

INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.
- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.
- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35mm slide or as a 17"x 23" black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35mm slides of 6"x 9" black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.

8708298

Lauer, Helen

REASON-GIVING STATEMENTS

City University of New York

PH.D. 1987

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

Copyright 1987

by

Lauer, Helen

All Rights Reserved

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy ✓
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages _____
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Dissertation contains pages with print at a slant, filmed as received ✓
16. Other _____

University
Microfilms
International

REASON-GIVING STATEMENTS

by

Helen Lauer

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

1987

©1987

HELEN LAUER

All Rights Reserved

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

December 10th 1986
Date

Arnold Koslow
Chair of Examining Committee

Dec 10, 1986
Date

Arthur W. Cole
Executive Officer

Charles Landmesser
David Rosenthal
Arnold Koslow
John Cunn
Supervisory Committee

PREFACE

One very obvious and rather marvelous fact about explanations in the human sciences is that the subject matter talks and sometimes literally explains itself. This dissertation is an essay about what sort of difference language participation makes to explaining what language participants do. For although it is commonplace to observe that explanation of human behavior diverges from explanations of other sorts, it is not at all easy to tell exactly what this divergence amounts to.

For instance it might seem obvious that when I give you my real reason for acting, what I say conveys the thoughts that caused my action and this is why you are correct to accept my reason-giving statement as explaining my action. But it's so far from being obvious just how this sort of causal explanation might work that some philosophers have questioned whether actions even have causes.

Currently action theorists are recruiting insights from philosophy of natural language to shed light on the relation between actions and the reasons that explain them. The four contemporary action theorists studied here maintain that certain thoughts do cause actions. And further they maintain that from an agent's own reason-giving statement

one can infer the particular thoughts that caused his action. In Chapter I, I review how Hempel's standard deductive-nomological model of explanation fails to show how reason-giving statements might divulge action-causing thoughts.

It is often supposed that certain linguistic forms and expressions are reserved exclusively for act description, constituting a distinct mode of discourse in natural language. In Chapter II, I try to clarify what characteristics mark off all act descriptions as a class from other ways of referring to occurrences.

Different causalist action theorists put the spotlight on different linguistic facts about action statements. An initial disagreement about what to count as significant and crucial in the data of action talk may account for some persisting controversies about the nature of actions themselves. This suggestion is illustrated by example in Chapter III. Donald Davidson reaches his conclusions about particular actions from his way of parsing the logical structure of singular action sentences taken as a class. On the other hand, to build a theory of action Alvin Goldman contrasts the different meanings associated with specific action predicates occurring in individual sentences within that class. An apparent disparity ensues between Davidson's and Goldman's claims about the number of particular actions whose occurrence is entailed by a true singular action sentence. On close inspection of their views this apparent

disagreement seems insubstantial. Yet the literature reports a standing controversy among received views of act individuation, and points to Davidson and Goldman as classic contenders in this debate.

In Chapter IV, I examine the claim that causal facts are entailed in a special way by true action sentences. Jennifer Hornsby derives this result from observing the grammatical or 'surface' form of certain action sentences. She argues that there is a non-truth functional form of entailment that is characteristic of action sentences, distinguishing them from other modes of description like causal statements.

In Chapter V, I examine whether any light has been shed on the structure of reasons leading to actions by analyzing the structure of thoughts involved in language participation. Donald Davidson and John Searle independently argue that a reason-giving explanation of an act's occurrence parodies the correct interpretation of a text or an utterance. They both appeal to 'holistic' illustrations of an agent's primary reason as a causal batch of interdependent thoughts of specifiable types. Strong parallels emerge between speech interpretation and action explanation, and this affords us an excellent method of hindsight with respect to our theorizing about meaning and about action: We can test our progress by checking and balancing our hypotheses about these two aspects of mind relating to the 'nonmental' world.

Overall, the results presented in this essay go a modest way, only so far as to demonstrate that we know less today about reasons and actions than might be supposed. These four versions of the move from linguistic facts to non-linguistic conjectures about the structure of causal thoughts are highly suggestive, but so far inconclusive.

I wish to acknowledge a grateful debt to the patience and persistence of David Rosenthal, whose advice and example have been crucial to my philosophical progress.

TABLE OF CONTENTS

PREFACE	iv
Chapter	
I. INTRODUCTION OF CAUSALIST ACTION THEORY; WHY THE D-N MODEL FAILS	1
II. INDETERMINACY IN REFERENCE TO PARTICULAR ACTIONS. .	38
III. ACT INDIVIDUATION AND THE LOGICAL STRUCTURE OF SINGULAR ACTION SENTENCES	75
IV. GROUNDING CAUSAL ASSUMPTIONS IN THE SEMANTICS OF SINGULAR ACTION SENTENCES	112
V. HOLISTIC APPROACHES TO INTERPRETING CAUSAL SEQUENCES AS ACTIONS.	152
Appendix	
A. SELLARS' "PRINCIPLES OF INFERENCE" FUNCTIONING IN REASONS.	198
B. DAVIDSON'S CONDITIONS FOR "RADICAL INTERPRETATION".	203
C. SEARLE'S "LOGICAL PROPERTIES" DISTINGUISHING TYPES OF THOUGHT AND ATTITUDINAL STATES	208
SOURCES CONSULTED	215

CHAPTER I

INTRODUCTION OF CAUSALIST ACTION THEORY;
WHY THE D-N MODEL FAILSIntroduction

Consider a statement that an agent regards as giving his own reason for something he did intentionally. Suppose this reason-giving statement describes what the agent thought he was doing, or what he was trying to do or intended to be doing at the time he acted. Now suppose further that one accepts this statement of the agent's own reason as a correct explanation of the occurrence called his action. How does the agent's reason explain what he did intentionally?

Some philosophers argue that from an agent's reason-giving statement one can infer the particular beliefs, desires or other attitudes that caused him to act. Proponents of this view might be dubbed causalist action theorists. They seem to argue that a reason-giving statement, or "rationalization" comprises "a species of causal explanation."¹

This thesis needs to be filled out and made more

¹Donald Davidson, Essays on Actions and Events (Oxford: Clarendon Press, 1980), p.3.

precise. But suppose that an agent's correct statement of his own reason does somehow identify the particular thoughts that caused his action to occur; how does it do so? In what way does the explanation of an action done intentionally differ from the explanation of any other causal sequence?

It has been claimed that statements conveying the fact that something has been done intentionally (for a reason the agent calls his own) share a distinctive feature that sets them off as a special class of descriptive expressions. The distinctive feature of certain act descriptions is supposed to be the indeterminate way they refer to particular nonrepeatable¹ sequences of events. What does this indeterminacy amount to? Is there a uniform feature or family resemblance that distinguishes descriptions of actions done intentionally from other ways of describing sequences of events?

For purposes of mentioning this perspective on action discourse in a general way let us use a cheap label: call this 'the uniqueness of indeterminacy in reference to actions'. Note various theses can be distinguished as falling under this label. On the one hand, one may claim that there is a unique or distinctive way of referring to actions; Chapters II through IV follow three versions of this claim. Alternatively, one may claim that there is a unique sort of thing called an action that gets referred to in a standard way; Chapter V explores two versions of

1. By repeatable and 'nonrepeatable' sequences and actions, I have in mind here the type and token distinction.

this second thesis.

1. First let us return to the 'deducibility of causal thoughts' thesis introduced in a very general way at the outset. For we can recognize this claim in the texts of some contemporary action theorists with relatively little digging.

To begin a cursory survey, note Donald Davidson's comment summarizing his project in "Actions, Reasons, and Causes" (1963). The comment appears in a footnote to his later essay, "Agency" (1971):

In "Actions, Reasons, and Causes" I try to show that although beliefs and desires (and similar mental states) are not events we can properly say that they are the causes of intentional actions, and when we say this we draw upon the concept of 'ordinary event causality' . . . [i.e.] the relation, whatever it is, that holds between two events when one is the cause of the other.¹

In "Actions, Reasons, and Causes" Davidson proposed that a statement "rationalizes"² an event called an act done intentionally just in case one can infer from it at least one of the agent's beliefs or pro-attitudes that caused that event to occur.

