussing the requirements for an agent's being justified in believing some course of action to be in his or her interests or to be likely to maximize his or her expected utility. Although, as was previously remarked (Page 162), the desire to further one's interests or expected utility may not be the sole,

or even a prominent or important aspect of the agent's reason for acting, nevertheless to have such a desire and a belief that some course of action would further one's interests and/or expected utility would always be to have a reason for acting, in the sense described on page 163; and to be justified in holding the latter belief would be to have a good reason for acting in the sense under discussion.

Notice that the notion of acting "reasonably" now has entered the discussion. It is impossible here to do justice to all the interconnections, similarities and dissimilarities uniting and dividing the expressions 'rational' and 'reasonable', 'irrational' and 'unreasonable'.[17] However, it is important to point out two things before we turn to a closer analysis of what makes a reason a good reason. The first is that it seems to be the conception of "rational" just described which corresponds and may be said to be roughly synonymous with the concept "reasonable" when the latter is used of an action. To fail to act rationally by acting for reasons judged not to be good reasons is to act unreasonably. Unlike "rational", "reasonable" is never used of an action merely because there were reasons for it; "reasonable" and "unreasonable" are wedded to the evaluative context.*

The second point is that the concepts "reasonable" and "unreasonable", unlike "rational" and "irrational", are ascribed equally naturally to beliefs and desires themselves, and to the holding of beliefs and desires, as well as to actions or behaviour. And, acting "reasonably" or acting for what we have called good reasons, is just acting upon beliefs and desires which may themselves also be described as "reasonable". Although not a natural way to speak, it seems possible to evaluate the forming or holding of beliefs and desires as "rational" to the extent that they meet the requirements for being said to be "reasonable".

There are several different kinds of requirement for something's counting as a good reason for an action and I shall deal with them in turn. First, the beliefs and desires comprising the reason for the action must themselves be reasonable or justified

---

* And, since it is only in one of the several ways in which the notion of "rationality" is characterized that it is interchangeable with "reasonable", I do not accept Peter Alexander's assumption that they are simple synonyms.[18]

and must meet certain standards of coherence in relation to the agent's other beliefs and desires. I have discussed already the requirements which must be met for a belief to be said to be justified. Further, I have remarked that the process by which we form a belief, just as the acting upon it, can be evaluated as reasonable or not. A belief which is reasonable, or one which it is said to be reasonable to hold, is a belief which, according to my earlier definition, is justified, that is, it is:
(i) held on the basis of what would generally be regarded as good grounds, and
(ii) when it is inferentially reached, it is reached on the basis of what would be regarded as sound inference - inference which is in accord with the laws of logic and whose premisses are drawn from all the available evidence.

Beliefs themselves sometimes are also said to be reasonable or unreasonable to the extent that they are true or that the propositions they express are likely to be true. The latter of these formulations at least is linked with the degree of reasonableness with which the beliefs are held. The degree of reasonableness with which a given belief is held may be said to be a function of the likelihood of that belief's being true. As I remarked earlier, a belief which the agent is justified in holding may be false. Similarly, it can be said to be reasonable for P to believe X even when X is false (if P has good reasons for doing so). But is P's belief a reasonable one? And can it be said that there is reason to believe X?

Now the notion "reasonable" and the notion of there being reason to believe something both seem to wobble unhelpfully here between characterizing an evaluation relative to the actual facts, and one relative to the beliefs which a person might be justified in holding - generating, in doing so, the paradoxical suggestion that there both can be said to be and not to be a reason for holding some belief. Because of this ambiguity in our ways of speaking, we seem to have no recourse but to arbitrate that just as it is possible to hold a justified belief which is false, so it must be possible reasonably to hold a false belief which there can be said to be reason to hold.

Some desires, also, can be said to be reached by a process which may be evaluated according to its reasonableness. Wants and desires which are passions or affections are not susceptible to such an evaluation. But, as I discussed in Chapter Two, some desires must be regarded as the product of a process of deliberation and reasoning. My wish to order steak for dinner might be an affection, an inexplicable whim which enters my mind without preliminary, but alternatively, it might be the carefully thought-out conclusion of a piece of conscious reasoning concerning my tiredness, my anaemic tendencies and the nature of iron deficiencies. And, just as we did with the reasonableness of belief, so we can say roughly what constitutes a reasonably held desire: it is one which is reached by way of a sound inference from all the available evidence.

A person proceeds rationally when his or her reasons for so proceeding both:

1. comprise only rationally held or "reasonable" beliefs and desires in the sense described, and
2. are coherent, in the sense of being consistent with that person's other beliefs, values and attitudes.

If, in choosing to buy one rather than another musical instrument, I cite as the reason for doing so the fact that the chosen instrument is pitched lower, then my choice would have been less than rational if, on the one hand, my evidence for the difference in pitch was from a source which any normal person might be expected to have known to be unreliable - for example, a confessedly tone-deaf salesperson, an obviously out-of-date manual, etc. - or if, on the other hand, I held some belief which conflicted with this choice, for example, if I believed that all things being equal, it was better to play higher-pitched instruments than lower-pitched ones as there is a more extensive repetoire available for them, (or if I had some taste, attitude, or value which was at variance with it, for example, if I preferred the sound of the higher-pitched instruments - assuming that I had no other overriding reason for favouring a lower-pitched instrument in this case).

There is, however, a qualification necessary for (2): A person may be said to proceed rationally on the basis of a reasonable

desire which is not consistent with that person's other beliefs and desires, but only if:

(a) the inconsistency is recognized by the agent, and

(b) the agent judges the desire upon which he or she proceeds to be the more rational or reasonable one.

Several different sorts of case prompt this qualification to requirement (2). One is when the said incompatibility of beliefs, attitudes and values is induced by the nature of the situation. For example, at the scene of a house fire, I might have a reasonable desire to rescue P, together with a reasonable desire to rescue Q, when time would permit me to act upon only one, and for some reason (for example, that P is a child, or that I am more likely to be successful in rescuing P), I might choose to rescue P, and act rationally in doing so. The incompatibility may not merely be a product of the situation, however: it may be intrinsic to the beliefs, values, attitudes or desires themselves. For example, I may have a reasonable desire for good health and, yet, I might give in to my desire (also, perhaps, reasonable) to eat stodgy cakes. But in neither case described, unless I had explicitly acknowledged the incompatibility of the two desires to myself and had decided to act upon what I judged to be the more rational one (in the sense of being the desire for which there seemed to me to be stronger reasons), should I be said to be acting rationally.

One further important observation must be made before we leave the questions of the consistency and coherence of the rational agent's beliefs and desires. It seems to go without saying that the above-mentioned requirement for a degree of consistency and coherence among the agent's beliefs and desires would be reflected in a similar consistency in that agent's actions. If, save for the qualifications described, a person acts on the basis of a consistent set of beliefs and desires, then we need only assume that that person's behaviour is an accurate reflection of his or her beliefs and desires to be assured of its consistency. And the above assumption is one to which we are entitled as long as we stipulate that the rational agent acts out of his or her beliefs and desires.

A second important point with regard to this notion of "rational" is that it must be the case not only that the agent believes that his or her action is likely to bring about his or her aim or goal (a requirement built in, as was remarked earlier, when we describe the action as one which is motivated or goal-directed) but that such a belief is itself justified.

Sometimes it has been argued that, in order for the latter stipulation accurately to convey what we mean by acting rationally or reasonably, it must further be the case that the action taken is the one most likely to achieve the state of affairs desired, or at least one as likely as any other possible course of action to do so. But, for reasons given earlier (pages 149-150), this requirement would seem to be over-stringent. While it would always be more rational or reasonable, all other things being equal, to choose the course of action most likely to bring about the agent's desired goal, still it seems that we should be prepared to accord some degree of rationality to the person choosing any action provided that it had some likelihood of achieving the goal (and the agent had good grounds for believing so), and provided, of course, that the agent did not know of and would not have been expected to know of the greater likelihood of achieving the goal being associated with the other course of action. If I have been told by a reliable source that some horse X will win a race, and I know of no reason to suppose any other horse to be likely to be superior, nor any other source on these matters which has in the past been as reliable, then I may be said to act rationally in putting my money on X, even if Y is more likely to win.

Notice that the considerations involved in this judgement exactly parallel those discussed earlier in relation to the extent to which an action is said to be in an agent's interests. For this reason, we need not concern ourselves this question any longer.

There is a third important point which must be made about the conception of rational action which I have been discussing. The notion "having a reason" seems now to be used in such a way

as to imply that the agent is aware of, or capable of becoming aware of, and of giving at least some of his or her reasons for acting. Part of what is meant by the term 'acting reasonably' is being able to give one's (good) reasons. Similarly, part of what is meant by the phrase 'being a reasonable person' is not only being guided by good reasons but being able to report, discuss, explain, justify and defend them.

If my intuitions are correct about this evaluative sense of the concept of rationality, then it follows, of course, that one might act upon good reasons without yet being said to act reasonably or rationally. And actions of the sort introduced in Chapter Three as counter examples to the thesis that it is of the nature of reasons to be known to, and communicable by, the agent - the actions of the mute and paralyzed person, who is incapable of communicating his or her reasons for acting, and the skill of the composer picker who can give or recognize no reason for his or her choices - while done for reasons which may be judged to be good ones, could not be said to be cases of "acting reasonably".

This appears to follow because an action's being said to be "reasonable" or "unreasonable" strongly suggests that a relatively leisurely, introspective process of reasoning and deliberation - a conscious process of comparing, weighing and choosing - occurred in the agent's mind before the action, and could accurately be described by the agent. However, the model here is misleading. The picture is one of a person who with full attention deliberates over the reasons for and against acting, forms his or her decision to act according to the principles of reasoning described earlier and then, in subsequent reason-giving reports on the process. But, while perhaps an accurate characterization of the paradigm case of a person's acting reasonably, this picture cannot be taken to express the conditions necessary for a description of a person's so acting.

It is true that giving (good) reasons for one's actions sometimes is akin to reporting an inner process rather the way in which we might report a series of sensations. But this fact,

coupled with our intuitive capacity sometimes to make a distinction between, on the one hand, knowing of good reasons for doing X and citing them and, on the other hand, doing X for good reasons which one knows and cites - a distinction which has been described as that between "justifying one's behaviour" and "explaining" it[19] - combine in inviting us to lay stress on the wrong capacity when we attempt to give an account of the nature of the reason-giving necessary in this sense of rationality.

In order for his or her action to count as rational in this sense, or as reasonable, the agent must be able to give or to think of good reasons for it. And it must be the case that those considerations did move him or her to act - that they were necessary conditions for the action's occurrence. But is not further necessary that the agent acknowledges them to be the considerations which seemed to move him or her to act.

It is possible at times to make a distinction of the sort described above. Sometimes, perhaps even most often when we cite a reason for an action, we can distinguish whether or not we regard it as one which would have been a good reason for so acting, or as one which not only would have been a good reason, but was the one which actually moved us to act. But, while it is in some cases and in the case of some reasons possible to make such a distinction, the thesis that we consistently and unfailingly can do so loses all plausibility once we question the notion that reasons for actions are limited to the discrete and finite desirability characterization(s) sometimes described as the primary or main reason(s). As was shown in Chapter Three, it is not possible to distinguish among the innumerable beliefs and desires which jointly are necessary for the carrying out of any particular action and thus the reason for an action, in the sense of the complete reason, could hardly be conceived in terms of a kind of a small number of discrete mental items which are objects of awareness in the way in which sensations are. Yet, only if "reason" were able to be so circumscribed would it be at all likely that we should enjoy (in relation to all of our reasons for acting) the intuitive capacity described earlier.

In fact, we could not and do not enjoy any such capacity. To return to the example used in Chapter Three (Page 104), if part of my reason for insulting X is that I believe that 'pianist' means one who plays the piano, my regarding it as part of my reason is not based on some special connection which I sense between that belief and my action. Rather, it is due to my taking it to be part of the set of connected syllogisms which can be traced from my action and other beliefs and desires comprising the reason. But, awareness of the latter relation is equally present when I cite what I take to be what would have been good reasons for my actions, that is, when I merely "justify" (in retrospect) what I do.

So, we must conclude that despite the purported distinction between those reasons we regard ourselves as having acted upon and those we merely regard as possible reasons for acting, all that actually is necessary for an action to be described as reasonable is that the agent is capable of citing good reasons for having so acted (and acts, of course, sincerely in doing so), and that he or she actually was moved by those reasons.

Rationality: A Summary

We might now sum up the different connotations of the expression 'rational'. There are three ways of conceiving of rationality to which we appeal when we speak of people and their capacities as rational:

(A) Being rational as having the human power of reason;

(B) Being rational as having developed and well-functioning reasoning powers;

(C) Being rational as being normal in respect of beliefs, reactions, etc.

Then there are ways of accounting for 'rational' as it is used in the notion of particular actions being said to exhibit rationality:

(1) Acting rationally as acting in such a way that one furthers one's interests, and acting rationally as acting in such a way that one maximizes one's expected utility (minimizes expected undesirability, minimizes regret, etc.).

(Although, as I have shown, there are differences between the two conceptions of rationality grouped together under (1), they nevertheless reflect the same kind of reconstruction of rationality.)

(2) Acting rationally as having reasons in the sense of a rationale or purpose in acting.

(3) Acting rationally as acting reasonably, that is, having and being moved by good reasons in acting.

Notice that (3) includes being justified in believing one's action(s) to further one's interests and/or expected utility.

Notice also that the use of the term 'rational' in the three people-related senses above (A-C) would not always be compatible with the use of the expression indicated in (1), (2) and (3). A person and/or his beliefs or actions might meet all of the conditions expressed in (1)-(3) and yet be described as irrational merely on account of being judged to be insan (and hence the victim of malfunctioning faculties, or abnormal beliefs and inclinations). Similarly a person's action might meet none of (1), (2) or (3) and, yet, it might be described as rational, if the contrast which is being drawn is the one between creatures born with the capacity for reason and those without.

Irrationality

Before leaving this area, let us look again at the class of things described as "irrational". I have referred briefly to that class as comprising those states which are traditionally characterized as affections or passions - emotions, moods and some desires - as distinct from voluntary actions (both mental and physical).

Note that, just as we can describe (the holding or reaching of) beliefs and (some) desires themselves, and their outcomes in action, as rational, so in making judgements of irrationality, we refer to as irrational:

(a) the emotion or feeling itself (for example, "Love is irrational");

(b) the behaviour which the passion produces (for example, "She acted irrationally because she was so upset"); as well as

(c) the procedure which led up to the having of that state: "She had an irrational fear of X" sometimes means that the fear is ungrounded, that is, she has no good reasons for feeling afraid.

In every way, the meaning of "irrational" is as complex and difficult as that of "rational". Very often, to say that passions are themselves irrational, are had irrationally, or produce irrational behaviour, is merely to say that they are not rational in the strictly normative sense described earlier (Page 145), that is, "irrational" is devalued to mean "abnormal" or "unusual". Much more than the class of voluntary actions (including those mental actions traditionally classified as voluntary belief(s)), the class of passions or affections is subject to descriptions based upon this solely normative distinction. We apply to them predicates like "disproportionate", "excessive", "uncalled for" - as well as "unjustified", "unwarranted" and "irrational", and sometimes seem to mean nothing more than "abnormal".

But we also treat "irrational" as synonymous with "involuntary" so that if a mental state or behaviour arising from a mental state, is a passion, something which happens to us, then it is said to be "irrational". Nor are passions and the behaviour which they occasion the only things to which irrationality is ascribed. From what has already been said concerning the concept of rationality, it is apparent that there are additional ways in which we may behave, come to believe, decide, etc., irrationally - viz. by failing to meet one of (1), (2) or (3). But, we must distinguish the case of a person who fails to proceed rationally according to one or some of (1), (2) or (3) though acting quite voluntarily (he or she could have acted rationally, but did not), from the case of a person who cannot proceed rationally because he or she is not in control of his or her feelings or behaviour (he or she could not have acted rationally).

(i) We may classify the strictly normative sense of irrationality - where it means "abnormal" - as irrational_1;

(ii) We can say that a person may be irrational_2 - where it means "non-rational" in failing to be rational despite the fact that he or she could have been;

(iii) We might call a person irrational_3 - what might be called "a-rational" - in failing to be rational because he or she proceeds involuntarily.

With regard to the characterizations of rationality expressed in (1), (2) and (3), notice that only (2) may be expected to pertain

to $\underline{\text{irrational}}_3$. In order for a person to proceed rationally, it must be appropriate to cite reasons for his or her procedure, and not merely causes. And, inasmuch as we proceed involuntarily, our behaviour is describable in purely causal terms, and fits into a causal explanatory framework. Similarly, by failing to meet one aspect of (3), a person may act $\underline{\text{irrationally}}_3$; for example, if a person gave R as her reason for doing A but was actually moved to so act by the effect of a drug which had been administered to her, then while not acting $\underline{\text{irrationally}}_2$ that person would have been acting $\underline{\text{irrationally}}_3$.

In summary, a person might be described as acting irrationally in exhibiting merely abnormal behaviour (irrational_1); in failing to act in such a way as to further his or her interests (expected utility, etc.), in failing to act with a reason or rationale, or acting on reasons which could not be described as good reasons (irrational_2); or in proceeding involuntarily (irrational_3).

Notes and References

1 Black, M., "Reasonableness". Reprinted in Peters, Dearden and Hirst, eds., Education and Development of Reason (Routledge and Kegan Paul, London and Boston, 1972).

2 It is this sense, of course, which Bennett explores (Rationality, Routledge and Kegan Paul, London 1964). See, for instance, his definition of rationality, as:

> "Whatever it is that humans possess which marks them off, in respect of intellectual capacity, sharply and importantly from all other known speciés."

(Page 5)

3 Interestingly, we do not usually say of children that they are irrational, but merely that they are not (yet) rational.

4 Fingarette explores and attempts to justify this strictly and explicitly normative sense in a recent book (The Meaning of Criminal Insanity, University of California Press, 1972). Note, for instance, his claim that:

> "Conduct is insane, crazy, mad, irrational when it is not shaped in the light of certain norms.....(not only) of correct inference or valid argument; there are also norms regarding what emotions or moods, or attitudes or desires are in some sense suitable or proper with respect to certain other aspects of one's situation. Clearly, there is much room for variety here.....but there are limits. These are the limits that distinguish the irrational and intelligible from the rational."

(Page 183)

5 The full quotation from Hobbes is taken from The Elements of Law, Chapter One, Section 6, where he states that:

> ".....it is not against reason, that a man doth all that he can to preserve his own body and limbs both from death and pain."

6 The traditional formulation of the relation between the concept of interests and that of rationality does not exhaust the ways in which these two concepts have been linked. There remains the possibility, in Sigwick's terminology, of Rational Benevolence, as well as Rational Egotism, (Methods of Ethics, MacMillan, N.Y., 1874); that is, it is arguable that at the very least acting so as to further another's interests as well as one's own may also be acting rationally - it may even be acting more rationally than merely acting so as to further one's own interests.

A more extreme version of this thesis, recently proposed by Nagel (The Possibility of Altruism, Oxford, 1970), is that acting so as to further others' interests may be the only rational way to act.

However, the normative issues arising out of a comparison of the relative rationality of acts of prudence and altruism need not concern us here. What we are attempting to elucidate is the established intuitive concept of "rationality" and, for better or worse, it does seem to remain firmly rooted in the theory of psychological egoism from whence it originated.

7 For further discussion of this point and, more generally, the concept of interests and what it is for a course of action to be in a person's interests, see L.R.V. Burwood's (unpublished) Oxford B.Phil. thesis: Interests, Morality and Action, 1973.

8 This claim has actually been made explicitly at times, for example, by Talcott Parsons, The Structure of Social Action, London, 1937, in defining rationality.

9 For example, Hempel's definition of rational action (from "Rational Action", in Proceedings of the American Philosophical Association, 1962):

> "An action is rational for X iff it has as much expected utility as any alternative X has, where the expected utility of an action is obtained by
>
> (a) multiplying, for each subjectively possible outcome of the action its subjective probability and its utility, and
>
> (b) adding these products."

10 Von Neumann. See Von Neumann and Morganstern, Theory of Games and Economic Behaviour, Princeton, 1944.

11 Savage, "The Theory of Statistical Decision", J. Amer. Statistical Assoc., 1951.

12 Similar as they may sound, decision theorists' formulae such as those mentioned embody, as has been shown (Milnor, "Games Against Nature", reprinted in Thrall, Coombs and Davis, eds., Decision Processes, Wiley, N.Y., 1954), different ones of the characteristics suggested in our purely intuitive conceptions of rationality, and none embodies all.

13 Which must be interpreted broadly to include avoidance of the various, informal, ignoratio elenchi fallacies, such as the fallacy of non sequitur, as well as adherence to procedures which are formally valid.

14 (a) One of the few attempts at such a non-normative specification is that offered by Mullane, in the context of his paper mentioned previously on the rationality of unconsciously motivated behaviour. Mullane sets out cases which he takes to illustrate unequivocally the difference between justified and unjustified beliefs, respectively:

"My turning the ignition key (even when, unbeknownst to me the car has no battery), is obviously rational because my relevant beliefs, though false, are justified - I have

> perfectly good reasons for believing that they are true. My believing that God opens liquor cabinets on request is, given my culture and my experiences, clearly unjustified, and so my imploring Him to do so is irrational."

and he states what he admits to be a rough rule for deciding on the rationality or irrationality of someone's belief; a false belief is unjustified if the person who holds it has had sufficient opportunity to know better. That this criterion is insufficient is apparent in the case of the hypnotist's subject who offers ingenious and elaborate rationalizations for carrying out instructions which were suggested to him under hypnosis, the memory of which experience has been erased. For it seems to work against our intuitive judgement by confirming that the belief upon which the rationalization is based, is justified: the hypnotist's subject, having been in a state of unconsciousness at the time at which he might have come to find out about the subsequent situation, must presumably count as not having had sufficient opportunity to know better.

(b) This is to be distinguished, of course, from the separate claim that rationality itself as an ideal or standard to be adopted, is unjustifiable. Such a claim has been proposed by Bentley in The Retreat to Commitment, N.Y., 1962, where he argues that:

> ".....a man cannot, without arguing in a circle, justify the rationality of his standard of rationality by appealing to that standard. Yet, if he holds certain beliefs - for example the standard itself - to be immune from the demand for rational justification and from the question "How do you know?" he can be said to hold them irrationally or dogmatically."

(Page 91)

15 a) See, for example, Jarvie and Agassi ("The Problem of the Rationality of Magic", British J. Anthropology, 1967), who attribute rationality to an action "if there is a goal to which it is directed".

16 See Evans Prichard, Nuer Religion, Oxford, 1956; Levi-Strauss, The Savage Mind, Libraire Plon, Paris, 1962.

17 For a careful account of the meaning of the term 'reasonable', and of some of the nuances referred to, see Max Black's paper on "Reasonableness", op. cit.

18 Peter Alexander (op. cit.) appears to claim that behaving rationally cannot be equated with simply behaving for a reason in this manner, since:

> ".....we should not say this (that the action was rational) if A's reason was a bad one. I may do something for a reason without its being the

reasonable thing to do, or a fortiori, rational."

(Page 329)

19 For example, by Peter Alexander, op. cit., Page 333.

15 b) See, also, Les Burwood and Carol Brady, "Married Women Students in F.E.: the meaning of coming to college", Journal of Further and Higher Education, Summer 1980, for an example of the type of sociology which attempts to show how explanations of rational action at the phenomenological level can be accounted for by wider social structural considerations.

CHAPTER FIVE

Summary

The possibility of assessing rationality of unconciously motivated behaviour is now considered in relation to the different conceptions of rationality and irrationality just described. First, the notion of behaviour which is "compelled" or involuntary is examined, and it is argued that on no plausible analysis of the meaning of compulsion could it be said that all unconsciously motivated behaviour is behaviour in which there is an element of compulsion and is, thus, to be judged irrational or "a-rational" on that account.

The possible rationality of unconsciously motivated behaviour is examined in the light of each in turn of the three conceptions of rationality distinguished earlier, and it is argued that according to each different conception, certain instances of unconsciously motivated behaviour can be assessed as fully rational. But, in addition, it is shown at the same time that not all unconsciously motivated behaviour fits the category of action which is rational or reasonable. Rather, among the various kinds of unconsciously produced behaviour, we find some which is simply caused, some which, while motivated, is irrational, and some which is both motivated and rational, even reasonable.

Finally, it is argued that although some unconsciously motivated behaviour can be shown to be rational, no procedural policy advocating repression or self-deception ever could be countenanced, and it is stressed that the only policy suggested by these findings is that of striving for self-awareness and self-knowledge, a policy traditionally proclaimed by psychoanalysis.

* * * * * * * * *

I wish to return now to the view mentioned earlier, which was attributed to Toulmin and Flew, or at least was shown to be a development out of their initial insight. This is the view that, since the unconscious mental states which are introduced to account for behaviour in psychoanalytic explanations are better classified as reasons and motives than as efficient causes, some at least of the behaviour which is the subject of those explanations, behaviour which has hitherto been regarded as irrational, is perhaps better regarded as rational. It is the view that because rationality consists in acting or proceeding appropriately in the light of one's particular state of mind, the behaviour of the neurotic or the unconsciously motivated person can be seen as rational if it is considered not merely in relation to the person's conscious beliefs and desires, but in relation to his or her unconscious mental states as well.

Unconsciously Motivated Behaviour and Compulsion

In my discussion of irrationality in the previous chapter it was shown that a person might be said to act irrationally in one sense (irrationally$_3$ or a-rationally), in failing to be rational through proceeding involuntarily. Being judged to have acted irrationally$_3$ would preclude forthwith the very possibility of being judged to have acted rationally. So, before we turn to the question of the possible rationality of unconsciously motivated, we must deal with the possibility that since all unconsciously motivated behaviour sometimes is said to be "compulsive" and so less than fully voluntary, it must be classified as irrational$_3$, and thus it cannot be expected ever to meet rationality requirements of the kind expressed in Chapter Four.

Let us turn then to the hypothesis that unconsciously motivated behaviour is irrational on account of being behaviour which is driven or compelled - behaviour which, in some sense, the agent has no choice in undertaking. I wish to approach this question by:
(i) establishing what is meant by the notion of an action's being "compelled" or "compulsive", and
(ii) showing that on no plausible account of the meaning of the concept of compulsion can it be said that all unconsciously

motivated behaviour is behaviour in which there is any element of compulsion.
In establishing (i) and (ii), I shall then have succeeded in rebutting the claim that because it is behaviour which is compulsive, or compelled, no unconsciously motivated behaviour can be treated as anything except irrational.

There has been some discussion of the role played by compulsion in the notion of unconsciously motivated behaviour.[1] In particular, the context in which this discussion has taken place has been those areas of jurisprudence and moral philosophy in which a distinction is drawn between behaviour which is undertaken freely and behaviour which is determined - or, as it is sometimes put, behaviour for which the agent is to be regarded as having been fully responsible and behaviour for which the agent is to be excused. Now, the position which is of interest to us here is one, very commonly seen in philosophical and jurisprudential discussions, which asserts that a person who can be shown to have been unconsciously motivated is to that extent to be excused from taking full responsibility for the action in question.[2] This position has been expressed, in philosophical writing, by Hospers[3] and I shall centre my discussion by examining certain of Hospers's arguments.