Judging from some of his other essays in the same collection,³ it seems that by 'infer' Davidson has in mind here strict logical inference or deduction. For instance, when developing an "account of acting on a reason" adequate for problem laden cases of "incontinent

1. Davidson, Actions and Events, p.49.

2. Ibid., p.3.

3. See for example "How is Weakness of the Will Possible?" (1970), "Intending" (1978) and "Hempel on Explaining Action" (1976) in Actions and Events.

action"¹ Davidson reconstitutes the thought sequence comprising an episode of practical reasoning by parodying the apparatus of probability logic. He suggests the best way to analyze the "pattern of practical reasoning" is to treat it "like reasoning from probabilistic evidence."²

In the retraceable steps of any practical argument³ Davidson reads each premise to be the propositional content of a belief, desire or related pro-attitude, or some conjunction of propositions representing a cluster of such attitudinal states. The conclusion of a practical argument is read by Davidson to be either an "all-out judgment"⁴ or an occurrence of action.⁵ Again, from "Actions, Reasons, and Causes":

Corresponding to the belief and attitude of a primary reason for an action, we can always construct (with a little ingenuity) the premises of a syllogism from which it follows that the

¹.Davidson, Actions and Events, p. 39.

².Ibid., pp.37-38.

³.Ibid., p. 16. Davidson notes that the "practical syllogism . . . cannot be subtilized into a reconstruction of . . . the relative force of various desires and beliefs in the matrix of decision . . . [i.e.] the weighing of competing reasons"

⁴.Ibid., p. 99.

⁵.Ibid., p. 32. In his portrayal of the conclusion in a practical argument as one that has a distinctive character, Davidson seems to favor preserving the overall structure of the syllogism as a valid deduction--at this period in his writing on the structure of practical reasoning.

action has some . . . 'desirability characteristic'.¹

Davidson prefers to read the premises of a practical syllogism as individual "attributions of desires, preferences or beliefs" rather than "axioms of decision theory."² He remarks that decision theory models cannot illuminate the way courses of action are decided under "ordinary," non-"static" conditions,³ where agents may act for "bad" reasons⁴ and on the basis of "conflicting desires."⁵ Yet all of Davidson's modifications of the decision theory model are predicated on his conviction that

being in a position to call a person irrational . . . or nonrational . . . presupposes that we have already found it possible to give reason explanations to his actions.⁶

Davidson quotes Hempel's description of causal explanations to depict the method of moving from premises to conclusion in a practical argument. Davidson takes no issue with Hempel's term "deduce" in this passage of Aspects.⁷ Later Davidson remarks

Hempel set out to show that reason explanations do not differ in their general logical character from explanations in physics or elsewhere. My reflections reinforce this view.⁸

Davidson's main disagreement with Hempel is over the content of law-like premises and the scope of the subject

¹. Actions and Events, p. 9. In footnote 4, Davidson acknowledges there is a controversy over whether the practical syllogism should be characterized as deductive or not. He then refers to an essay wherein his own depiction of practical reasoning depends upon deductive rules of inference, see above p.4,n.2.

². Ibid., p. 273. ³. Ibid., p. 271. ⁴. Ibid., p. 266.

⁵. Ibid., p. 269. ⁶. Ibid., p. 267. ⁷. Ibid., p. 274.

⁸. Ibid., p. 265.

matter in conclusions that follow from successful reason explanations. Davidson has launched a famous objection to the ambiguous way Hempel presents the role of causal laws in explaining individual episodes of acting intentionally.¹ I will not be able to address the rich subject of how causal laws contribute to reason-giving explanations. It suffices for our purposes here to note that Davidson regards reasons as causes of the individual actions they explain:

Central to the relation between a reason and the action it explains is the idea that the agent performed the action because he had the reason.²

In A Theory of Human Action (1970), Alvin Goldman "contend[s] that reason explanations are a species of causal explanation."³ He argues that "acts are caused by mental events."⁴ Quoting Goldman:

. . . that S flipped the switch in order to turn on the light implies that . . . S's having the indicated want . . . and belief caused or resulted in his flipping the switch.⁵

In Actions (1980), Jennifer Hornsby remarks:

It can be said that no event is an action unless the agent's thoughts and wants cause that event

¹. Davidson, Actions and Events, p. 274.

². Ibid., p. 9.

³. Alvin I. Goldman, A Theory of Human Action (Princeton: Princeton University Press, 1970), p. 80.

⁴. Ibid., p. 91.

⁵. Ibid., p. 78.

to occur.¹

For each episode of an action done intentionally, Hornsby posits the unlocatable occurrence of a 'trying' inside the agent causing the occurrence of events locatable at the surface of the agent's body and beyond his body.² Without tracing her argument now, it is worth noting that Hornsby presents her theory about the existence of inner tryings as based upon a line of argument offered by H.P. Grice in defense of his causal analysis of visual perception.³

John Searle also appeals to an analogy with visual perception in his causal analysis of acting intentionally. In his essay "The Intentionality of Intention and Action" (1979) Searle claims there are two components of each act done intentionally: the agent's "intention in action" and some bodily movement or other physical event caused by the agent having just that specific "experience" of doing something intentionally.⁴ A correct description of the agent's "experience of acting" identifies both the "intentional" cause and the "physical" effect constituting distinct "aspects" of the "complex event" called his act

¹·Hornsby, Actions, p. 5. ²·Ibid., p. 33.

³·H.P. Grice, "The Causal Theory of Perception," Aristotelian Society Proceedings 35, supplement (1961): 121-152.

⁴·John Searle, "The Intentionality of Intention and Action," Inquiry, 22 (1979): 253-280, p. 261 (hereafter cited as Searle, "Intentionality").

done intentionally.¹

2. Before going further in this survey with different theorists' formulations of the 'indeterminacy' of act descriptions, it is appropriate to clarify the particular sense in which a reason or cause may be given by an agent's own reason-giving statement.

We will be focussing exclusively on a reason-giving assertion that correctly gives the 'real' reason for what an agent has done intentionally or for what he thinks he will do. I set aside many cases where agents make false assertions about why they act. Agents are sometimes motivated to be deceptive about the reasons for their behavior and at other times they are mistaken about why they are acting a certain way. These are different sorts of case where it makes sense to distinguish the statement an agent actually gives as his 'own' reason from another that identifies the 'real' reason for his action.

There are still other cases where an agent states a reason for doing something that is not his own reason. There may be reasons for doing things that are not recognized by any agent as being his 'own' reason. For example there may be a reason to begin rebuilding the Graduate Center this afternoon. Perhaps such reasons exist in some way that is parallel to the existence of a proposition that can be stated but that no one believes to be true. In any case, it is easy to recognize a contrast between the state-

¹ John Searle, "Intentionality," p. 175.

ment of a proposition or the utterance of a sentence on the one hand, and the state of believing a proposition or the state of believing the content of a sentence on the other.

I assume that there is a comparable difference between stating a reason for doing something and being in the state of having one's own reason for doing something. In this study we will focus on just those occasions when an agent's stating a reason indicates that he is in the state of having his own reason for something he has done or is doing or thinks that he will be doing in the future. I will be neglecting the different sorts of case where someone clearly does something intentionally yet for 'no apparent reason'. I set these restrictions in order to study how the sentence an agent utters when he states his own reason can illuminate the very thoughts¹ causally connected to the occurrence of his action.

3. Here are some grounds for regarding an agent's own reason-giving statement as a distinctive sort of causal explanation.

According to one classic model, an explanation is a set of statements characterized by two components called

¹. Throughout this essay I use the term 'thought' in a broad sense to refer to all sorts of propositional mental attitudes. In doing so I follow Vendler's usage. See Zeno Vendler, Res Cogitans (Ithaca: Cornell University Press, 1972).

an explanans and an explanandum.¹ An explanandum statement, e.g. 'The bridge collapsed', is a description of some event or phenomenon whose occurrence "was to be expected"² by one who believes another statement called an explanans, e.g. 'The support beams fractured from metal fatigue.' On this model an explanans is often figured as a conjunction of statements. But for some cases the model will figure a single statement containing both the explanans and the explanandum components of explanation:

The bridge collapsed because the support beams fractured from metal fatigue.

It is controversial whether the relation between an explanandum and its explanans is characterized solely as 'logical entailment'. But no matter how the main connecting terms in an explanation (e.g. 'because') are best defined, an explanans is supposed to provide "good grounds for believing" or for "predicting . . . deductively or with reasonable probability" the occurrence described by the explanandum.³

Now consider an explanation where the 'explanandum' statement describes the occurrence of an act done intentionally. The explanandum might refer to an act that has already

¹Carl G. Hempel, Aspects of Scientific Explanation (New York: Macmillan, 1965) p. 247 (hereafter cited as Hempel, Aspects).

²Ibid., p. 337.

³Ibid., pp. 178, 368.

been carried out, or to one that is anticipated for the future.

As an example suppose that one has grounds for believing the truth of this act description:

(1) Jack sabotaged the bridge at 3:00.