Hospers reaches the position that unconscious motivation is an excusing condition by way of the argument that since, in being unconsciously motivated we are caused rather than free, we must be regarded as being less than responsible for those of our actions in which we are unconsciously motivated. Thus, of the compulsion to wash found in the typical "obsessive-compulsive", he remarks that:

> ".....it has unconscious causes inaccessible to introspection, and moreover nothing can change it - it is as inevitable for him to do it as it would be if someone were forcing his hands under the tap. In this, it is exactly like the action of a powerful external force: it is just as little within one's conscious control."
>
> ("What Means This Freedom?", Page 31)

and

> "If we cannot be held responsible for the infantile situations (in which we are often passive victims), then neither, it would seem, can we be held responsible for compulsive actions occurring in adulthood that are inevitable consequences of those infantile situations.....Their occurrence, once the infantile events have taken place, is inevitable, just as the explosion is inevitable once the fuse has been lighted."
>
> ("What Means This Freedom?", Page 31)

I have selected these particular passages in order to emphasize two distinguishable strands which emerge from Hospers's general argument. First, there is the simple claim, mentioned already, that because unconscious mental states are themselves 'causes', they act as standard efficient causes do in producing neurotic behaviour. In response to this line of argument, it seems sufficient to draw attention to reasons given previously to show that introducing the notion of unconscious mental states as causes and that of unconsciously motivated behaviour as caused is unhelpful, since we have no satisfactory way of showing the difference between caused and non-caused behaviour (or causes and reasons) which in any way coincides with the distinction between the class of behaviour which is said to be unconsciously motivated and that which is not.

The second distinguishable line of Hospers's argument may be expressed in the following way: neurotic behaviour is caused by childhood experiences, and since a person has no control over his or her childhood experiences, and thus no control over what follows inevitably from them, then he or she cannot be held responsible for his or her neurotic behaviour. But, expressed this way, the argument contains a non sequitur. For, from the fact that some mental state, motive or inclination is caused in us, it does not follow that we are to be excused from taking responsibility for the action which that state, motive or inclination occasions. For example, our bodily appetites are caused in us (and their causes are also usually unknown to us, just as in the case of our unconsciously affected behaviour and states), but normally, unless they are especially strong and compelling, we are treated as acting freely and responsibly (although perhaps not as fully rationally)

when we act upon them. If I am assailed by a yearning for some kind of food, then despite this yearning's itself perhaps having been caused in me, I am usually held responsible for actions taken to satisfy it. I am not held responsible for having the craving, of course. Having the craving and acting upon it are quite distinct. The latter is 'caused' in the sense implied, but the former, often at least, is not. The argument which I have been examining fails either: in supposing, falsely, that every state or appetite which is caused in us is so strong as to compel us to act upon it (in which case, for a separate reason, we would be said to be less than responsible in acting upon it); or: in omitting one step in the account of the motivation in question: the step between unconscious cause, on the one hand, and resultant action, on the other.

Now there are, as we have seen cases where no conscious inclination or state mediates between the unconscious state and resultant 'action', for instance, verbal slips. But in other cases which I have considered of unconsciously motivated behaviour, there was shown to be an intermediate step, a conscious inclination to wish or do X, between the unconscious state and the action (X). And, in fact, as I shall argue, the concept of "compulsion", as distinct from that of "propulsion", would only apply when such a conscious inclination was present.

So while Hospers's general conclusion may be the correct one, neither of the arguments which he proposes appears to be sufficient to establish that behaviour stemming from unconscious states (including motives) is always behaviour which ought to come under the category of "compelled" or "compulsive".

The Concept of Compulsion

Let us turn to an examination of the concept of compulsion in our attempt to investigate whether, and if so why, (all) unconsciously motivated behaviour might still necessarily be said to be behaviour which is compelled, despite the failure of Hospers's arguments to prove it so.

My argument will be that the notion of the production of behaviour - whether caused or motivated and whether conscious or unconscious - is a different notion from the notion of behaviour being compulsive. Behaviour is "compulsive" when it is irresistably strong and I use examples to establish rough behavioural criteria to identify behaviour which might correctly be described as compulsive. Compulsion is usually associated with such things as threats carried out at gun point, drug addictions and certain obsessive-neurotic rituals. In knowing that a piece of behaviour has the predicate "compulsive" correctly applied, we do not, I suggest, know how that behaviour was produced: whether caused or motivated. It is always a separate question to ask.

What does it mean to say that some behaviour was compelled, or was compulsive in nature? Notice, first, the breadth of the notion: we speak of the compelling force of reasons and beliefs (I was compelled to sacrifice my queen; my sense of duty compelled me to attend the function), or threats of physical and other violence (I was compelled to march by my captor), of bodily needs and habituations and addictions (the addict is compelled to search for another 'fix'), and of applications in clinical psychopathology (the kelptomaniac is compelled to steal). And notice the dictionary's wide-ranging: "to urge irresistably, to constrain, oblige, force, a person to do a thing".[4]

In addition to the one expressed earlier in my discussion of Hospers's position, the argument that in being unconsciously motivated we are compelled to act and that in being compelled we are caused is open to an obvious criticism, which might be mentioned at this point. The claim that being caused to act is to be equated with being compelled to do so appears to abuse ordinary usage. For, in the everyday kind of cases in which we are said to be caused to act, as when we are pushed from behind, we are said to be "propelled" rather than "compelled". The concept of compulsion, whatever its other looseness, at least implies always that the person in question is aware of doing what he or she is doing and, in the broad sense discussed in Chapter Two (page 59), wants to do so. Whether acting at gun point or getting a 'fix' or obsessively washing his hands, the agent is aware of what he

is doing. To be compelled to act is always to have some desire(s) or impulse(s) (and also, perhaps, some belief(s)) concerning the action in question. And being pushed from behind and propelled implies the presence and mediation of no such mental states.

The dictionary definition of 'compulsion' given above included the expression 'force'. At first sight the inclusion of this concept in the definition of 'compulsion' may seem to contradict my claim in the previous paragraph that "compulsion" implies the presence of awareness of his or her action on the part of the agent. The use of 'forced' in: "He forced me off the cliff" is a case where no awareness is involved on the agent's part. In this case 'force' is being used in the sense of propulsion: "He propelled me and thus caused me to fall". In the case of propulsion 'force' is not, and would not be, followed by an infinitive: someone could not, in this sense, be said to force me to fall ("fall" involves no awareness on my part). I could only be said to be caused to fall. If someone forced me to jump ("jump" involves awareness on my part), then I would indeed be said to be compelled to jump, rather than to have been propelled off the cliff. The key point here is that if an agent is compelled to do X, he or she is aware of doing X - although not necessarily aware of why he is doing it.

Now both having been propelled and having acted under compulsion are treated, in certain contexts, as excuses for behaving in particular ways. And there is certainly, as that would suggest, an instructive analogy between the case of being propelled, by being pushed from behind for example, and that of being compelled. The similarity between the two rests in the quality of irresistability which characterizes the force which moves the person in each case.

Consider the case of being compelled by threats of physical violence. Having a gun pointed at one's head and being told to move, while different from the case where physical force is literally applied and one moves forward on account of having been pushed, is regarded as also involving an irresistable force. One could refrain from moving, despite the gun, but it is usually

understood that for all but extraordinary people, the wish to avoid certain death would itself be so powerful as to act upon them with something of the force and insistence associated with ordinary physical pressure like that suffered through the push from behind. The same applies in the case of the drug addict. The heroin addict is genuinely compelled to keep taking a 'fix'. He knows what he is doing but he is subject to an irresistable force. At least, if he does resist, the short-term consequences are very unpleasant.

As well as cases of compulsion of the kind just described, there are, as I have mentioned, a class of unconsciously motivated or unconsciously produced actions associated with particular pathological syndromes, or neuroses. Again, in these cases, the person acting is moved to act by an impulse which is so strong as to be apparently irresistable. In such cases, the force of the compulsion would appear to be at least as strong as that exerted when a person is compelled to take some action by threats of physical violence. The wish to repeat a neurotic ritual, for example, seems to have much of the same quality of irresistable force attached to it, as is revealed clearly in the following excerpt from an actual case study of a typical obsessive-compulsive neurosis:

> "Sally returned home.....with a fully developed hand-washing compulsion which did not go away. She soon had an ugly, painful dermatitis.....ugly and painful or not, the hands and forearms had to be scrubbed every time Sally went to the toilet and every time she had a "sensual sensation" or thought about contamination. If she resisted the impulse to wash and scrub, she had an anxiety attack. Her heart would pound, she would break out into a sweat and feel panickly and breathless - 'as if something were pressing the life' out of her, she said."[5]
>
> (My underlining)

From this illustration, we can establish some rough behavioural criteria by which clinical compulsions might be characterized. We might say that an impulse to act is irresistably strong just when:

(i) enormous emotional discomfort, stress and anxiety follow attempts on the the agent's part to resist the impulse, and/or

(ii) attempts to distract the agent from the impulse would be

unlikely to be successful - it insistently demands the agent's full attention until it is satisfied,

(iii) the impulse is recurring or continuous, and is thus insatiable.

While apparently isolating key features of the phenomenon of clinical compulsion as it is described in psychiatric case studies of the kind quoted, notice that we have not established a way of distinguishing this kind of compulsion from the compelling force which holds sway in the sorts of non-pathological context mentioned. Criteria (i)-(iii) in the previous paragraph would, indeed, count equally as behavioural criteria for a case such as the one in which a person is compelled to march by external threats. Inasmuch as the pressure persists, the impulse to march might be quite as insatiable, as attention-consuming and as impossible to disobey as the obsessive-compulsive's impulse to wash. This, I would suggest, is because clinical compulsion differs only in being associated with the behaviour of persons suffering from particular psychiatric conditions. It does not otherwise differ in any fundamental way from the "compulsions" found in non-clinical situations. (Although there is one additional clue to distinguishing it, which Austin notes:[6] only in psychiatric cases do we employ the adverbial expression and speak of acting "compulsively". In non-clinical cases we may speak of the agent as having been compelled - but not of his or her having acted "compulsively".

Let us note that the concept of clinical or pathological compulsion is not one which is sharp-edged. Behaviour is either caused or motivated, but there are degrees of compulsion. In mentioning a class of cases of clearly compulsive actions, it is necessary to speak very loosely. For, between the cases of compulsion described and cases of ordinary action, there is a gradation. Both in situations where an external threat compels (the gun held at the head), and among psychological compulsions like the one in the case study quoted, there are, in fact, a range of cases over which it becomes increasingly less plausible to speak of compulsions and of the reasons to act as compelling ones. And, there is a vast, fuzzy central area over which it would be difficult to decide whether or not the notion of compulsion has application. So, my account is not sufficient to enable us unfailingly to pick compulsive from non-compulsive behaviour. It

will merely let us distinguish clear cases of compulsive behaviour from clear cases of non-compulsive behaviour.

It is important to notice that the kind of compulsion described in the case study quoted characterizes by no means all clinically significant, or pathological behaviour. It is associated particularly with conditions centring around distinct cravings like alcoholism, drug addiction, certain sexual perversions and fetishes like exhibitionism, kelptomania, etc., on the one hand, and, on the other, with a distinct psychological and behavioural syndrome called "obsessive-compulsive" neurosis, or reaction, where a pathological concern with questions of right and wrong, and the concept of dirt and cleanliness combines with, and explains, irresistable impulses to act in certain characteristic ways, ways with which we are familiar from the typical "compulsive hand-washers". But those mentioned comprise only a small number of the behaviour patterns and syndromes usually classified as neuroses. There are many neuroses, for example, those associated with affective states, depressive reactions, mania and manic-depressive cycles, in which irresistable impulses, or compulsions, play no part at all.

Nor, more importantly, does such compulsion characterize all unconsciously motivated behaviour. Although it may not be possible always to distinguish behaviour which is compelled from that which is not, according to my rough definition, nevertheless, as I have said already, we do have paradigm cases of compulsive behaviour of the particular kind in question. And these cases contrast starkly with much unconsciously motivated behaviour, for instance with the kind of unconsciously motivated behaviour where someone refuses an invitation to a party, knowing that she did so partly because of unconscious reasons. There is no temptation, on the face of it, to describe this case as one of compulsion at all. There could be argument about whether behaviour was caused rather than motivated, but the standard criteria of applicability of "compulsive" are absent.

The production of behaviour - caused or motivated - is one thing, but being compulsive is another. Of course, it is always

open for someone to use "compelled" synonymously with "caused". But, if this is so, then the question of whether unconsciously motivated behaviour is compelled is a trivial one, because it is true by definition that it is not. There is always the problem, as we have seen in previous chapters, of deciding whether a particular piece of behaviour is caused or whether it is motivated, but that is a different problem.

Thus, there appears to be no way of interpreting "compulsion" which would encourage us to vary our initial assertion that not all unconsciously motivated (and otherwise produced) behaviour is behaviour which can be described as "compulsive" or "compelled". And we must conclude that any attempt to establish that unconsciously motivated behaviour is irrational in being compelled or compulsive, must fail. It is no more true that all unconsciously motivated behaviour is compelled or compulsive, than it is that all behaviour which is compelled or compulsive ought to be described as caused, in the sense Hospers employs.

We can now leave the question of the irrationality$_3$ of unconsciously motivated behaviour and concern ourselves with its rationality.

The Nature of the Thesis

From my analysis of the concept of rationality in the previous chapter, it can be seen that to ask whether unconsciously motivated behaviour is ever rational is actually to ask a number of different questions. In particular, it is to ask whether unconsciously motivated behaviour can be described as rational in any of the three different ways of characterizing the rationality of action described in Chapter Four (page 172). Clearly, unconsciously motivated behaviour may prove to meet one or more of those different characterizations without thereby meeting the others.

So the general claim, made by Toulmin and Flew and their followers, that all unconsciously motivated behaviour is rational, is too general to be informative. So, too, is the contrary claim, made subsequently by Alexander and Mullane,[7] for example, that unconsciously motivated behaviour never is, and never could be, rational. We must reject both these claims. The more correct view, as I shall show, is that while some unconsciously motivated

behaviour may be regarded as rational in each of the different ways described, some may not. It is inaccurate to characterize all unconsciously motivated behaviour either as rational or as irrational.