Grounds for believing this sentence might include finding Jack's fingerprints on a bomb plunger after 3:00 nearby the shattered bridge.

Or suppose conditions warrant one's asserting this description of a future occurrence:

(1') Jack will sabotage the bridge at 6:00.

Conditions that warrant one's asserting sentence (1') might include finding Jack nearby the bridge installing and setting a time bomb.

In what follows, the example of Jack's sabotaging the bridge is described so that the relevant verbs in each of the numbered sentences comprising our example can be conjugated into the future tense. This is so the example can be altered to illustrate the common run of cases where agents give reasons for actions that they anticipate themselves doing as well as the reasons agents give for things they think they have already done.

Now let's try to demonstrate the deducibility of causal thoughts thesis using this model of explanation. According to the thesis under investigation an act done intentionally can be redescribed so that one can deduce another statement identifying the content of the agent's

causal beliefs, desires or other attitudinal states.¹

Suppose a redescription of Jack's activity at 3:00 comes about in this way. Jack is asked what he was doing or why he was in the area of the bridge at 3:00. In reply Jack gives this statement as his own reason:

(2) I was pleasing my uncle.

Jack's asserting (2) ordinarily indicates that

(3) Jack believes that he was pleasing his uncle at 3:00.

As a general assumption ordinarily made under so-called standard conditions, this much ascription of mental attitude on the basis of what Jack sincerely asserts is uncontroversial. By asserting (2) as his own reason in reply to a query about what he was doing, Jack's saying (2) also indicates that he believes that his pleasing his uncle was related to what else he thinks he was doing at 3:00, namely, sabotaging the bridge.

If we accept the proposed thesis of deducibility of causal thoughts, we will assume that Jack's asserting (2) as his 'own reason' indicates not only (3) but further that

(4) Jack's belief, stated by (2) and reported by (3), is part of what caused his acting as he did at 3:00.

Here is the central problem we face in accounting for the way Jack's reason causally explains his action: It is not obvious how sentence (4) is logically entailed by sentence (2) or (3) or by the fact that Jack asserts (2) as his own reason for doing what sentence (1) describes him

¹Davidson, Actions and Events, pp. 10, 156.

as doing.

Perhaps the problem lies with the example that I constructed to illustrate the purported deduction.

Generally no one counts a speaker's having a particular belief as sufficient cause for him to do anything except perhaps to assert or assent to a particular sentence. Even then, no one supposes it is obvious which set of sentences a speaker would be disposed to assert or assent to granted he had a particular belief. Having a belief may involve one's being disposed to assent to more than one sentence type and perhaps disposes one to do other things as well depending upon the type of belief that one has. Many suggestions flow from the apparent parallel between attributing propositional attitudes and dispositional states to speakers given what they say on the one hand, and attributing action-causing thoughts to agents given what they do. We will explore and exploit this parallel at different points in what follows.

For the present, this consideration--that expressions of belief are no grounds for anticipating action--might lead one to reject my sentence (2): it doesn't count as a good example of how any reason-giving statement can reveal the cause of something an agent would have done intentionally or would intend to do.

This objection reflects a basic tenet accepted by causalist action theorists, to which we will give central attention in Chapter V. Let me just introduce this basic postulate of a 'holistic' account of reason-giving

explanation here. Consider Davidson's claim that a reason-giving statement "rationalizes" an act done intentionally just in case it describes or indirectly refers both to a causal belief and to some related pro-attitude.

Whenever someone does something for a reason . . . he can be characterized as (a) having some sort of pro-attitude towards actions of a certain kind, and (b) believing (or knowing, perceiving, . . .) that his action is of that kind.¹

Hempel seems to be making a closely related point by the examples he gives when he says "belief attributions and goal attributions are epistemically interdependent."²

This matter of the interdependence of different types of thought constituting a reason counts as one of three basic features of the holistic model for reason-giving explanation proposed as an alternative or amendment to the D-N standard. Before we examine this alternative it is appropriate to see whether the holistic point (about the interdependence of different types of thought functioning causally) can be incorporated into the standard D-N model. For the D-N model has been widely applied to different sorts of explanation by making appropriate modifications.

Without otherwise changing the example as it was

¹. Davidson, Actions and Events, p. 3.

². Hempel, Aspects, p. 45.

initially presented, suppose then that Jack expressed a desire rather than a belief when he told why he sabotaged the bridge. Suppose he said,

(5) I wanted to please my uncle.

Another objection might be raised at this point: one should not be too hasty in assuming that Jack's saying he wants X always correlates with his having a desire for X or with his being causally disposed to do something in order to bring about X. People may express wants or report themselves as having wants for many reasons. But without studying different uses of the terms 'want' and 'desire' one might suppose there is a broad sense in which either term can stand in for a whole gamut of pro-attitudes.¹ One might further suppose that sometimes when agents give their own reasons, the terms 'want' and 'desire' function in this general way.

Thus on the basis of Jack's correctly asserting either (2) or (5) as his own reason for sabotaging the bridge one can assume that

(6) At 3:00 Jack had a desire to please his uncle.

Here is a list of other assumptions one might make if Jack gave (2) or (5) as his reason for the sabotaging:

(7) Jack is a language participating agent;

(8) Jack has an uncle or believes he has an uncle;

(9) Jack believed that by sabotaging the bridge he was pleasing his uncle.

¹. Davidson, Actions and Events, pp. 4, 6.

On Davidson's formulation of the deducibility thesis, if Jack did do something or other intentionally at 3:00¹ or if Jack believes that statement (2) describes what he was doing at 3:00, then (6)-(9) are correct assumptions to make. They are warranted on the basis of the fact that Jack says (2) or (5) when giving his reason for what he did.

Davidson argues further: given assumptions (6)-(9) are true it is correct to make the further assumption that Jack formed an "all-out judgment" at 3:00, based on his "prima facie" beliefs and desires overall.² One may thereby assume Jack was in a particular mental state:

- (10) At 3:00 Jack had the thought 'I shall sabotage the bridge now.'

This formulation of the thought constituting the conclusion of an episode of practical reasoning is adopted from Sellars' reconstruction of Davidson's causalist account of acting for a reason.³ In Chapter V we will compare Sellars' with Davidson's accounts of the structure of a mental state that can cause an action.

I should mention that when I emphasize Davidson's text or another causalist theorist's views, I do not presume to defend my reading as definitive. I do mean to adopt

¹Davidson, Actions and Events, p. 99.

²Ibid., p. 98.

³Wilfrid Sellars, "Actions and Events," NOUS 7 (1973): 179-202, p. 194.

either the most accessible interpretation or one that improves my understanding of the issues at stake, as these are presented in the corner of current literature on action theory with which I am most familiar.

The method just proposed for deducing causal thoughts derives collectively from interpreting Davidson's essay "Actions, Reasons, and Causes" (1963) and "Intending" (1978),¹ as well as from the 1973 article of Wilfrid Sellars noted just above.² In later chapters Davidson's main argument in "Causal Relations" (1967)³ will be adopted to test the consistency of certain assumptions he shares with other causalist action theorists about truth and reference of action sentences. In this way I will sometimes carry an implication further than what is stated overtly in the original texts. Indeed the whole point of Chapter V will be to consider certain implications of recent holistic theories of meaning and interpretation in philosophy of natural language, to see whether these might be stretched to do some work in the field of action theory.

Let us return now to take stock of sentences (1)-(10) comprising our example of Jack explaining what he was doing down by the bridge.

¹. Davidson, Actions and Events, pp. 3-20 and 83-102.

². See p.16, n. 3.

³. Davidson, Actions and Events, pp. 149-162.

Let's grant that the truth of sentence (1) describing Jack's sabotage, conjoined with the fact that Jack gave (2) or (5) as his own reason for what he was doing, does imply at least some of the assumptions about what Jack was thinking, about Jack as an agent with an uncle or with a belief about having an uncle, and so on, as listed in (6)-(10).

Note here that it is not obvious why one is warranted in drawing any causal conclusions from this set of assumptions and known facts. On what grounds does the correctness of assumptions (6)-(10) support the causal conclusion stated in (4), to the effect that Jack's belief and desire about pleasing his uncle caused the sabotage described in (1)?

Here is the nub of the problem: A description of what Jack did--(1)--together with the fact that he gave his reason regarding his uncle's pleasure--(2) or (5)--might warrant drawing assumptions (6)-(10) about Jack's thoughts and status as an agent. But the converse does not hold. Conjoin the fact that Jack gave (2) or (5) as his own reason with the assumptions about his being an agent with a bomb and an uncle and a language as per (6)-(10). The truth of the resulting conjunction of sentences is consistent with (1) being false. Yet (1) describes the explanandum in question. So how is statement (2) or (5), complete with the appropriate auxiliary assumptions, supposed to causally explain the occurrence described in (1)?