Further, according to the different ways of using the term 'rational', it is possible to evaluate both consciously motivated and unconsciously motivated behaviour alike in terms of their rationality, and to conclude that just as some consciously motivated actions are rational in each of the ways described, while some are irrational (in both of the philosophically interesting senses: irrational$_2$ or non-rational, as well as irrational$_3$ or a-rational), so some unconsciously motivated behaviour or action, in particular some at least of that which is attributed to psychologically normal, as distinct from neurotic persons, is rational, in some of the different ways of conceiving of rationality described in Chapter Four, while some is irrational. Thus, I would draw a sharp distinction within the class of behaviour which is plausibly explained in terms of unconscious beliefs, feelings, wishes and motives. Some behaviour of that class, for instance the compulsive handwashing of the obsessive, is completely irrational, while other instances of behaviour in the class - for instance, much of the behaviour which we encounter in what Freud called the psychopathology of everyday life, as well as some of the unconsciously motivated behaviour exhibited by neurotics - is rational.

Rationality (1): Unconsciousness and Acting in One's Interests, Maximizing Utility, etc.

In order to discover whether and when unconsciously motivated behaviour would be described as rational, we might begin by considering the characterization of acting rationally given under (1) in Chapter Four as acting in such a way that one furthers one's interests (or maximizes one's expected utility, minimizes one's expected disutility, minimizes regret, etc.). If it can be shown that some instances of unconsciously motivated behaviour fulfil one or some of these requirements, then at least the behaviour of the unconsciously motivated person sometimes can be said to be rational according to this criterion.

It may be supposed that unconsciously motivated behaviour is unlikely even to be rational in the way mentioned. Certainly, we are inclined to associate the idea of unconsciously motivated behaviour with the repetitive, pointless and characteristically self-defeating behaviour of the full-blown neurotics who find their way into psychiatric case studies: the obsessive with her foolish ritual, the patient whose life is spent in avoiding an inexplicable phobic object, or pursuing some meaningless fetish, Freud's hysterics and suchlike. And, most certainly, the behaviour of such persons is unlikely to meet the requirements for rationality expressed in characterization (1), if only because the mere fact that these people have had to seek clinical help seems sufficient to indicate that in these cases, their behaviour was at least likely to bring about consequences outweighing in undesirability the prospective desirability of the goal or aim it was intended to achieve (in this case, such undesirable consequences being consciously experienced: discomfort, failure, dissatisfaction, etc.). But these are by no means the only sort of case in which it seems plausible to offer unconscious states in explanation or partial explanation of behaviour. For instance, the whole range of explanations in terms of unconscious mental states arising in non-clinical settings (explanations of what may be described, in Freud's expression, as the psychopathology of everyday life) involve behaviour which would seem a much more likely candidate for being shown to be rational in the way of characterizing rationality under discussion.

In The Psychopathology of Everyday Life, as I have already remarked, it is proposed that much everyday behaviour, both that which is ordinarily regarded as intentional action and much which has hitherto been thought to be completely involuntary and the result of apparently insignificant causal factors, is really a reflection of unconscious beliefs, feelings and wishes, so that the behaviour of normal, psychologically 'healthy' (despite the reference to pathology in the description of this class of behaviour) and, we may add, rational-seeming people as well as the behaviour of neurotics, may be given a plausible explanation in terms of the agent's unconscious motives or mental states. Of those actions already recognized as such, for which the agent does

have some consciously held reason or reasons, the suggestion is that additional unconscious beliefs and wishes may be offered to supplement or, in some cases, to replace, the reason which is believed by the agent to explain the behaviour. And of that behaviour which was not hitherto regarded as requiring an explanation in terms of reasons or motives, the suggestion is that this behaviour is to be classed as action, and thus that an explanation in terms of reasons or motives, albeit unconscious ones, is to be sought in accounting for it. Since on the face of it, we might expect more likelihood of rationality among the psychologically normal than in the clinically significant behaviour of confirmed neurotics, it is towards this class of behaviour that we ought first to turn in an attempt to find cases of unconsciously motivated behaviour meeting one or several of the criteria expressed in rationality formulation (1).

In discussing the nature of unconscious motivation earlier, I distinguished, on the one hand, between cases of behaviour which may be said to be solely unconsciously motivated and behaviour which is only partially so, and, on the other, between cases of unconscious motivation in which one or several of the belief(s), desire(s) and emotion(s) comprising the motive itself were unconscious, a belief about the relation between the belief(s), desire(s) and emotion(s) comprising the motive and the behaviour in question, was unconscious, viz. the belief that that particular behaviour was motivated by those particular beliefs, desires or feelings. Let us begin by examining a case in which the behaviour in question is solely unconsciously motivated, and involves the type of unconscious motivation where (some part of) the motive itself is unconscious.

Example A. In The Psychopathology of Everyday Life, Freud describes the following case of unconsciously motivated forgetting:

> "Due to unknown causes, Jones left a letter for several days on his desk, forgetting each time to post it. He ultimately posted it, but it was returned to him from the Dead Letter office because he forgot to address it. After addressing and posting it for a second time, it was again returned to him, this time without a stamp. He was then forced to recognize the unconscious opposition to the sending of the letter."

If there were an unconscious wish to avoid sending the letter, as Freud suggests here, then the series of 'deliberate' forgettings described constitute motivated actions, in the sense described earlier. They are likely to achieve, at least temporarily the fulfilment of the wish. Since we are not told the unconscious reason for the "counter will", it is not so easy to judge whether these forgettings also meet criteria of the sort expressed in rationality formulation (1) - whether, for instance, they may be seen to have been in the agent's interests (and believed by him to have been so). Let us suppose that the letter contained a cheque to cover a bill which Jones believed (consciously or unconsciously) to have been unjustifiably high. In this case, the 'action' taken might well prove self-defeating and irrational by merely hastening the unpleasant and troublesome advent of debt-collectors or lawyers' threats. But let us suppose, instead, that the letter contained a proposal of marriage to a woman with whom Jones was besotted. We can well imagine that his unconsciously expressed reservations might be justified in this case, although he could not consciously recognize or acknowledge them, and that were the marriage to proceed, it would be a failure. So, we can alter the case until it conforms to part of the criteria expressed in rationality formulation (1). In this case, Jones might be said to have an unconscious belief that severing the relationship would be in his interests, at least in one way. But could it be said to be in his interests? It might be objected that the only rational thing to do in this situation, given that one had some (unconscious) reservations about the relationship, would be to become acquainted with them, in the first place, and then to act upon them, by explicitly severing the relationship or explicitly and knowingly ceasing communication. Yet, if the considerations involved in being in one's interests or, indeed, if those relating to the notion of maximizing utility, or minimizing disutility or regret, are taken into account, then it can be argued that the anguish which such a direct approach would create in the besotted man would be a consequence outweighing in undesirability the prospective desirability of the expected consequence, and that the more devious approach, which at once protects the man's feelings while ensuring that such action is taken as to avoid producing unfortunate long-term consequences, is at least as rational as the direct one. Similarly,

it may be proposed that while the forgetting behaviour (unconscious action) could be said to have been a rational response to the situation, nevertheless engaging in the same behaviour, that is, failing to post the letter while consciously being aware of what he was doing, must count as having been a more rational one. However, the feeling of felt discomfort brought about by the conflect of wishes (the wish to post the letter and the wish to avoid doing so), were the unconscious wish to enter the agent's conscious mind, seems likely to be sufficiently undesirable a consequence of this case for us to say that the former one at least equals and perhaps exceeds the latter in its rationality. This then is one case of solely unconsciously motivated behaviour which may be considered to be rational according to rationality formulation (1). It is a rational thing to do, or way to proceed, given the situation; it may even be said to be the most rational thing to do.

Let us consider another case. Many of the examples which Freud uses to illustrate the notion of the unconscious motivation present in everyday life are those of verbal slips, written and spoken. It will be remembered from my earlier discussion of the nature of unconscious motivation that these cases exhibit certain peculiarities which make them less clear-cut cases of unconscious motivation than the sort of error of memory which I have just considered, but since Freud himself treated them as in every way analagous to other unconsciously motivated 'errors', we might look briefly at one notwithstanding that.

Example B. This time Freud quotes from Stekel:

> "An unpleasant trick of my unpleasant thoughts was revealed by the following example: To begin with, I may state that in my capacity as a physician, I never consider remuneration, but always keep in view the patient's interests only: this goes without saying. I was visiting a patient who was convalescing from a serious illness. We had passed through hard days and nights. I was happy to find her improved, and I portrayed to her the pleasures of a sojourn in Abbizia, concluding with: "If, as I hope, you will not soon leave your bed." This obviously came from an unconscious selfish motive, to be able to continue treating this wealthy patient, a wish which is entirely foreign to my waking consciousness, and which I would reject with indignation."
>
> (The Psychopathology of Everyday Life, page 16)

Assuming Stekel's own hypothesis about the unconscious motive to be the correct one, we can see that his slip clearly expressed or was

a sign of his unconscious wish for continuing treatment of and remuneration from the wealthy patient. Yet, it seems to be ill-designed as a method of achieving such an end. Saying "I hope you will not soon leave your bed" is most unlikely to bring about the fulfilment of the wish for continuing treatment. And, while the goal may be said to have been in Stekel's financial interests, without further information it is difficult to judge whether or not it might be said to have been in his interests simpliciter. For the very same reason that it barely counts as a clear-cut case of motivated action, as distinct from action in which an unconscious wish is reflected, or revealed, at all, it will meet that part of the rationality requirement which stipulates that it bring about no consequences outweighing in undesirability the prospective desirability of the aim or goal. And, this will be true of the general class of unconsciously produced speech acts, written as well as spoken. This is because they generally go disregarded and even unnoticed, since away from the psychoanalyst's couch, they are usually not taken to have any significance at all. They are, thus, rendered at once ineffectual and harmless - they simply have no significant consequences, desirable or otherwise.

We can, of course, view cases of unconsciously motivated speech acts in a slightly different light by transforming them into what look to be potentially rational (1) actions of a somewhat different kind. To do so, we need merely to postulate that these verbal acts are motivated by an unconscious wish to express some belief, wish or attitude held, in this case, unconsciously. If we concede that directly expressing the belief, wish or attitude is a way of fulfilling the wish to express it, then we appear to have at least a promising candidate for an additional sense in which unconsciously motivated speech acts might be said to be rational according to the criteria expressed in formulation (1).

Ceteris paribus it is not against our interests to give verbal expression to our mental states, whether they be conscious or unconscious. But is it actually in our interests? Clearly, if we accept the fact mentioned previously, that verbal slips qua

actions, generally have no significant consequnces, desirable or otherwise, for the person whose actions they are, then in order to say so, it must be the case that expressing one's beliefs, wishes, attitudes, etc., counts as intrinsically desirable. But is it? A conclusive answer to this question would require too extensive a departure from the main topic at issue. It is perhaps sufficient to say that self-expression has long been touted by philosophers and educators, both as an instrumental and as an intrinsic good. We can conclude that it might well be that self-expression has intrinsic value.

With respect to Example B, it might still be objected that the possibility of this kind of unconsciously motivated speech act's being regarded as rational according to the criteria expressed in rationality formulation (1), presents difficulties in permitting us to do violence to our intuitions by allowing us to drive a wedge between the notion of the expression of the belief and of the belief itself and to judge that the expression of the belief might be rational although the belief itself was held irrationally. Thus the possibility seems to be allowed of its being rational to express an unreasonable belief. If the following conditions were met: the person

(a) wished or intended to express the belief;

(b) did so in a direct way without circumlocution or metaphor; and

(c) if his or her doing so was unlikely to lead to harmful consequences,

then regardless of the unreasonableness of the belief, the person's expression of the belief must be treated as rational.*

---

* The above class of restrictions may actually suffer a further restriction. For, except in the case of unconsciously motivated speech acts of the sort which Freud describes, we rarely have a single purpose when we speak. As well as wishing or intending to assert a belief which one often does when one's speech is consciously motivated, one also often wishes or intends to describe, inform, comfort, agree, warn, frighten, command, embarras, impress, bamboozle, etc., or some combination of these, and it may be said to be the rationality of the combination of different wishes involved, or at least of the most significant wish involved, which must decide the rationality or irrationality of the total speech act. So that even if fulfilling one's wish or intention to express some belief were to satisfy the rationality criteria, if one also spoke with the intention of comforting or reassuring, but chose for the purpose to express a belief with little likelihood of producing that effect in the hearer, we might want to say that overall such total speech acts were less than rational.8

While seemingly odd, such a description is not, however, in fact inconsistent with any of our intuitions about what counts as rationality. We are not accustomed to speak of speech acts as rational or irrational, it is true; usually we restrict that evaluation to gross bodily actions. But this seems to be nothing more than custom. While bodily actions are the primary object of application for evaluations in terms of rationality, an extension of that application to speech acts would seem to be innocuous enough since speech acts do fit into the broad category of intentional behaviour.[9]

However, before leaving the discussion of verbal slips, I would repeat that, unlike that of errors of the sort described in Example A, the case of verbal slips is one which, unless we resort to an interpretation in terms of unconscious wishes to express, can only be dubiously included in the category of action which is motivated rather than merely causally produced by unconscious mental states.

To treat verbal slips the way Stekel's remarks indicate that he did and Freud himself often did, as efforts to fulfil the wish they reveal, is to have them fail consistently to meet what we have been regarding as the mark of something's being a motivated action - its appropriateness for or fit with the aim for which it was supposedly undertaken: it must have some possibility at least of fulfilling that aim. For that reason, it seems more apposite to see the slip as standing in the relation of being a sign of, or revealing or reflecting the particular unconscious wish it expresses, than as being an action designed and taken in order to carry out that wish. So unless we introduce the alternative account and treat the "wish to express the unconscious wish mentioned" as the unconscious wish, verbal slips are not, strictly speaking, unconsciously motivated.

As was remarked earlier, if we should resort to the latter interpretation of treating the wish to express as the unconscious motive, then we would be faced with the additional difficulty that to introduce a wish to express some verbal act is never to add anything by way of explanation of that act which could not have

been added merely by stating that it was an act, that is, that it was, broadly speaking, intentional. And since to give a motive is always to explain, merely to refer to some slip as indication of an unconscious wish to express some future wish, is not, strictly speaking, to give an account of motivation at all.