This is not a new question in the field of action theory. It has been posed in a more refined way as 'the problem of deviant causal chains' by Roderick Chisholm, Daniel Bennett and others.¹ Chisholm et al address the following challenge as one that specifically plagues explanations about actions done for agents' own reasons. I say their formulations are more refined than the one above because they draw attention to two idiosyncracies of causally explaining events by citing agents' own reasons. The way I have initially presented it, the problem of accounting for reason-giving explanation reflects a more global problem of explanatory uncertainty that may always face singular causal descriptions of nonrepeatable occurrences.² I do not address this more generic concern about verifying any singular causal description.

Here is a way of stating the 'problem of deviant causal chains'. Suppose one is correct in attributing to an agent the onset of certain beliefs and related attitudes that jointly rationalize something he is doing. How can one tell whether it is the occurrence of these very thoughts that caused his behavior on this occasion? How is one to

¹ Roderick Chisholm, "Freedom and Action," Freedom and Determinism, ed. Keith Lehrer (New York: Random House, 1966) 28-44; Daniel Bennett, "Actions, Reason, and Purpose," JP 62 (1965): 85-97.

² Donald Davidson writes that he has "despaired of" the matter as a version of this global problem in "Freedom to Act" (1973), Actions and Events, pp. 79-80.

distinguish an episode where Jack intentionally sabotages the bridge because he has his own reason for doing so, from an episode where Jack inadvertently succeeds in blowing up the bridge that (coincidentally) he happens to have his own reason to sabotage?

Two theoretical problems are packed into this challenge. These will constitute the main focus of concern throughout our study of reason-giving as a distinctive mode of causal explanation.

One theoretical problem is that of reconstructing a line of reasoning from the generality of an agent's desire to the particularity of an event caused by his once having that desire. For it is generally supposed that when a causal mental state like a desire is attributed to an agent, any member of a disjunctive set of events might count as a causal effect of that dispositional state operating in the agent. Further, it is recognized by causalist action theorists that agents are neither exhaustive nor completely effective in carrying out the desires, intentions, needs and other dispositional states attributable to them. There's no telling what may come of Jack's desire to please his uncle, at least not precisely. Yet Jack's reason-giving statement, if it explains anything at all, is supposed to explain an individual occurrence. And the explanation is supposed to be conveyed partly by the fact that Jack was in the dispositional state loosely characterized by Jack as wanting to please his uncle.

A second problem with accounting for the way reason-giving statements explain particular actions is that of spelling out how an onlooker settles upon the right interpretation to give particular effects that are brought about by an agent's acting for his own reason. Which causally related events count as Jack's acting the way he intended to, and which events constitute what he brought about inadvertently or unwittingly? How does one tell whether Jack's attempt succeeded or failed? When should events--that are causally connected to the occurrence of Jack's making some specific attempt--count as incidental to what Jack was intending to do, rather than indicating whether his attempt failed altogether or 'coincidentally' succeeded in some unintended way?

Since agents often do not know about many of the effects of their actions, Jack himself may not know on a given occasion whether his attempt has succeeded or failed, nor how. His attempt may succeed or misfire without his realizing it; that is, he may be wrong about what he is doing in more than one way.

Thus if an agent's thoughts can ever be causally related to his action it is not because he knows all about what he is doing. And the fact that an agent has a particular belief that is causally related to what he is doing does not entail that the content of that belief is factual.

We are supposed to countenance an explanans statement because its truth entails the certainty or the likeli-

hood of another factual report of what has happened or what will happen. However in the case of correct reason-giving statements this assumption is out of place. The problem for analysis here is that an agent's account of what he did in terms of his reason for doing it may reveal the thoughts that caused his action, but not because the agent provides a factual report of what has happened.

Suppose Jack is wrong about what he did: suppose (1) is false. Perhaps the bomb failed to trigger and the bridge's sudden collapse at 3:00 resulted from metal fatigue in the support beams. Or suppose Jack did trigger the bomb explosion but not according to plan. Perhaps he triggered the bomb out of nervousness or because of unconscious anxiety he had about the plan. Perhaps the bomb went off because Jack's elbow was jarred unexpectedly just as he was ready to press down on the bomb plunger.

Jack is warranted in asserting (2) as his reason for intentionally doing what (1) describes provided he believes that sentence (1) describes what he did intentionally and provided he believes that (2) is why he did it.

Yet conditions that warrant Jack's sincerely asserting (2) as his reason are conditions that may not include the occurrence described by (1). Or the occurrence that actually resulted from Jack's carrying out his intention might be misrepresented by the act description in (1). Jack's believing that (2) is related to what he thinks he did is no guarantee that (2) or that (1) describes events

related in the causal sequence that occurred.

And this is why the standard deductive-nomological model fails to illustrate the way causal thoughts can be conveyed by reason-giving statements. If reason-giving statements do causally explain occurrences of action then it appears they must do so in a way that contrasts with the relation between the explanans and explanandum depicted on the standard D-N model.

4. The general adaptability of the deductive-nomological model to many other types of explanation might lead one to suspect that there is something anomalous about the way reason-giving statements describe the occurrences they explain. Agents are characteristically vague and imprecise in making references to particular events when we describe what we are doing intentionally or why. This fact regarding our linguistic conventions might be held accountable for the failure of an agent's reason-giving statement to entail a causal description of what he did. Perhaps there is a characteristic discrepancy between the way act descriptions refer and the way causal descriptions refer to nonrepeatable sequences of events.

What might this discrepancy amount to? Is it tenable to claim that act descriptions always refer in some special way that other sorts of description never can? Does the indeterminacy of act descriptions reveal anything distinctive about the way acts done intentionally occur, marking them off from other sorts of occurrence?

Since different action theorists have proposed different formulations of this distinctive indeterminacy in reference to actions, it's not obvious what the characteristic amounts to exactly; so we will spend Chapter II just aiming at its clarification.

There may be no grounds for expecting any resolution to come from comparing different theories about the way singular act descriptions relate to nonrepeatable event sequences. One may even welcome discrepancies between alternate viewpoints, since the expressions called 'act descriptions' may comprise a very broad and varied class of utterances. Perhaps a lack of consensus about the way actions get described accurately reflects a disparity in the data that needs to be addressed if one's theory of reason-giving is to be complete.

Still it seems important to clarify whether a causalist action theorist is warranted in treating act descriptions as a special class of referring expressions, however broadly or loosely defined. For on certain formulations the claim about unique indeterminacy in reference to actions appears inconsistent with other claims that causalist theorists propose. The most acute examples of this inconsistency involve the claim that causal facts can be deduced from reason-giving descriptions of action. These inconsistencies will be spelled out later in Chapters III and IV.

For the moment consider that some theorists who

claim that particular actions are identical to nonrepeatable individual events also claim that singular act descriptions can never refer to individual events. The apparent tension here is hard to pinpoint or to resolve since the thesis about the unique indeterminacy in reference to actions is open to different interpretations.

Davidson, for example, characterizes a "rationalization" as a singular sentence whose truth is accounted for by a causal relation existing between two distinct events. But this type of causal description is peculiar since a rationalization must redescribe the causal sequence in a way that reflects the agent's relevant thoughts and attitudes.¹ Thus he says reason-giving constitutes "a species of causal explanation . . . set apart from other sorts."² Yet it's not exactly clear what there is about a redescription of an agent's thoughts regarding what he's doing that makes it impossible for the redescription to refer directly to what the agent is doing.

Not all the causalist theorists quoted so far agree about the way to analyze the reference of singular act descriptions, and this can complicate our understanding of what each theorist means when he claims that the reference is indeterminate in some special way.

For instance Hornsby believes that the indeterminacy of act descriptions is rooted in the fact that a particular

¹. Davidson, Actions and Events, pp. 3, 10.

². Ibid., pp. 3, 9.

action can only be described indirectly, in terms of some other event that counts as an effect of that action. She claims there is no *prima facie* reason to assume that act descriptions functioning in "everyday" speech must refer to the very same particular events that neurophysiologists refer to in their causal descriptions.

It is not merely that reductionism is false, and that we should not expect to find any simple neurophysiological predicate with the same extension as any action predicate. Even in the particular case, why should we expect that when we refer to some action what we refer to is something caught by the neurophysiologist's classification?¹

We will examine Hornsby's "two vocabulary" thesis² at length in Chapter IV.