In discussing the different kinds of solely unconscious motivation, I distinguished cases of the sort mentioned in Examples A and B from those where the agent did something knowingly and in doing so also did something else unconsciously. An example of the latter case used was of person X who, while (knowingly) conversing with P, was also unconsciously hurting P's feelings. While X's conversing with P was said to be an action with mixed motivation, some conscious and some unconscious, her hurting P's feelings was given as an instance of a solely unconsciously motivated action.

Example C. By slightly altering that case, we can construct an example of another action which is solely unconsciously motivated and which appears to be rational in the manner under discussion. Let us say that in choosing examples to illustrate what she is saying, P unconsciously selects those which flatter her listener Y. If we postulate that P has an unconscious wish to make Y respond warmly towards her, and a justified, conscious or unconscious belief that having Y respond warmly towards her is in her interests, then we seem to be able to say that as long as Y registers the flattery and accepts it in the right spirit (a supposition which seems likely enough), P's unconsciously motivated flattery might be described as rational in the sense of rationality (1). Nor would it be evidently more rational in this sense for P to have consciously acknowledged her unconscious wish and either to have expressed it more directly or to have engaged in the same flattering strategy with conscious intent. It may be plausibly argued that approaches of either kind would be self-defeating: openly expressed a wish that someone will be your friend is notoriously more likely to produce the opposite effect. Since conveying self-conscious and intentional flattery is a skill which few master, the person may sense and resent the art in the flattery, making it a dangerous strategy at best for all but the most skilled practitioner, and we can build it into our case that P is not one.

The fourth example of solely unconsciously motivated behaviour which I mean to discuss is one which includes the feature described earlier of "rationalization". Here, the agent falsely believes himself or herself to have a conscious motive or reason for the action when in fact the action is motivated solely by some unconscious mental state or states. In addition, I wish to introduce another new feature. In contrast to examples given so far, in each case of which some part of the belief(s) and wish(es) comprising the motive itself was unconscious, I wish to introduce a case where the only unconscious state is a belief concerning the relation between the belief(s) and wish(es) on the one hand, and the behaviour which they motivate, on the other. The motives are not unconscious, but the motivation is. In order to use an illustration from the clinical field, to contrast with the everyday life examples considered so far, and thus to show that even some neurotic symptoms may be rational in the way expressed in formulation (1), I shall adapt the example of Freud's obsessive patient to which I have referred before.

Example D. A part of Freud's obsessive patient's elaborate bedtime ritual was opening the door which led into her parents' bedroom. In insisting that this door be kept open, Freud explains, the girl was unconsciously motivated by a wish to prevent sexual intercourse from occurring between her parents, since she regarded this as likely to result in a rival to herself. As was remarked in earlier discussion of this case, it is at least possible to postulate from the rather meagre evidence which Freud allows us on the background of this patient, that at some time during the development of her neurosis, the patient was unconsciously motivated, in observing this particular part of her nightly ritual, without it being true that the nature of her wish was totally unknown to her. She knew of her wish, yet at the same time she was unaware of the relation between her behaviour and the wish. Thus she had an unconscious belief that the one affected the other.

In order to complete the example we need merely add a detail which the original case did not possess. Let us say that serving to distract the patient's attention from the above-mentioned unconscious belief about the relation between her (conscious) wish

and her behaviour, was an additional consciously held belief, this time a false one, about the reason for her wish to keep the door open: the belief that doing so assuaged her fear of the dark. Her rationalization for wanting the door open was that she felt less frightened in the dark with it open, knowing that her parents were close, while the real reason why she wanted it open was her wish to prevent her parents from having intercourse.*

Can we say that Freud's patient's action was rational? Certainly, as was observed before, this particular action, unlike many others undertaken in the course of the nightly ritual, has the character of a motivated action: it is appropriate to the wish in the light of which it is undertaken; without the aid of added assumptions it can be seen as a possible way of fulfilling that wish. Assuming normal inhibitions on the part of the patient's parents, we might go further and suppose it to be a likely way of effecting the wish. And we can similarly suppose it to be the case that in thus fulfilling her wish, the patient was acting in her interests. Nor is it obvious that any less devious method of carrying out the wish would have been as free of undesirable consequences. It may be supposed that had she been aware of the true reason for her behaviour, the patient would either have refrained from carrying it out at all and thus would have suffered the pains of frustration, or would have done so with increased embarrassment and/or guilt feelings. So it would seem that in carrying out this aspect of her obsessive ritual, at least, Freud's patient may be said to be acting rationally (1): she may even be said to have been doing the most rational thing, given the situation.

Example E. Before leaving the question of the rationality (1) of the various sorts of solely unconsciously motivated behaviour distinguished in Chapter Two, we must examine the kind of case introduced on page 71, in which the action can be seen as motivated by an unconscious belief or wish only if we assume some symbolic identifications. Let us turn again to the example there discussed, of Freud's patient's nightly ritual of separating bed and bolster. It was remarked of this kind of case that the action could be interpreted as being unconsciously motivated in two different ways. The action may be regarded as fulfilling what we

there called the secondary unconscious wish, viz. the wish on the patient's part to express her primary unconscious wish (to prevent sexual union between her parents). On the other hand, it may be seen as an appropriate way of fulfilling the primary unconscious wish if we assume the symbolic identification of bolster and bedstead with the patient's own mother and father, respectively.

If we regard the behaviour as unconsciously motivated in the former way - that is, as fulfilling the secondary unconscious wish - then the question of the action's rationality, in the sense under discussion, might be answered in the way it was when the same sort of analysis was given in terms of unconscious motivation behind verbal slips. It can be imagined that, like verbals slips, rituals of the sort involved in our case might be without significant consequences. And we have then merely to insist that because the action of expressing the unconscious wish itself would be good for the agent, the ritual can be judged to be in the agent's interests. That is, if it can be said that self-expression is an intrinsic good, which seems to be at least possible, then both verbal slips and such symbolic actions as the one under discussion might be said to count as rational according to the sort of criteria expressed in formulation (1).

On the second interpretation of the case in which the action is analysed for rationality, it is seen as motivated by what we have called the primary unconscious wish, viz. to prevent intercourse between her parents. As was stated in earlier discussion of this example, on the assumption implied in the identification (that bed and bolster are father and mother), but only on that assumption, would the method employed be 'appropriate to the desired aim'. However, in reality no such aim would be achieved by this method and therefore the action could not be deemed to be in the patient's interests, and is thus not rational.

It would seem that whenever cases involving symbolic identification are introduced and treated in terms of the rationality of fulfilling the primary unconscious wish rather than that of fulfilling the secondary one (the unconscious wish to express the unconscious wish), difficulties will arise. As long as it is held

that rationality requires both not only that the action be in the agent's interests, but also that the agent hold the justified belief that it is so, then such cases will fail to meet the requirement for an action's being rational.

We can now conclude in general that even solely unconsciously motivated behaviour, both that exhibited by neurotics and that found in the psychopathology of everyday life, may be said to be rational (in certain instances) in the way described in rationality formulation (1), that is, rational in the sense of furthering one's interests. When we turn to the case of behaviour whose motivation only partially springs from unconscious sources, we find, as might be expected, that the same is true. Much partially unconsciously motivated behaviour seems to meet the requirements expressed in rationality formulation (1).

I have mentioned already in Chapter Four the alternative strategies available when an assessment must be made of the rationality of an action stemming from more than one distinguishable motive, and the possibility of assessing the action under each separate motive description available of it. This strategy, as was there remarked, leads to a more stringent rationality requirement since we may stipulate that the action is rational if and only if it meets the rationality requirement as assessed under each separate motive description. This strategy is the one best employed where the mixed motivation in question is partly conscious and partly unconscious. I shall hope to establish that even on this stronger criterion, some unconsciously motivated behaviour can be shown to be rational. For as long as we were to apply only the weaker strategy - assessing the action's motives as a whole - our account risks being countered with the objection that qua unconsciously motivated action, the action in question is not rational, that it derives its rationality from the rationality exhibited by the action regarded in relation to the conscious motive underlying it.

Consider the following case:

Example F. P likes Q and is aware of no ill-feeling on her part towards Q. P refuses an invitation from Q, giving as her reason, which she sincerely believes to be her only one, that she

has pressing work to do on her thesis. It is true that the work must be done, but unlikely that that alone could be sufficient to motivate (party-going, invitation-accepting) P to refuse. Now, if there are independent reasons for believing that P feels some measure of unconscious hostility towards Q, then either one of the following explanations for P's refusal might be plausible:

(1) She refuses the invitation because:

(i) she believes that she ought to stay at home and work on her thesis, and

(ii) she unconsciously wishes to avoid seeing Q.

While neither (i) nor (ii) alone would be sufficient to produce the refusal, the two conditions together are jointly sufficient to do so.

(2) She refuses the invitation because:

(i) she believes that she ought to stay at home and work, and

(ii) she unconsciously wishes to avoid seeing Q.

Either her unconscious belief or her unconscious wish alone would be sufficient to make her refuse the invitation, but in this case both are operate in her decision, that is, her decision is "over-determined".

When we consider this example in the light of the rationality criteria expressed in formulation (1), we find that the behaviour clearly is motivated (refusing the invitation is an appropriate way of achieving the fulfilment of both goals, the conscious wish for time for thesis-writing and the unconscious wish to avoid seeing Q). And also, since there is no particular reason to suppose that the refusal would bring about adverse consequences outweighing in undesirability the prospective desirability of the goals, and since we can suppose that P is justified in believing it to be in her interests to work on her thesis, and that she does consciously believe it, and is similarly justified in believing that it is in her interests to avoid seeing Q and, albeit unconsciously, does believe it - it is also rational in the sense of being the or a rational procedure given the situation. And again, as we have done in previous cases, we may compare the behaviour undertaken in this case with other possible strategies for dealing with the particular situation in question, strategies which involve no

unconscious motivation, and conclude that according to this criterion for ascribing rationality, the strategy which involves unconscious motivation may actually be shown to be more rational than the alternative strategies. That is, it could be argued that since the amount of consciously experienced unhappiness for P which an acknowledgement of her unconscious hostility towards Q would be likely to produce counts as an undesirable consequence of the alternative strategy, the initial strategy must be regarded as being in P's interests to a greater extent than the alternative, more open and insightful strategy would be, and hence more rational.

In conclusion, then, we may say that it is possible for both solely and partially unconsciously motivated behaviour to be correctly described as rational.

Our next question must be whether the unconsciously motivated person can be said to act rationally according to either or both of the two alternative formulations of rationality described in Chapter Four: formulations (2) and (3). In being unconsciously motivated, whether solely or in part, are we ever rational in the ways described in (2) and (3)? It is important to stress that the same behaviour which fulfils the requirements expressed in formulation (1) may not be that which fulfils those expressed in (2) and/or in (3). A person may act in such a way as to further his or her interests, or to maximize his or her expected utility, without having reasons in acting or acting out of these reasons.

Rationality (2): Unconscious Motivation and Acting with a Rationale

In considering whether unconsciously motivated behaviour might be described as rational according to the formulation of rationality expressed in (2) - in which acting rationally is simply acting with a motive, rationale or purpose, we have only to remember the definition of a motivated action for our answer. Any behaviour which is correctly described as unconsciously motivated, according to the characterization of motivation given in Chapter Two, may be described as rational according to the formulation of rationality expressed in (2).

Let us turn to the formulation of rationality which is expressed in (3). This formulation, it will be remembered, is the one which equates acting rationally with acting reasonably, that is, having and being moved by good reasons in acting.

Rationality (3): Unconscious Motivation and Acting Reasonably

Three different questions, corresponding to three distinguishable requirements, suggest themselves:

(a) Are unconsciously held reasons "reasons" in the required sense? (see below)

(b) Are unconsciously held reasons good reasons? (see page 210)

(c) Do unconsciously held reasons move us to act? (see below)

Question (c) requires but a perfunctory answer: unconscious "reasons" are indeed what move us. How they do so may be questioned, but not, assuming that the very concept of unconscious reasons is accepted at all, that they do. Let us then turn to the first question: are they reasons at all?

In answer ~~the~~ to (a), it will be remembered from Chapters Three and Four that in the sense of 'reason' required in a discussion of rationality, only those cases of unconscious beliefs and desires of which we happen to be aware could be said to be (unconscious) reasons. This is because, although in Chapter Three it was argued that not all that we call reasons need meet the awareness condition, nevertheless, in discussing the sense of 'reason' required in the demands for reasonable behaviour, it was argued that a simple version of the awareness condition must be met (where a person has a reason, then he or she must be conscious of it).

However, even if we concede the general possibility that some unconscious beliefs and desires (viz. those which we are able to deduce by feat of self-analysis or with the aid of an analyst at the time at which the behaviour for which they are the reasons is taking place) can be said to be the reasons for what we do, still, when we come to examine this possibility in relation to particular cases, we find that a number of factors detract from the likelihood of this possibility's actually being realised.

The first of these factors becomes apparent if we turn to

the case of Jones's unconsciously motivated forgetting to post the letter, discussed in Example A. Solely unconscious motivated omissions of memory of this kind would seem by their nature to be precluded from a realisation of the possibility described above, since in these cases an absence of awareness on the agent's part of what he or she is doing is a necessary condition of its being done at all. It is simply a contradiction to suggest that a person might at once be aware of - and hence, remember - and be unaware of or forget to do something.*

Although I have shown that the case of verbal slips, illustrated in Example B, fails to meet the rationality criteria expressed in formulation (1) and perhaps even fails to meet the criteria by which we distinguish actions which are motivated, at all, it is nevertheless worth looking briefly at the case of verbal slips also in light of this discussion. Being aware either of:

(a) the unconscious wish, or 'counter-will', for example, the wish to say "not-X", or of

(b) the unconscious wish and its likely outcome in behaviour,

and then, at some subsequent time, making the slip, seems to be possible. The difficulty with describing a person as at once being aware of (a) and (b), and at the same time making the slip, is that the most plausible explanation of the slip would in that case always seem to be not that the unconscious wish motivated the behaviour, but that the conscious awareness (of (a) and/or (b)) caused it, acting in the manner of a standard mental cause. So for the case of verbal slips, as well as for the case of other kinds of unconsciously motivated omissions, like the forgetting discussed earlier, it seems true to say that knowing, in the sense of having deduced, one's unconscious motivation would make simultaneously acting from that motivation impossible.