Searle accounts for the indeterminacy of act descriptions much differently. He believes that some occurrences of acting intentionally bear no linguistic representation or descriptive correlate at all. He says there are some events that agents bring about just by knowing how to do so, without being able to describe exactly what it is they are doing. That is, some actions occur for which there are no words available to refer directly to the nonrepeatable events constituting those actions. Yet he argues that a "logical"³ connection must exist between the agent's awareness of what he is 'directly' trying to bring about and what he is in fact 'directly' bringing about: usually some bodily movement. Searle's

¹·Hornsby, Actions, p. 105. ²·Ibid.

³·Searle, "Intentionality," pp. 262-263.

analysis will be the focus of sections in Chapter V.

Both Hornsby and Searle agree with Davidson that every particular occurrence of action is identical to some nonrepeatable event, but not all causalist theorists view actions this way. Goldman for example takes issue with Davidson's construal of the truth conditions for a singular action sentence.

Both Goldman and Davidson recognize that the truth conditions of a singular action sentence do not limit the situation described by that sentence to an individual, nonrepeatable event or sequence. They both recognize an action sentence as failing to refer directly to a particular occurrence in the strict sense that proper names might be said to refer directly to unique particulars.¹ Yet they account for this failure in different ways. Davidson regards action predicates as describing classes of events; Goldman regards action predicates as describing dispositional properties of agents rather than classes of events. In Chapter III we will examine what the dispute may amount to when theorists who believe that particular actions have specifiable causes, deny that an action sentence can designate any event in particular.

Suppose these theorists are correct that an agent's

¹. For an example of Davidson using the phrase 'sentences refer to events' as synonymous with 'sentences about events' see his paraphrase of Kim's theory in "The Individuation of Events" (1969), Actions and Events, pp. 163-180, p. 170 et passim.

reason does somehow identify the particular thought sequence causally related to his action. Then in what ways can the singular description of his action differ from the description of any other causal sequence?

5. One possibility we haven't looked at yet in accounting for the peculiarities of reason-giving explanation is to hypothesize about the special composition or structure of mental states that cause actions.

Each of the four theorists mentioned so far have a different label for the exclusive type of thought or mental state that can account for the truth of a reason-giving statement. Davidson calls this type of thought an "all-out judgment." Goldman calls it an "occurrent want." Hornsby calls it an episode of "trying" and Searle labels it an "intention in action."¹ Thoughts represented by these four labels seem to have the following three characteristics in common.

One salient feature is the 'interdependence' among the mental attitudes constituting a primary reason, as cited by Hempel and by Davidson. Other action theorists seem to agree with this 'holistic' claim: One can identify the propositional attitudes constituting an agent's primary

¹. Davidson, Actions and Events, p. 98; Goldman, Theory of Human Action, p. 86; Hornsby, Actions, p. 33; Searle, "Intentionality," p. 258.

reason only in batches or in pairs (like a belief conjoined with a desire).

For instance by attributing to me the belief that there is bread in the store across the street one cannot explain why I crossed the street and entered the store unless one also assumes that I wanted or felt obliged to buy bread. Likewise, by attributing to me the desire for bread one cannot account for my leaving the house with money unless one assumes that I believe I can satisfy my desire for bread by leaving the house with money.

A second generally recognized feature of causal thought clusters has to do with identifying events as a particular kind of act done intentionally because an agent had a particular cluster of thoughts.

There is supposed to be an interdependence between the content of an agent's thoughts about what he is doing on the one hand, and the correct way to interpret a causal sequence as the kind of act he is doing intentionally.

Thus it is only by assuming I am having some particular combination of thoughts like this belief and desire about bread, that one can correctly interpret my exiting the house as an act of going to the store intentionally to buy bread, rather than interpreting the sequence as some other type of action like going to change a large bill, or failing to keep a fast, or mistaking the grocery store for a drug store.

A third widely accepted premise on the 'holistic'

view is that people must know what they are talking about and what they are bringing about in order for their speech and their actions to be interpretable in the right way.

Although there is controversy about whether this knowledge is propositional, all the causalist theorists quoted here share this assumption that a speaking agent always knows something about what he is doing. It is readily shown why this assumption is so attractive but it is less obvious why the assumption is warranted.

It does seem obvious that the most straightforward way to find out what Jack was doing intentionally at 3:00 and why is simply to ask him. Certainly with regard to many of the actions he does for his own reason it is supposed that Jack can be correct or at any rate corrected about which of his attitudes and thoughts caused his action and about what happened as a result.

But on consideration the claim that an agent always knows something about what he is doing does not seem consistent with other widely held beliefs about agents acting for their own reasons. For instance, suppose that Jack could be depended upon to know precisely what he was doing and why under some set of conditions. Then indeed an entailment relation could be expected to hold between Jack's correct primary reason-giving statement about his action and a true sentence describing a causal sequence of events that occurred under those conditions.

Yet if it were solely Jack's knowledge that provided

the basis for deducing the correct causal description of his action on this occasion, then it's hard to see what grounds there could be for suspecting Jack of any misjudgement or error about what he's done on any other occasion.

It seems that our standards for adequacy in accounting for the relation between an agent's reason and his action are directed by a tension between two competing sets of intuitions.

On the one hand we apparently presume agents can know something about what they are doing since we rely on what agents say about their actions in order to surmise what is going to happen in the near future.

Yet on the other hand it is uncontroversial to assume Jack will frequently fall short of acting in a predictably rational way and that he will frequently fail to be doing what he thinks he is doing. Indeed it is a classic difficulty to define precisely how Jack would act if he were ideally or completely rational. We recognize agents as acting on the basis of erroneous beliefs, oblivious to the consequences of what they do and for reasons that may appear to them in retrospect to be contrary to their own best judgment,

In light of the fact that agents are recognized as being fallible in different ways, it is not obvious what the consistent role is supposed to be for an agent's beliefs in distinct episodes of his practical reasoning.

More generally, it is not obvious how much generalizing is warranted about the types of thought that must occur whenever something happens because an agent has acted for his own reason.

The 'holistic' assumptions outlined here parody the "interdependence of belief and meaning"¹ proposed on Davidson's theory of radical interpretation. We need to examine whether it is plausible to apply the results of his theory about the preconditions for interpreting sound patterns as sentences to illuminate the relation between events called actions and attitudinal states called reasons. What conclusions can be drawn from the apparent parallel between interpreting an utterance as someone's asserting his own belief and interpreting an event as someone's acting for his own reason?

A holistic model highlights the interdependence between identifying an agent as being in a specific causal state and interpreting an event sequence as a certain type of action that occurred because the agent was in that particular mental state. Hence a holistic model is supposed to be a more sensitive instrument than the traditional D-N model for observing how reasons cause actions.

So we need to test the holistic model to see if it offers a solution to the problems that arose when we tried to use the D-N model to demonstrate the relation between an

¹. Donald Davidson, "Radical Translation," Inquiries into Truth and Interpretation (Oxford: Clarendon Press) 1984, pp. 125-139, p.134 (hereafter cited as Davidson, RT).

action and the agent's own reason for acting.

Recall the reasons for rejecting the D-N model: It left open the possibility of misinterpreting a causal sequence as the wrong sort of action from the standpoint of the agent's thoughts about what he meant to be doing. For instance the model might yield false results by failing to distinguish between something being done because an agent has acted for his own reason, and something being brought about by that agent as his "inadvertent success", to quote Chisholm's formulation of this problem.¹

Can we avoid the problem of misconstruing 'deviant' causal chains by adopting a holistic model of reason-giving explanation?

What else should we look for to test whether a holistic account offers a viable alternative to (or a substantive modification of) the standard D-N model of explanation?

Recall again that reason-giving statements differ from other kinds of causal description also in the way that reason-giving statements exhibit explanatory adequacy. An agent's own reason explains an individual, nonrepeatable episode of action. The explanatory adequacy of this type of statement is exhibited in a case by case fashion. The auxiliary role that law-like statements may play in the verification of individual agents' reason-giving statements is a controversial question that I do not address here at all.

¹Chisholm, "Freedom and Action," p. 37.

Again the parallel emerges between interpreting individual actions done for individuals' own reasons, and interpreting "the infinity of sentences [a] speaker might utter."¹

We will consider whether an adequate model for reason-giving explanation is likely to show that certain definitive types of thought must occur whenever an agent acts for his own reason. I do not have the resources to pursue an answer to this question. My aim here can only be to show why it may be worth reconsidering the sort of non-linguistic facts about reasons one should expect to follow from linguistic facts about reason-giving statements.