---

* A phenomenon somewhat like this is said to occur in cases of what is called 'self-deception'. But in such cases the two 'attitudes', awareness of X, or awareness that one ought to do X, on the one hand, and unawareness of X, or failure of awareness that one ought to do X, on the other, are both, strictly speaking, in consciousness. One is merely, somehow, given less attention or emphasis than the other; the item to be ignored or 'forgotten' is merely, at most, preconscious, in Freud's terminology. It is not unconscious, in the full sense defined in Chapter One, at all.

Let us turn to Example C, of a person's doing something which was entirely unconsciously motivated (when the agent was unaware not merely of the motivation but of all or part of the belief(s) and desire(s) comprising the motive itself): the case of P's unconsciously selecting examples to flatter her listener Y. We can imagine that if P were an insightful person, and practised at analysing the unconscious elements in her behaviour, she might deduce the significance of the particular selection of examples which occur apparently randomly to her, in illustration of what she is trying to say, at the time at which she is speaking. And we can imagine that if P did not actively disapprove of flattery, and had no other qualms about using it, she might continue to use the examples which came to her, rather than purposely thinking up alternative ones or making alterations in the original ones to disguise the flattering element in them. In order to use this case we must emphasise that P was indifferent to the flattering effects of, and intention in, her examples - and had no additional conscious motives with regard to them. (We must presumably say that she had one changed attitude as a result of her insight - in order to have continued to use examples we must hypothesise that consciously she wanted to do so. But the ascription of a mere want of this sort is not sufficient for us to have to describe her as having become (partially) consciously motived.) Given the above qualification, then, we do seem able to allow this as a case of solely unconsciously motivated behaviour in which the agent continues to act on unconscious reasons of which she is aware - a case which is a candidate for being regarded as rational in the sense of reasonable; but it becomes rational only when the agent becomes aware of her previously unconscious reasons.

Let us compare this with the case given in Example D, in which while the behaviour was said to be unconsciously motivated, none of the belief(s) and desire(s) constituting the motive itself were unconscious beliefs and desires; the only unconscious item was a belief about, or awareness of, the relation between the motive and the action. This is the case of Freud's patient opening the door between her own and her parents' room, knowing that she wished to avoid having a younger child in the family, knowing also that her parents' having intercourse would be likely

to produce a child, perhaps even knowing that she did not wish her parents to have intercourse lest it result in a child, knowing that opening the door would likely prevent them from having intercourse - yet believing (falsely) that she opened the door to assuage her fear of the dark. Since her reason for acting is simply those belief(s) and desire(s) comprising the motive from which she acted, in this case of solely unconsciously motivated behaviour at least it would seem that the agent can be said to have been aware of her reason for acting at the time at which she did so, although clearly she was not aware of it as her reason, since she believed her sole reason to be her fear of the dark. Technically this case seems to meet the requirements expressed in the rationality formulation with which we are concerned. However, what we normally understand by 'having a reason' seems to include, not only being aware of that reason (or being able to become aware of it), but also being aware of it as, or believing that it is, one's reason. Because of this qualification, I do not think that this case can after all unqualifiedly be proposed as a case of solely unconsciously motivated behaviour for which the agent can be said to have reasons.

Let us notice that if we change the example in accordance with this qualification on the notion of having a reason, then although it would seem probable that if the agent became aware of her previously unconscious motive, her changed state of awareness would create in the patient an additional conscious motive to undertake the action in question, this need not be so. However, on the contrary, although it seems in this case unlikely, nevertheless the patient may continue to act with full awareness of her 'unconscious' reason for doing so and yet without any further alteration in the state of her conscious beliefs and wishes. Her sole consciously felt wish may still be to avoid being alone in the dark, and therefore to keep the door open. In this case we should say that the agent did not act rationally$_3$ in keeping the door open, since she did not act out of a conscious reason.

Thus, of the different kinds of action classified in Chapter Two as solely unconsciously motivated, only two can be shown to be such that they may qualify for the evaluation of 'reasonable'.

First, those things done while doing something else, may qualify. And, second, so may those things done which, if not actually partially consciously motivated, are at least accompanied by an ordinary, conscious wish or inclination to do them.

In contrast with the cases just discussed, behaviour which is only partially unconsciously motivated of the sort described in Example F, as we would now expect, is not affected by the introduction of the concept of reasons. If P had sufficient self-knowledge to guess that she refused the invitation partly out of an unconscious wish to avoid seeing Q, she might still refuse, still feel consciously nothing but warmth and liking forQ. (In acknowledging that her unconscious state was a necessary condition for her refusal, and that her conscious reason was not a sufficient one, P would not have consciously to want to avoid seeing Q in order to persist with the refusal - let us assume that her flash of insight occurred just as she was beginning to speak to make her apologies.) So we may conclude that both behaviour which is partially unconsciously motivated and behaviour which is solely so, sometimes can be said to be done for reasons, in the sense of that expression understood in rationality formulation (3), and can thus be a candidate for being described as rational in the sense of being reasonable.

On page 206 I raised three questions relating to the requirements of what it is to act reasonably. I have now dealt with questions (a) and (c); I now turn to question (b): Are unconsciously held reasons good reasons? In order to say of any unconsciously motivated behaviour that it is done rationally or constitutes rational action according to the formulation by which rational action is action which is reasonable, an added requirement must be met. Not only must the agent have reasons for what he or she does, those reasons must be good reasons.

The requirements for good reasons outlined earlier may be said to be two: in the first place good reasons were seen to comprise none but rationally held beliefs and rationally reached desires; in the second place it was seen that in order to be good reasons, the sort of beliefs and desires comprising one's reasons must be coherent, in the sense of being consistent with one's other

beliefs, values and attitudes, or, if they are not so consistent, it must be the case that:

(a) the inconsistency is recognised by the agent, and

(b) the agent judges the belief, attitude or value upon which he or she acts or proceeds to be the more rational or reasonable one of the two, or the most reasonable or rational of the number of the conflicting beliefs, values and attitudes however many there be.

Let us begin by considering the second stipulation concerning the reason's coherence. I have remarked previously upon the claim that our unconscious beliefs and wishes are always contrary to our consciously held beliefs and wishes. Were this claim true, then clearly no unconscious reasons would ever succeed in fulfilling the coherence stipulation. But as I have already stated, there is little reason to embrace the claim as long as we remain uncommitted to Freudian theories of repression.

It would appear that we may not need to investigate particular examples to establish that sometimes unconsciously motivated behaviour can be done for good reasons in the sense of reasons which cohere with the agent's other consciously held beliefs, values and attitudes. For we would seem to have established that already in showing that sometimes we can have reasons for our unconsciously motivated behaviour, since we are able, in certain cases, to become aware of our unconscious motives at the time at which we act. It has been argued that it follows that if, in these cases, we were to continue to do whatever we were doing, and in doing so to gain admission for our own action to the category of actions for which there can be said to be reasons at all, then necessarily we must also in those cases be acting upon unconscious motives which were compatible with our other, consciously held, beliefs, attitudes and values: our very allowing ourselves to (continue to) act upon them is, as it were, the guarantee of their meeting the coherence requirement, and thus being, to that extent, not merely reasons, but good reasons. Since we could not act upon our unconscious motives if they were not compatible with our consciously held beliefs, attitudes and values, it appears to follow that if we do act upon them, then they must be so.

In reply to the latter a priori argument which, if it is valid, will have established that all unconscious reasons properly so-called are good reasons in the sense specified, there is one obvious objection. It is not the case, it may be argued, that any time that an unconsciously motivated person continues to act upon his or her unconscious motive even after coming to acknowledge it as one of the reasons for his or her action, the belief(s), wishe(es) or attitudes comprising that unconscious motive are consistent with the person's conscious beliefs, values and attitudes. This may usually be the case when a psychologically normal, or 'healthy' person so acts, but in the case of the action of neurotics, the element of compulsion often enters in, so that even though his or her continuing to act upon the acknowledged unconscious motive is a source of shame, guilt or despair simply because of the incompatibility between that motive and the person's avowed and consciously held beliefs, attitudes and values, nevertheless the motivation is strong enough to make the person powerless to resist it. We can imagine, for example, that even were Freud's patient whom we discussed earlier to have gained sufficient self-knowledge to understand that in addition to her ordinary, conscious, wish to assuage her fear of the dark, the door-opening part of her nightly ritual was motivated by her unconscious wish to prevent her parents from having intercourse, and even if the idea of this additional motivation were highly repugnant to her, nevertheless she might find herself so uncomfortable and anxious until the door was opened, that she might have been unable to resist her urge to open it.

And needless to say, the incompatibilities arising between a person's unconscious motives and consciously held beliefs, values and attitudes, such as these, are unlikely to be accommodated by way of an application of the qualification, mentioned earlier, that if there is inconsistency then it must be the case that the inconsistency is recognised by the agent and the agent judges the belief, attitude or value upon which he or she acts or proceeds to be the more rational or reasonable of the two, or the most rational or reasonable of the number of the conflicting beliefs, values or attitudes, however many there may be. For in the case of the neurotic pursuing a compulsively followed ritual, it is

likely that while continuing to act upon them he or she would acknowledge the beliefs, attitudes and values underlying the compelling motive to be far from the most reasonable or rational he or she holds (apart from acknowledging them to be abhorrent in more distressing ways).

Because of the possibility of the kind of compulsion described, it might be argued that the class of cases of unconsciously motivated behaviour in which it is possible to attribute reasons to the agent is not coextensive with that in which the agent can be said to have good reasons in the sense of reasons which are coherent with his or her other, consciously held beliefs, attitudes and values.

But the case of compulsive behaviour would appear to be the only kind which it is possible to use in support of the above argument. And as long as this is so, the argument still would remain undefeated. For the case of behaviour stemming from unconscious states which thus compel or force us to act, cannot strictly be called a case of unconsciously motivated behaviour at all. Rather, it is behaviour which must be said to be produced or caused by unconscious mental states. In these sort of cases unconscious mental states do seem to function much more like brute causes of, than like like reasons for, our behaviour. So it would seem that the suggestion made initially still might be the correct one: as long as any unconsciously motivated behaviour can be shown to meet the requirements for being done with a reason at all, then it will usually also meet that part of the requirements for good reasons which is concerned with coherence and consistency.

A more serious objection remains, however. Our generalisation may be questioned on the grounds that the agent's own acceptance of his or her unconscious reasons cannot be used as a criterion for their compatibility with the set of that person's consciously held beliefs and values, since we are notoriously imperfect in the practice of detecting incoherence and incompatibility among the body of our beliefs and values. This argument is not, in relation to what are usually regarded as the facts, a particularly convincing one. For what is regarded as the nature

of most unconscious states is such as to make them easily recognisable as either cohering, or being sharply incompatible, with our ordinary, consciously held beliefs, attitudes and values; typically they appear to have a simplicity and a starkness far removed from the kind of complication and obscurity which lead to failures to detect incoherence and incompatibility between particular ones of our conscious beliefs, attitudes and values. But we cannot be swayed by such considerations, for such supposed 'facts' about unconscious mental states rest on the assumption of the very Freudian theories of repression and psychic structure of which we wish our account to remain innocent.

We must then consider the question in light of cases, and rest content with sufficient support for the generalisation that, sometimes at least, it is true that unconscious unconscious reasons meet the requirement in question and thus may be said to count as good reasons.

Let us look first at a case of solely unconsciously motivated behaviour. In Example C given earlier, P unconsciously selects examples to flatter her listener Y out of her unconscious wish to have Y respond warmly to her. It seems plausible to suggest that no incompatibility need obtain between P's consciously held beliefs, desires and attitudes (including her attitude towards Y, which might, for instance, be one of sheer indifference), and the beliefs, desires and attitudes comprising her unconscious motive.

In contrast, as it stands the case described in Example F of P's partially unconsciously motivated refusal of the invitation, clearly fails to meet the coherence requirement. Since it is built into the case that P's dislike of Q is inconsistent with P's consciously experienced and expressed attitude towards Q, P's consciously motivated refusal must be said, in that respect at least, to fail to be occasioned by good reasons, and hence must fail to be a reasonable action (although it might well still be the or a rational thing to do according to rationality formulating (1), in the sense of being in the agent's interests). But we can change the example in order to make it accord with the coherence requirement. Let us say that far from liking Q, P had always tended to

dislike her, and is well aware of this. Nevertheless, when the invitation is issued, and P begins to express her refusal, she does not experience her dislike as any part of her motive for refusing. Were she stopped in mid-sentence and accused of letting her dislike influence her decision, she would say, sincerely, that while she did dislike Q, her need to work on her thesis, and not her dislike, had prompted her refusal on this particular occasion.

But unbeknownst to her, P's dislike of Q has (partially) motivated her. Now when, in mid-sentence in making her refusal, P's flash of insight comes to her, she continues to dissociate her feelings of dislike for Q with her not wishing to go, but she readily admits to the possibility of this additional, hitherto unconscious, link between her dislike and her refusal. There are actually two possible accounts of what might have occurred in this case. It might be supposed that the case is now analogous to the one described in Example D, where all that is unconscious is a belief about, or awareness of, the nature of the relation between the motive (some mental state(s)), and the behaviour. In that case, P's flash of insight involved becoming aware that her conscious dislike of, and her disinclination to see, Q, had been influential in bringing about her behaviour (the refusal), or, more strictly, was at that very time influential in bringing it about. But the difficulty with this interpretation is that it allows that after her flash of insight, P was no longer unconsciously motivated at all: her motivation was totally conscious as she finished her apology. So in order to retain our example as one of unconsciously motivated behaviour, we must give an alternative account of what occurred. We must say that P disliked Q both on a conscious and on an unconscious level. It is compatible with this account to say either that P's conscious dislike of Q was influential in effecting the refusal, or equally, that it was not: the important dislike, and the dislike of which P's flash of insight made her aware, was the unconscious dislike. (P's disliking the same person both consciously and unconsciously would have to involve something like the following: consciously P dislikes Q because she is selfish and politically conservative, while unconsciously she dislikes her because Q symbolises her hated rival sibling; thus P consciously dislikes Q as selfish, conservative,

and unconsciously dislikes her as hated sibling.)

The above, then, is one case in which the person who is partially unconsciously motivated may be aware of all of his or her reasons at the time at which the action takes place, and may continue to act upon them, when the reasons themselves are good reasons inasmuch as they meet the coherence requirement set out earlier. As stated, that case has some features which make it unnecessarily complicated. In particular, because the unconscious motive and part of the conscious motive are not merely compatible, but are, at least in one respect, identical - both consisting, as they do, of an attitude of dislike felt towards Q, we were forced into speaking of the motive as being at once conscious and unconscious at the same time: a manner of speaking which is a little confusing.

However, we can avoid this difficulty, since bearing a relation of identity to it is not the only way in which one mental state or attitude can be compatible with another. Let us alter the example in the following way: P accepts the invitation, and her acceptance is motivated in part by her conscious wish to see her cousin, who will be at the party, and in part by her unconscious wish to see her old flame X, who is also likely to attend. Midway through writing a letter of acceptance, light dawns and P realises that an additional reason is involved. Her consciously felt attitude towards seeing X might be anything from indifference to very mild apprehension, at least it is not eagerness. But her desire to see her cousin is strong, and despite her consciously experienced indifference towards the prospect of seeing X as well, she decides to accept anyway - while admitting that some of her enthusiasm for her cousin might be deflected from her unconscious wish to see X, that is, while admitting that her decision might be partly unconsciously motivated. Now, since P's wish to see her cousin and her unconscious wish to see X are in no way incompatible, and neither are her consciously felt attitude of indifference or mild apprehension towards the idea of seeing X and her unconscious wish to see X, then we seem to be able to conclude that her unconscious wish, once its presence is known to P, counts as a good reason at least in as much as it is not incompatible with any of

P's other beliefs, attitudes and values.

We can conclude that if not always, then at least sometimes, both in cases of solely and of partially unconsciously motivated behaviour it is possible for 'unconscious reasons' to count as good reasons according to one criterion by which that evaluation can be made: coherence.

With regard to the second half of the requirement for a reason's being a good reason, viz. that it comprise none but rationally held beliefs and rationally reached desires, we face additional difficulties. These difficulties are multiplied if we hypothesise that the unconscious beliefs and desires in question were formed at an unconscious level: the whole notion of unconscious 'thinking', or 'reasoning', brings with it so many further conceptual problems and is, at least on the face of it, so vague and tenuous as to be best left unmentioned. And we are not obliged to take the further step of introducing it. For in the first place, there is one kind of unconscious motivation, as I have discussed already, which merely concerns consciously held beliefs and desires. (The kind of unconscious motivation to which I refer is that in which the person is unconscious of the connection between his or her (consciously held) beliefs and desires, and the behaviour which, unbeknownst to him or her, they occasion.) And there is no objection to the notion that ordinary consciously held beliefs and desires involved in unconscious motivation might not themselves have been reached by ordinary, conscious rational procedures.

The example used to illustrate the latter kind of motivation was the one of Freud's obsessive patient. In discussing the possibility of this girl's neurotic actions being in her interests, I raised the possibility that the beliefs and wishes which motivated her were not necessarily wholly unreasonable in the sense of being unjustifiable. As long as they could be seen to be defensible at all, then it seems possible to hypothesise that the patient herself had orginally arrived at them by way of a conscious rational procedure.

With regard to the other kind of unconscious motivation, in which the beliefs and desires motivating the behaviour are unconscious (although the agent must be aware that he or she has them, in order to fulfil the requirement for reasons expressed in rationality formulation (3)), we can insist that, at least in the cases of unconscious reasons with which we are dealing, the beliefs and desires in question were formed at a conscious level, and merely became unconscious after they were formed. And if they were formed by the same sort of process of conscious deliberation and reasoning with which we form the more rational of our ordinary, consciously held beliefs and desires, then there seems to be no objection in principle to the idea that some at least of these unconscious reasons may count as good reasons. For instance, in the final formulation of Example F, page 216, in which P's acceptance of the invitation is motivated in part by her wish to see her cousin and in part by her wish to see her old flame X, we can hypothesise that the latter wish may have been rationally reached at an earlier time and at a conscious level. P may have believed that only by seeing X again could she ascertain the strength of her feelings for him and thus decide whether or not to marry his successor Y, and on the basis of this reasoning might have concluded that she ought to see X again - an eminently reasonable strategy, perhaps. So P's unconscious reason might itself have been well-grounded and rational. On the other hand, we can also imagine that her unconscious wish to see X might have failed to meet the criterion expressed in this half of the formulation under discussion. Although it had once been a conscious wish, which only subsequently became unconscious, her wish to see X may have been based on faulty reasoning (she may have based it on a romantic and unrealistic illusion that if X were to see her again, he would sweep her off her feet and they would live happily ever after). Or it may not have been 'grounded', strictly speaking, at all (it may have been a mere whim, which entered her mind without any preliminary at all). Had P's unconscious reason had either of the two non-rational histories described, then while it may be said to be a reason, it could not be said to be a good reason because it fails to meet the requirements expressed in rationality formulation (3). So only in the first case, in which P's reason was shown to have had a rational history, could it be said that P's action was

a reasonable one.

Thus, in conclusion, we can say that it is possible for some - but not necessarily all - behaviour which is partially unconsciously motivated and for behaviour which is solely so, to be described as rational action according to the account of rationality in formulation (3): that is, acting reasonably, or acting for good reasons.

Unconscious Motivation and Irrationality

Just as I have shown that not all unconsciously motivated behaviour is rational, and of that which is, not all is rational in the same way (some being rational inasmuch as it meets the requirements expressed in formulation (2), formulation (1), or both, or by meeting those expressed in formulation (3) for acting reasonably), so not all that unconsciously motivated behaviour which fails to be rational does so in the same way. If we consider the two philosophically interesting senses of 'irrational': irrational$_2$ or non-rational, where the person fails to proceed rationally despite the fact that he or she could have done so, and irrational$_3$ or a-rational, where the person lacked the necessary control of himself or herself to be said to be able to proceed rationally, we can see that only with the case of clear compulsions of the sort described previously, do we have instances of unconsciously motivated behaviour which is irrational$_3$. Much other unconsciously motivated behaviour which is less than rational, for example the case described on page 218 of P's accepting the invitation when it was merely in respect of its history that her wish to see X was seen to be less than fully rational, would appear just to be irrational$_2$ or non-rational. If it were possible to establish that unconscious mental states were themselves merely causes, and that the explanations introducing them were simple causal explanations, then of course this would not be so, and all unconsciously motivated behaviour would indeed count as irrational$_3$. But, as was shown, this cannot be established.

The Rationality of Consciously and Unconsciously Motivated Behaviour Compared

I have shown that certain kinds of unconsciously motivated behaviour, both that which is only partially unconsciously

motivated and some which is wholly so, can be described as fully reasonable. It is interesting to note that the sorts of cases which do count as rational in this strong sense are neither those cases which are associated with the commonest actions of full-blown neurotics, that is, actions in which symbolic identifications are involved, nor with those which Freud thought of as the typical unconsciously motivated behaviour of everyday, and thus normal, life, that is, the class of slips, errors and omissions he classified as parapraxis.

In discussing the rationality of Examples A-F in the earlier part of this chapter, I examined the question of whether they met rationality requirements of the sort expressed in formulation (1). In addition, I considered the rationality of undertaking each of the actions described relative to undertaking the same action with the same motivation were the motivation thoroughly conscious. And it was concluded, in each case in which it was judged rational, that it was at least as rational, and apparently sometimes likely to be more so, in the sense of 'rational' under discussion - that is rational$_1$ - to perform the action out of the unconscious motive as it would have been to have performed the same action from the same motive had that motive been fully conscious. When we turn to a consideration of the same contrast in the case of the two actions which have been shown to meet the stricter rationality requirements of formulation (3) discussed in the latter half of the chapter, however, we find that the situation is somewhat different.

If we compare the case of the girl continuing to choose flattering examples although aware of her unconscious motives for doing so, with the case of a person who did so out of the thoroughly conscious motive of wanting to flatter her listener, then as long as we build in that the motives in each case meet the requirements for being good reasons, there seems to be nothing to choose between them in relative reasonableness. The two actions described would appear to be equally rational in sense (3). Similarly, if we compare the case of the girl (P) accepting the invitation when her motivation, while unconscious, or partially so, was nevertheless known to her, with that of someone (R)

having the same motives as conscious motives, and if we build in the same qualifications, then there would appear to be no reason to suppose that P or R acted any more or less rationally$_3$. The difference between acting on a conscious motive and acting on an unconscious one of which one is aware does not alone, all things being equal, make a difference to the rationality of the action. What is important is that the reason (whether conscious or unconscious) on which the action is based has a rational history.

What, then, is the significance of the contrast drawn in the case of the criteria first discussed, and the apparent possibility that someone's unconsciously motivated action might be more rational according to the characterization given in formulation (1) than the same action undertaken by a thoroughly consciously motivated person? To the extent that rationality (1) can be said to reflect a genuinely distinct formulation of the notion of rationality, then we have a counter-intuitive conclusion, viz., at least according to this formulation of rationality, it is sometimes more rational to act out of unconscious motives than out of conscious ones.

Procedural Policies

Can we from this last generalisation derive a rule of action, perhaps what might be described as a policy advocating the wisdom of forgetting? It is distinctly counter-intuitive, but such a policy, or regulatory motto, deserves a little further attention.

We must first distinguish two formulations of the policy, a more general one to the effect that one ought in general to forget or be unaware of one's mental states and a more modest one to the effect that it sometimes is rational to forget or be unaware of one's mental states. The latter perhaps hardly rates the description of a policy at all, yet it is all which can safely be proposed. For we could only propose the former, more general, policy, were it the case that what we have found to be true could be shown to be so in a great number of cases, a number greater than the number of cases in which an open mind and knowledge were associated with the furtherance of one's interests. Without the

necessary empirical evidence underlying the assertion of the general policy, we are bound to restrict ourselves, clearly, to the more modest one. But even if we limit ourselves to the more modest formulation, could it ever be correct? However modest the policy, there seem to be difficulties in the way of proposing that a conclusion concerning the rationality of some unconsciously motivated behaviour alone could lead us to adopt some policy of action.

To begin with, proposing a policy about the wisdom of forgetting seems to fly in the face of good sense since neither forgetting, nor its technical corollary in Freudian theory, repression, are in the ordinary way considered voluntary or intentional actions at all. We cannot in any ordinary or simple way consciously decide to forget, nor can we forget on demand. And a regulatory policy of the sort expressed in the slogan above seems at least to imply that the activity involved entails such a capacity.

Someone might now propose the following objection. If forgetting or repressing is not itself to be regarded as a consciously intentional and voluntary activity, then how can its outcome in behaviour (viz. cases of unconsciously motivated behaviour like those I have been discussing) itself be assessed as rational or otherwise? For to judge some action as rational is to imply a degree of consciously experienced intentionality on the part of the agent. This objection can, however, be shown to rest upon an unnecessary and dangerous conflation of several distinguishable phenomena. I remarked earlier (page 143) that while a person's being in the grip of a passion is not usually described as rational or irrational, since it is merely a passive state which the person undergoes, rather than something done , nevertheless it is possible that an action taken out of that state may be so evaluated. Refusing companionship out of my grief may be judged irrational, though grieving (within certain normatively established limits) would not. In the same way, while the forgetting or repression which precedes the unconsciously motivated behaviour may not be in any way intentional or voluntary, and thus may not itself be a candidate for being assessed according to its rationality,

still the possibility remains that behaviour or actions stemming from such repression may be so assessed.

But will not the concept of self-deception help here? Although we cannot speak of the rationality or otherwise of forgetting or repression, can we not speak of the rationality or otherwise of those activities which are said to comprise or occasion (according to whether it is regarded as a motive or an activity) self-deception? Pretending to oneself, engaging in wishful thinking, neglecting to attend to certain beliefs and desires or the evidence supporting or destroying other beliefs and desires, and refusing to "spell out" in the phrase of one theory of self-deception,[10] all may be undertaken fully consciously and intentionally. According to philosophers writing about self-deception, activities such as those listed above effect or bring about forgetting or repression. And while, as I remarked in Chapter One, the activity of self-deception may not always involve such deep or thorough forgetting as that implied in the technical sense of 'unconscious of' and 'unconsciously' used here, it involves some degree of unconsciousness nevertheless. It may be said that, while forgetting and repression are not directly under our control, yet indirectly they can be voluntarily and intentionally brought about, by the practice of the above kinds of activity. In the same way, it may be said, we cannot directly and voluntarily alter our beliefs or our emotional attitudes by a mere act of will, but indirectly we can do so, in certain cases with considerable success, by, for example, forcing ourselves to dwell upon certain information (I may be able to alter my attitude of loathing for X by reminding myself of what X has suffered). And we seem to be able to do something towards forgetting or repressing mental items in just the same kind of indirect way.

There are several reasons why it seems that philosophers may have been premature in assuming that the activity of self-deception can be introduced in this context and that we can helpfully discuss regulatory principles of the sort described above in relation to it. The first such reason is empirical: it seems to be an ordinary empirical question, as yet in want of ordinary empirical proof, to say that engaging in activities of the kind described

above as self-deceptive actually leads to the forgetting or repression of mental states.

The only thing which may be said to follow (conceptually) is that not attending to P is part of what 'forgetting P' and 'repressing P' mean. But of course since it follows, if it does at all, conceptually, it would be to beg the question at issue to assume that the intentional action of withdrawing attention from P is likely to bring about the situation in which the agent has forgotten or repressed P. Certainly we can say that attention is withdrawn from P, but whether the withdrawal causes or occasions the forgetting or repression and may thus be a conscious and intentional strategy on the agent's part or whether it merely comprises the non-intentional forgetting itself, we cannot say.