Conclusion to Chapter I

We have been examining how a standard, widely applicable model of explanation fails to provide an understanding of the way an agent's own reason-giving explanation might causally explain his action. To account for this failure, some causalist theorists stress the element of language participation as a factor overlooked by the D-N model. These theorists emphasize the following. To understand what it is to act intentionally for one's own reason, one must understand what it is to produce sounds that have linguistic meaning. The parallel suggested was to compare the interpretation of events as actions done intentionally with the interpretation of utterances as sentences.

¹. Davidson, Truth and Interpretation, p. 127.

The significance of this parallel will be the focus of Chapter V.

Before that we will canvass different ways of characterizing act descriptions as a distinct class of linguistic expression. Next in Chapter II we will try to clarify a sense in which 'indeterminate reference' is idiosyncratic of act descriptions, a characteristic that contrasts them with causal descriptions of event sequences. In Chapter III we will study whether anything distinctive about the occurrence of actions follows from the way that Davidson construes the logical form of singular action sentences. In Chapter IV we will examine the claim that causal facts are entailed by the presence of key terms and grammatical structures in true action sentences. We will try to refine the thesis that two vocabularies exist for describing occurrences called actions done intentionally. We will examine the way that causal assumptions can be warranted on non-truth functional grounds, by studying examples of the way causal facts are surmised about occurrences generally, not just about occurrences of acts done intentionally. We will study whether it is coherent to claim that 'action' sentences can't be substituted in truth functional contexts for any other sentence describing a causal sequence of events.

The point of studying truth functional and non-truth functional grounds for drawing causal 'inferences' is to locate what it is that makes reason-giving statements

seem like a special sort of causal explanation. After these attempts to capture the distinctive features of act descriptions as a class of expressions, we will turn to study the specific function of act description in agents' own reason-giving statements. Can anything be learned about reasons and the actions they explain from theories about the relation between linguistic utterances and the thoughts they 'express'?

CHAPTER II

INDETERMINACY IN REFERENCE TO PARTICULAR ACTIONS

Introduction

Clearly, on any occasion of their use the nominals 'Jack's arm raising', 'Jack's making a legal turn' and 'Jack's bridge sabotage' can each refer to one of a whole range of measurable positions that Jack's arm might assume. Does it follow that the references of 'arm raising', 'making a legal turn' and 'bridge sabotaging' are less able to be determined than are the references of the nominals 'arm rising', 'bridge collapsing' and 'Jack's tie'? This question is the subject of this chapter.

We will be focussing on certain terms and phrases that can describe actions done intentionally. In order to do this let it be assumed that we already have a way to resolve the ambiguity in reference that may seem to result from Davidson's way of construing singular sentences containing act-descriptive terms and phrases. In the next chapter we will address Davidson's analysis of all singular sentences about actions as cases of existential generalization. In this chapter we will treat it as given that the following singular action sentence

is about one and only one occurrence. The occurrence that this sentence is about happens at a fixed time and location and it verifies or accounts for the truth of the sentence for one occasion of its use. Here is the sentence:

Jack's intentionally making a legal turn by raising his arm took place at 3:00.

A more familiar paraphrase, though less perspicuous, might be the sentence:

Jack intentionally made a legal turn by raising his arm at 3:00.

The method of construing reference for both these sentences recommended just above is one that Sellars' analysis suggests.¹

If we accept Sellars' recommendations then we are free to examine the reference of component expressions e.g. 'arm raising' and 'making a legal turn' rather than attending to the reference of the sentence as a whole.

Of course one cannot take for granted that Sellars and others are correct to assume that there is one and only one event being described here rather than two, or none, or indefinitely many. Some of the stands taken in this controversy will be studied in the next chapter when we focus on the individuation of actions.

Before considering singular action sentences as a particular class of act descriptions it may prove useful to raise the broader question regarding act descriptions

¹. Sellars, "Actions and Events", p. 197.

generally: What motivates causalist action theorists to treat descriptions of acting intentionally on particular occasions as a distinct class of expressions in the first place?

Since Melden wrote Free Action (1961)¹ there has grown a tradition in action theory of stressing the significance of the "everyday vocabulary"² constituting statements about acting intentionally. Of course many subjects of conversation in everyday life reflect a speaker's provincial indifference towards specifying particulars that he is talking about. Indeterminacy of reference does not characterize act descriptions exclusively, if the indeterminacy in question boils down to a general lack of specificity about which limb movements or other occurrences count as evidence that a certain type of action has taken place. Note for instance the different sorts of failure in specificity occurring in the following statements:

Put it over there.

It may rain.

It hardly rained at all.

There was a fire on Elm Street yesterday.

I saw a table covered completely with a cloth.

Despite the vagueness that pervades conversation on most topics, some action theorists argue that certain linguistic conventions characterize a special sort of

¹. (London: Routledge, 1961).

². Hornsby, p. 105.

vagueness peculiar to statements about acting intentionally. We will study two arguments that derive this uniqueness in reference to actions from this assumption: actions are causes of specifiable bodily movements and of other observable effects occurring beyond an agent's body surface.

The two arguments will be the focus of sections two and three. In section four we will consider an inconsistency that seems to emerge from claiming both that actions can only be described indirectly in terms of their effects, and also claiming that reference to actions must be indeterminate in a way that reference to the effect of an action never is. Lastly, in section five we will observe two different strategies developed by the causalist action theorists Davidson and Searle, for determining the 'content' of a thought that caused a particular action. On both their methods the alleged indeterminacy in reference to actions is overlooked, in order to interpret a particular occurrence in the right way as the type of action done intentionally for a reason. Later in Chapter V we will evaluate these two strategies for discerning the causal thought counting as the reason an act was done intentionally.

1. Neither Davidson, Goldman, Hornsby nor Searle denies that agents are able to recognize and to correctly describe what other agents are doing intentionally.

As an onlooker one can know by 'direct observation'

that Jack's arm is rising and from this and other facts one can assume correctly on many occasions whether Jack is raising his arm intentionally or not.

But on other occasions, one can know that Jack's arm is rising yet be uncertain whether he raised it intentionally. Knowing whether Jack's arm is rising does not entail knowing whether Jack is raising his arm intentionally.

Perhaps someone besides Jack can know what Jack is doing intentionally in the way that someone else knows when Jack is angry.¹ Sometimes an agent other than Jack knows that Jack is angry because it is obvious: Jane for instance can usually tell when Jack is angry just by looking at him. But observing any particular body surfaces of Jack does not wholly comprise her seeing that Jack is angry. And even Jane might be uncertain whether to count the tightening of Jack's jaw and fist on a particular occasion as evidence that he is angry or angry enough to hit her.

It is not obvious what there might be about Jack's jaw or fist that Jane could know that would indicate whether or not Jack tightened it because he was angry, nor what would indicate to her the degree or kind of anger that did cause such tightening.

Nor is it obvious what there is about the whole set of circumstances that can indicate the sort of action

¹Paul Ziff, "About Behaviorism," Analysis (1958): 132-136.

intended when it is clear to an onlooker that Jack is raising his arm intentionally.

Consider Jack's arm rising on a particular occasion. What about it can indicate the type of act being performed? Is this occurrence the effect of Jack's trying successfully to raise his own arm or is it the effect of Jack's failed attempt to lower his arm? Or is the arm rising the effect of Jack's wanting to do something else he believes (accurately or not) that he can accomplish by raising or lowering his arm 'just so'?

Under the circumstances suppose Jack appears to an onlooker to be making a legal left turn. How does one tell whether he is intentionally making a left turn or whether the signalling on this occasion is an inadvertent byproduct of Jack's pointing to something as his car approaches an intersection? How does one tell whether Jack's trying or doing anything in particular is the cause of his arm's rising at that very moment? Perhaps the arm rose because Jack took too much Thorazine, so that his arm jerked up in a spasm at the very moment he wanted to be making a left turn.

Ordinarily such questions can be resolved just by asking the agent what he thinks he's doing. For example one may ask Jack whether he intended his arm to rise in just the way it did on this occasion and why he intended it. Jack might reply:

- (1) I wanted to make a legal left turn.
- (2) I believed that by raising my arm
I would be signalling a left turn.
- (3) I was checking to see if the drug effect
had worn off by trying to raise my arm.
- (4) I was pointing.