In addition, if we accept Freud's word on this matter, such a regulatory principle is quite impossible. For, contrary to what he had at a theoretical level assumed, Freud discovered from his patients that repression, or what he then called 'defence', takes place without the patient's conscious awareness. While Freud supposed that defence or repression was no less intentional because of that (and we are free to draw the same conclusion), still a policy of the kind we have been considering has no application for the concept of action or activities which are anything less than consciously intentional and voluntary.

Thus we do not seem entitled to conclude any further policy of behaviour from the truth of the assertion that to act out of unconscious or partially unconscious motives sometimes can be to act rationally.

There is still one policy which our findings invite us to embrace: the policy of striving towards self-knowledge and self-awareness. For since, as has been shown, in the case of unconsciously motivated behaviour, only those actions stemming from unconscious reasons of which the agent has become aware, may merit the description of action which is reasonable, a person can be said to act rationally to the extent that he or she can give or is aware of the reasons or motives underlying his or her action.

Rather than conclude, then, that repression or forgetting can themselves be said to be rational procedures or strategies, we need only conclude, with Freud himself, that the ideal of knowing onself, in the sense of coming to know of one's unconscious motivation and mental states even if not to fully embrace them as one's own, is a rational ideal to pursue.

This last claim has usually been interpreted in psychoanalysis as a claim based on the value inherent in the consequences of pursuing such an ideal. Self-awareness, it has been thought, leads to admitting, embracing and 'working through' one's unconscious motives, and, in doing so, eventually losing them. And only by doing that, it has been supposed, would we come to act more rationally. But I have shown that, at least in principle, rationality might be the goal of self-knowledge in a much more direct way. For to act on an unconscious motive of which one is aware might itself be to act perfectly rationally.

* * * * * * * * *

Overview and Conclusion

Historically there has been a link between unconscious motivation and Freudian theory. Freud's view was that unconscious mental states can be referred to in explaining certain types of behaviour. He believed that such behaviour was intrinsically irrational and was produced in a manner quite unlike the way in which ordinary conscious beliefs and desires bring about action. In Chapter One, however, I argued that it was both desirable and possible to disentangle the concept of unconsciously motivated behaviour from Freudian theory and for it to be examined in its own right. My ultimate intention was to show that the concept of rationality is neutral between consciously and unconsciously motivated behaviour.

In Chapter One I distinguished the everyday concept of the unconscious from the technical Freudian concept and argued for the need for an account of a technical concept which is at once free from additional theoretical hypotheses and distinguishable from the everyday non-technical concept with its implication that the unconsciousness in question is a merely contingent and alterable matter.

It was seen that sometimes unconsciously motivated behaviour is defined by using the aspect of hindsight usually associated with post-therapeutic revelations of unconscious states whereby unconscious states are those which we cannot be aware of at the time at which they are influencing our behaviour. Despite this use of hindsight, the plausibility of also describing cases of people who continue to act upon what seem to be unconscious motives even after having become aware of them was postulated and subsequently established through argument based around and drawing on some of Freud's case studies. The nature of this awareness is, however, somewhat problematic. I eventually arrived at a technical definition such that unconscious mental states are those of our mental states of which we are unaware and of which we cannot become aware at the time at which the behaviour and states which they are introduced to explain are occurring, except in a way which is always qualitatively distinguishable from the way in which we are aware of our ordinary conscious mental states. The difficulties of characterizing this kind of awareness of unconscious motivation were explained and some analysis of the related notion of self-

deception was made.

In Chapter Two I approached the question of how it is that unconscious motives might be said to give rise to action by an analysis of the general problem of how causes and motives give rise to behaviour. Two types of motive and motive type explanations were identified, only one of which is as precise and purposive as to imply the presence of particular goals or ends on the part of the person to whom the motivation is ascribed, and beliefs about the likelihood of the behaviour in question bringing about those goals or ends. I then contrasted explanations citing motives with various other kinds of explanations such as that involving causes and, particularly, attention was drawn to the way in which motive type explanations differ from explanations citing signs and symptoms.

Freud's theory of unconscious motivation was then discussed and it was shown that, while some of the phenomena which Freud cited seem to require the description of "motivated by" unconscious mental states, other behaviour would appear to be better described by saying that it is a sign or symptom produced or caused by such states. In analysing the cluster of concepts relating to motives, reasons and intentions, I treated, for the purpose of the thesis, "motive" as the central concept. I accepted the claim that intentional action is a class of action for which there are reasons and that some actions for which there are reasons are unmotivated. The reason for any motivated action is just the motive with which it is undertaken.

In Chapter Three I argued that the notion of having a reason for an action is not logically tied to being able to be or to become aware or conscious of that reason by a simple act of introspection. I showed that there might be an indeterminate, rather than finite, number of beliefs and desires which might be reasons for a particular action and that, therefore, the agent will always be unable to give all his reasons for his actions. Moreover, I argued that in certain special but everyday situations, which differ from ordinary ones only in the special respect concerned, we allow that reasons are present of which the agent is unaware and

unable to become aware. I discussed the inconclusive and moot nature of conceptual arguments relating to unconscious reasons and intentions, but I concluded that there is, ultimately, no reason to doubt the conceptual possibility of unconscious reasons and intentions.

In Chapter Four the notion of rationality was explained and three different ways in which 'rational' qualifies particular actions were distinguished. I argued for three main types of rationality:

(1) 'Acting rationally' as acting in such a way that one furthers one's interests;

(2) 'Acting rationally' as having a rationale or purpose in acting; and

(3) 'Acting rationally' as acting reasonably, that is having and being moved by good reasons in acting.

Having analyzed in some detail these three main accounts of what it is to act rationally, I made certain points about the notion of irrationality.

In Chapter Five I attempted to draw the thesis together, drawing on the work of earlier chapters and using detailed case study examples. The possibility of the rationality of unconsciously motivated behaviour was thereby established. From the analysis of the concept of rationality in Chapter Four, it was shown that to ask whether unconsciously motivated behaviour is ever rational is actually to ask a number of different questions. In particular, it is to ask whether unconsciously motivated behaviour can be described as rational in any of the three different ways of characterizing the rationality of action described in the previous paragraph. The general claim that all unconsciously motivated behaviour is rational was rejected as too general to be plausible, as was the contrary claim that unconsciously motivated behaviour never is, and never could be, rational. I tried to argue for a middle position in which while some unconsciously motivated behaviour may be regarded as rational in each of the three ways described, some may not. My belief is that it is inaccurate to characterize all unconsciously motivated behaviour either as rational or as irrationl.

Further, I argued that, according to the different ways of using the term 'rational', it is possible to evaluate both consciously motivated and unconsciously motivated behaviour alike in terms of their rationality. I concluded that just as some consciously motivated actions are rational in each of the ways described, while some are irrational (in both of the philosophically interesting senses: irrational$_2$ or non-rational, as well as irrational$_3$ or a-rational), so some unconsciously motivated behaviour or action, in particular some at least of that which is attributed to psychologically normal, as distinct from neurotic persons, is rational, in some of the different ways of conceiving rationality (described in Chapter Four), while some is irrational.

Thus I drew a sharp distinction within the class of behaviour which is plausibly explained in terms of unconscious beliefs, feelings and motives. Some behaviour of that class, for instance the compulsive handwashing of the obsessive, is completely irrational, while other instances of behaviour in the class - for instance, much of the behaviour which we encounter in what Freud called the psychopathology of everyday life, as well as some of the unconsciously motivated behaviour exhibited by neurotics - is rational.

* * * * * * * * *

Notes and References

1 See John Hospers's "What Means This Freedom?", reprinted in Berofsky, ed., Free Will and Determinism, Harper and Row, 1966. All page references are to this edition.

2 Among philosophers and psychologists writing on this topic, the only significant dissenting voice, interestingly enough, is Freud's own. Although he nowhere dealt fully with the question, remarks of Freud's ("Moral Responsibility for the Content of Dreams", Collected Papers, Vol. V) suggest that he believed us to be responsible for our unconscious, as well as for our conscious, thoughts, feelings and wishes, and we can perhaps extrapolate from that to a presumption that he might have supposed us similarly to be responsible for actions which derived from those mental states in being motivated or produced by them.

Freud's arguments in support of this thesis deserve comment. In the first place, he argues that because evil thoughts (and, we may add, acts) are perpetrated, and some agent must be responsible, the person from whose unconscious mind they spring must take responsibility for them since no one else will:

> "Obviously, one must hold oneself responsible for the evil impulses of one's dreams. In what other way can one deal with them? Unless the content of the dream.....is inspired by alien spirits, it is part of my own being....and if.... I say that what is unknown, unconscious and repressed in me is not my 'ego' then I shall not be basing my position upon psychoanalysis, I shall not have accepted its conclusions.....It is true that in the metapsychological sense this bad repressed content does not belong to my 'ego'..... but, to an 'id' upon which my ego is seated. But.....for any vital purpose, a separation of the ego from the id would be a hopeless undertaking."
>
> (Page 156-7, op. cit.)

And, in the second place, he argues from the fact that even normal people do feel guilt for their unconscious states as well as for their conscious ones, to the conclusion that the implied assumption of responsibility which such guilt expresses, is itself justified and proper:

> "Experience shows that I.....do take....responsibility.... Psychoanalysis has made us familiar with a pathological condition, the obsessional neurosis, in which the poor ego feels itself responsible for all sorts of evil impulses which are brought up against it in consciousness but which it is unable to acknowledge. Something of this is present in every normal person.... It is just as though we could say that the healthier a man is, the more liable he is to contagions and to the effects of injuries. This is no doubt because

conscience is itself a reaction-formation against the evil that is perceived in the id."

(Page 157, op. cit.)

The former argument appears simply to beg the question. The latter argument rests on an explicit non sequitur unless one accepts, as Freud does, that the normal, as distinct from pathological, feelings of guilt which the normal person has precisely are feelings of guilt which are justified and appropriate, that is, feelings of guilt over an action or thought for which the agent is responsible - in which case it fails as an argument in following necessarily from exclusively psychoanalytic truths. In fact, we would want to dismiss the psychoanalytic manner of distinguishing normal from pathological, or as it might better be expressed 'appropriate' from 'inappropriate' guilt, as overly crude. For we want to allow for the possibility of the normal person having guilt which is inappropriate. And once we do so, then we are, of course, committed to saying that if unconscious motivation is correctly to be regarded as an excuse, then a normal person's feeling guilty over his or her unconsciously motivated behaviour, is a case of inappropriate guilt.

3 Hospers, op. cit.

4 Oxford English Dictionary.

5 Cameron, N., Personality Development and Psychotherapy, Houghton Mifflin, Boston, 1963, Page 392.

6 John Austin, "A Plea for Excuses", reprinted in J.L. Austin, Philosophical Papers, Oxford, 1970. All page references to this edition. Page 204.

7 Hospers, op. cit.

8 In terms of the standard Austin/Grice/Searle discussions of speech acts, I am at most proposing a further class of illucutionary acts: the illocutionary force of the speech acts mentioned lies in their expression of unconscious beliefs and thoughts. (According to the Gricean analysis of what it is to mean something, verbal slips must surely constitute a degenerate class, since, if my analysis of them is correct, they involve meaning without intending to produce an effect in the hearer by getting the hearer to recognize that intention.)

9 Graham Dawson provides some interesting discussion and examples of the rational, irrational, justified and unjustified uses of speech acts in his article: "Keeping Knowledge Under Control: Philosophy and the Sociology of Educational Knowledge", Journal of Further and Higher Education, Winter 1977.

10 Fingarette's, op. cit.

Only those books and articles of special help and relevance are listed here.

Alexander, P., 1962, "Rationality and Psychoanalytic Explanation, Mind.

Anscombe, E, 1957, Intention, Blackwell.

Bennett, J., 1964, Rationality: an essay towards analysis, Routledge, and Kegan Paul.

Collins, J., 1969, "Unconscious Belief", J.Phil.

Davidson, D., 1963, "Actions, Reasons and Causes", J.Phil.

Edwards,P., 1967 and 1972 edns., Encyclopedia of Philosophy, Collier-Macmillan.

Edwards, P., and Pap, A., (eds.), 1973 edn., A Modern Introduction to Philosophy, Collier-Macmillan.

Fingarette, H., 1969, Self-Deception, Routledge and Kegan Paul.

Flew, A., 1949-50, "Psychoanalytic Explanation", Analysis.

Freud, S., 1900, The Interpretation of Dreams, Avon Books edn., 1965.

Freud, S., 1957 edn., Collected Papers of Sigmund Freud, Hogarth Press.

Gustafson, D.F., (ed.), 1967, Essays in Philosophical Psychology, Open University Press.

Hart, H. and Honore, A., 1959, Causation and the Law, O.U.P.

Hollis, M., 1977, Models of Man, C.U.P.

Kenny, A., 1963, Action, Emotion and Will, Routledge and Kegan Paul.

Longford, G., 1972, Human Action, Macmillan.

MacIntyre, A., 1958, The Unconscious, Routledge and Kegan Paul.

Melden, A., 1960, "Willing", Phil. Rev.

Melden, A., 1961, Free Action, Routledge and Kegan Paul.

Mullane, 1971, "Psychoanalytic Explanations and Rationality", J.Phil.

O'Connor, D., 1972, Free Will, Macmillan.

Peters. R.S., 1959, The Concept of Motivation, Routledge and Kegan Paul.

Peters, R.S., Dearden, R.F. and Hirst, P.H. (eds.), 1972: Education and Development of Reason, R.K.P.

Richards, D., 1971, Theory of Reasons For Action, O.U.P.

Rieff, P., (ed.), 1963, Freud: General Psychological Theory, especially "The metaphysical essays".

Ryle, G., 1949, The Concept of Mind, Penguin.

Shope, 1973, "Freud on Conscious and Unconscious Intentions", Inquiry.

Strachey, J. (ed.), 1953-1964, The Standard Edition of the Complete Psychological Works of Sigmund Freud, Hogarth Press.

Toulmin, S., 1948-9, "The Logical Status of Psychoanalysis", Analysis.

White, A. (ed.), 1968, The Philosophy of Action, O.U.P.

Wolheim, R., 1971, Freud, Fontana.