Yet even if Jack provides sufficient evidence in virtue of his reply to decipher what he meant to be doing, his reply cannot preclude the problem that something 'deviant' about the causal sequence in question might have occurred without his knowing it. The resultant moves of his body might be mimicking what he meant to have done or the way he meant to have done it. Granted, most of the time agents seem to know what they're about and where their limbs are, as well as they seem to know anything about the details of what is going on in the course of everyday activities. Why is such stress laid upon cases where it is uncertain what an agent knows about his own action?

Introducing the possibility of deviant causal chains of events occurring unbeknownst to an agent can be misleading. It may seem to suggest that it is obvious what there is that distinguishes 'deviant' sequences from the 'nondeviant' ones. But this is precisely one of the points where the relation between act descriptions and causal sequences seems to need clarification. By formulating illustrations of the disparity between what is happening and what the agent thinks is happening, we can press the question of what it is exactly that an agent is right about

when he is right about his attempt succeeding or failing in a particular way.

In Chapter I when we were testing the efficacy of the D-N model of explanation, we noted that the correct interpretation of an event or sequence as a type of act done intentionally does not follow from facts about the case on truth functional grounds. That is, given a set of statements describing what is going on together with what an agent thinks is going on or what he wants to go on, one cannot infer by truth functional logic any particular sequence of events. Given what Jack thinks or says he was doing intentionally, he cannot deduce the specific fixed bodily movements that occurred. Knowing what Jack thinks he is doing intentionally, or knowing the belief and desires that caused him to try some particular thing intentionally does not entail knowing which particular bodily movements and other events to count as the causal effect of his thoughts. This inability to deduce particular events obtains even when Jack is correct about what he is doing and even when he is successful in carrying out precisely what he means or wants to be doing.

In the current example where Jack is driving, consider the many different sets of circumstances where any one of Jack's replies (1)-(4) describes what he was doing intentionally though his arm didn't rise at all. Suppose Jack is indeed making a legal left turn and is doing so intentionally: (1) is true if he honks his horn,

or if he yells 'Left turn!' or if he blinks a light by pressing a pedal with his foot. Likewise, any one of (1)-(4) may be true if Jack's arm is suspended at a 90° angle from the trunk of his body or at a 98° or a 50° angle from his body. Jack may know whether it is his hand rather than his foot that he moves when he tries to make the legal turn, without being aware exactly of what it is he does with his hand and not his foot at the time. To show this, one might ask Jack to replicate the bodily movement that constituted his legal left turn immediately after the occasion in question. Or one might ask Jack to identify from an array of snapshots the one that illustrates the precise position of his arm while he intentionally moved it to make his turn. It doesn't seem generalizable whether Jack will remember where his limbs were; this may or may not indicate that he ever knew exactly where they were at the time he made the turn.

Notice that the elusiveness in reference to particulars is not restricted here to the contents of thoughts or awareness an agent has about the limb movements involved in what he's doing. Surely this factor varies widely for most individuals on different occasions. Nor is the variety peculiar to knowledge of one's own actions. Suppose one were intimate enough with Jack to predict just what he will say when offered sugar in his coffee. One can be accurate about the type of expression Jack will evoke on a

certain nonrepeatable occasion without having a clue as to the exact motions of his tongue, larynx and vocal chords that comprise his verbal behavior on that very occasion. And different onlookers may vary in the degree to which they are astute observers. Consider a golf caddy who impersonates his boss. He may be able to replicate exactly the angle and every movement comprising his boss's stance and swing. But pantomime is wordless. The linguistic expressions called act descriptions do not seem to share this splendid precision in representing particulars.

These diverse considerations have led some people to draw the following generalization: Descriptions of acting intentionally involve an indeterminacy in reference to the particular limb movements and other occurrences located at measurable distances beyond the surface of an agent's body.

At this point it is tempting to seek some way of condensing this conclusion into a concise formula, applicable to all act descriptions or at least to a specifiable class. But our latest considerations suggest that it may not be possible to formulate a univocal pattern for the mode of reference to particulars involved in act description. For it is still unclear what sort of particular it is that an act description doesn't refer to directly or precisely.

So far we have noted that success or failure of an

act done intentionally is a function of two or three variables: the type of action performed and its success depends in part on what an agent thinks is happening and what he wants or approves of happening and what in fact is happening. Consequently one would expect that the truth conditions for a sentence reporting an action done intentionally should reflect the interdependence of these variables.

Quine has proposed a strategy for deciphering "opacity" of reference in different verb constructions, to contrast with sentences wherein reference to particulars is "transparent."¹ At the end of Chapter Four in Word and Object (1960), Quine suggests that transparency or opacity of reference is determined for a sentence according to the diverse conditions determining the truth value of that sentence on each occasion of its use. Quine's analysis involves reference to particulars constituting the direct object of an action named in certain verb constructions. We do not have the space here to test how Quine's paraphrases might be applied to illustrate the indeterminacy involved in reference to actions. The nonrepeatable particulars constituting the whole of an act done successfully or one that failed include the bodily movements the agent may or may not be aware of making, as well as further effects that determine the success or failure of his action.

¹. Willard Van Orman Quine, Word and Object (Cambridge: MIT Press, 1960), pp. 151-156.

Given agents' susceptibility to failure and mistake, perhaps contrary referential constructions are warranted when describing certain episodes as things done intentionally. To simplify the point consider a case where an agent is both successful and knowledgeable with respect to what he's actually bringing about. For example suppose it is patently clear to Jack and to others that what Jack is doing is sabotaging a certain bridge. Even then it's not obvious that a correct report of what he's doing intentionally must fall exclusively under one mode of reference to nonrepeatable particulars and not another. It is true that Jack intentionally exploded that particular bridge over there and because he did so it is correct to say that he succeeded at his sabotage. For had it been any other bridge that he exploded his sabotage attempt would have failed. Yet it is also true to say that his action failed in that he did not undermine the communist terrorist plot as he intended. This latter failure occurred at the very time and place that his successful sabotage took place. Yet his failure to undermine the communist plot is a broader campaign as Jack sees it, whose ultimate success or failure cannot be determined by referring to the outcome of any particular bridge caper in isolation. It might be correct to say that Jack's endeavor to undermine the plot involves no particulars beyond himself if there is no communist plot to undermine. Yet for Jack his progress towards the eventual success of undermining the plot can be

graphed in accord with his cumulative successes at different sabotage capers. If there were a plot to undermine, his gradual success would be measured in the same way.

Is it altogether correct to assimilate Jack's plot-undermining here to the paradigm case suggested by Quine in the example of lion-hunting?¹ Recall that Jack's lion-hunting can proceed steadily without any relation to specific lions in a locale where no lions roam. But not even Jack would consider it progress towards the eventual undermining of the terrorist plot if none of the bridges he had targetted were ever vanquished according to his plan.

It is not clear what the nature of the disparity is between lion-hunting where there are no lions and totaling up capers to undermine a nonexistent plot. Considerable testing of different examples seems in order before even tentative conjectures can be formulated about the role of nonrepeatable particulars in determining the success or the failure of an act done intentionally for an agent's own reason. We must leave unexamined the many suggestions that follow from Quine's models contrasting "opaque" reference with "straightforward relative" terms² functioning as main verbs in certain sentences. In Chapter V we will look further at the role of the agent's thoughts about

¹. Quine, Word and Object, pp. 154-156.

². Ibid., p. 155.

what he is doing at the time he acts. We will examine the interdependence of what the agent thinks and what in fact happens, in determining whether a sequence is to count as a specific type of successful action or as a failure of some other type.

2. One might argue that acts done intentionally, along with their descriptions, are unusual because only the effects of acting intentionally can be observed by onlookers. Jack himself may know what he intends to do or expects to make happen. But perhaps his doing it or his making it happen can never be observed independently of at least one occurrence caused by one incidence of his doing or his making happen. Thus his action is a distinctive type of occurrence because its descriptions are always indirect, referring to observable effects of what he does.¹

Perhaps a more suggestive verb to use here instead of 'observing' or 'seeing' actions is 'witnessing' actions. Hornsby claims for instance that "what we see is the fact that someone is acting and not that we see his action".²

Why can't we see actions done intentionally in themselves?

Hornsby gives two related reasons in answer to this question. One reason is that actions are identical to events that occur "inside the body".³ She says that every

¹·Hornsby, Actions, p. 106. ²·Ibid., p. 102.

³·Ibid., p. 103.

act done intentionally "can be identified with"¹ a bodily movement caused by some episode of "trying" that occurs somewhere unspecifiable inside the agent's body. The second reason is that whenever one observes an occurrence or monitors a change beyond or beneath the surface of the agent's body, one is observing something brought about by the agent's acting. It is wrong to confuse the cause with the effect: if bodily movements are caused by an agent's acting then they are not to be confused with the acting itself. Suppose it is true that acting intentionally is always described indirectly in terms of some observable effect.² Does it follow that the proper referent of any act description must be indeterminate?

To illustrate the contrast between indirect and 'indeterminate' reference ascribed to act descriptions and the reference to bodily movements that is standardly assumed to be 'direct', consider these two sentences. Suppose their truth is evidenced on one occasion by the same nonrepeatable limb movement:

(P) John raised his arm intentionally at 3:00.

(Q) John's arm rose at 3:00.

Hornsby supposes that sentence P is about something Jack did and therefore whatever P is about must be distinct from whatever Q is about. For Q is supposed to describe

¹.Hornsby, Actions, p. 10.

².Ibid., p. 106.

(on the date in question) a causal effect of whatever it is Jack did intentionally, as described by P. Suppose the occurrence that Q is about can be determined in the way that the situation any sentence is about generally gets determined. Suppose for instance we assimilate Q to cases illustrated by the sentence 'The cat is on the mat', whereby the sentence 'The cat is on the mat' gets construed clearly as being about the cat.¹ Suppose further that Hornsby is right in claiming that whatever P is about at 3:00 caused whatever Q is about, and that for this reason the occurrence described by P cannot be the same as the occurrence described by Q.²

Does it follow that whatever occurrence P is about must be indeterminate?

To answer this question, consider a parallel case that does not involve action and its effects, but does involve descriptions of a cause and its effects. Suppose that under certain conditions the only evidence an observer can have to establish the occurrence of a particular nonrepeatable flow of electrons is the flash of lightning that he sees between two cloud formations in the sky.

The observer might be warranted in stating these

¹.The reader may recall here a literature on the subject of deciding reference for sentences like 'The cat is on the mat', but this literature does not address the question we are concerned with here.

².Hornsby, Actions, p. 13.

two sentences in virtue of the very same occurrence of lightning striking at 3:00:

(R) There was a flow of electrons at 3:00;

(S) There was a flash of lightning at 3:00.

There is no inducement here to conflate or to substitute the observer's report of the flash that he saw with his report of the electron flow that caused it, even though both reports are confirmed by the same occurrence of lightning. On the other hand, there seems to be nothing counterintuitive in treating both of these sentences as being about the very same event or occurrence, the one that verifies or accounts for the truth of each.

It is uncontroversial that one sentence describes the occurrence of the lightning strike in terms of its cause (the electrical discharge) and the other in terms of one of its effects (the flash). Cause and effect can be recognized as distinct states of a single continuant or system; cause and effect can also be recognized as distinct aspects or features or properties of a single occurrence.

This parallel may not suit; for one may not regard actions in relation to their effects as comparable to the relation between an electron flow and its perceptible results in lightning, thunder, or traces in cloud chambers. Hornsby does not accept this parallel, probably. In Chapter IV we will look closely at the specific considerations motivating Hornsby to deny that the proper reference of an

act description can ever be counted as any observable change or movement whose location is measurable relative to the surface of an agent's body. She claims, as do other action theorists, that one must recognize the difference between two sorts of discourse. One form of discourse contains sentences like P characterized by key terms and grammatical structures; the other mode contains sentences like Q with a different set of conventions, "patterns of entailment,"¹ assumption, and expectation associated with these sentences by language participants. Part of learning a natural language is supposed to involve learning to recognize the circumstances in virtue of which certain grammatical patterns and key phrases indicate that agents are expressing their intentions to do certain things. Sellars has proposed that a child learns sound sequences containing the term 'shall' in a manner that causes him to move his arm, for instance when he thinks 'I shall move my arm now.'²

Here we are asking whether the indeterminacy attributed exclusively to act descriptions can be accounted for solely by the way a certain class of expressions conventionally occurs. Later in Chapter IV we will consider other postulates of the "two vocabularies" thesis of Hornsby that was just mentioned.³

¹.Hornsby, Actions, p. 124.

².Sellars, "Actions and Events," p. 192.

³.Hornsby, Actions, p. 105.

3. So far we have considered a parallel that suggests the reference to a cause in terms of its effects need not involve any sort of distinctive indeterminacy on that score alone. In this section we will consider why it seems inconsistent to say both (1) that act descriptions are indeterminate in the nonrepeatable particulars they are about, and (2) that the way to identify what type of action has been done intentionally is always in terms of its observable effects or its anticipated effects for a particular occasion.

We will elaborate Hornsby's view of actions in fuller detail in Chapter IV. For now, suppose she is right to claim that stating an action sentence is a covert way of making a causal claim or of conveying a causal fact. That is, suppose Hornsby is right that by saying 'Jack intentionally raised his arm at 3:00' one is conveying the causal fact that 'Jack caused his arm to rise at 3:00.' Then on what grounds is it plausible to regard one's action statement as indeterminate with respect to every causal sequence of events that occurred at 3:00?

Let us spell out the apparent inconsistency by considering a set of conditions warranting belief in these five statements:

- (A) Jack intentionally raised his arm at 3:00;
- (B) Jack tried to raise his arm at 3:00;
- (C) Jack's arm rose at 3:00;
- (D) Jack tried to raise his arm at 3:00 and was successful, (that is, he tried successfully).

(E) Jack caused his arm to rise at 3:00.

Hornsby's theory of acts done intentionally is that they are identical to successful tryings.¹ That is, A is true just in case D is true. And D is true just in case B, C and E are true.

How do we know sentences B through E are true, and thereby also know that the action sentence A is true? For reasons detailed in Chapter IV, Hornsby argues that we can't observe agents' tryings, that is we can't observe whatever it is that D is about and learn from this observation that D is true. Instead what we can do is witness reports of agents trying to do things and then observe the evidence of their successes or failures. Thus by observing 'directly' an occurrence that warrants believing C (seeing Jack's arm go up) and by hearing Jack say something that warrants belief in B, we are also warranted in believing D.

Now we should go slowly here. To interpret the evidence for B and for C as evidence that also confirms the truth of D, is to interpret that evidence as an indication that a particular nonrepeatable causal sequence has occurred: We hear Jack say he's tried or is trying to raise his arm and then we see his arm go up. We assume that whatever internal occurrence or state that Jack reports

¹ Hornsby, Actions, pp. 10, 14, 37-39, 45.

as his trying to raise his arm is something that caused the arm's going up. Elliptically this causal assumption might be stated as E.¹ If we believe E, then we call Jack's trying 'successful' or we say that he tried successfully, i.e. we also believe D. Suppose that calling an agent's attempt or trying 'successful' is a covert way of asserting that a particular nonrepeatable causal sequence of events occurred. That is, what is conveyed by saying that Jack tried successfully at 3:00 is this: on the particular occasion in question a specifiãble causal sequence took place. This is a reasonable reading of D if we are supposing that D is accepted on the basis of **evidence for E** being true. It is warranted to accept D on the basis of evidence for E if we believe that the events D is about are identical to the events that E is about.

Now, if Hornsby is right that acts done intentionally are identical to successful tryings, then the events D and E describe are identical to the events that A is about. But that seems to be just a paraphrase of saying that A is about the same nonrepeatable causal sequence that D and E are about.

¹.According to Hornsby, believing E need not involve subscribing to a special mode of agent causality. E can be construed as an elliptical way of saying that something Jack did (his trying) caused the effect described; thus E is assimilated to an 'ordinary event causality' model. See Hornsby, Actions, p. 15, n. 8 and chapters five and seven. The pros and cons in the literature regarding a special 'agent causality' model is a subject I could not pursue in the scope of this essay.

Yet Hornsby denies the action sentence A is about a causal sequence: action sentences are about tryings and their reference is always indeterminate with respect to particular events in causal sequences.¹ This is hard to see. If one can correctly infer from the meanings and the truth values of sentences A through E that the situation A is about is identical to the sequence that accounts for the truth of D, then how can one regard D as being about a particular causal sequence and also regard A as not being about any causal sequence?

These considerations do not show that there can't be any cogent way to regard the act description in A as indeterminate with respect to particular events. Hornsby claims A and D are covert ways of stating causal facts. Maybe statements of causal facts about actions have nothing to do with particular events.

Here is an argument that is useful in making more vivid the inconsistency I perceive in the position Hornsby is taking about action sentences and causal sequences of events. The argument appears in Davidson's "Causal Relations" (1967), where he used this so-called Fregean approach to show that the connected terms of a singular causal statement must be singular referring expressions designating particular nonrepeatable events, and not sentential expressions.

¹ Hornsby, Actions, p. 105 and the whole of Chapter Eight.

Consider the singular causal statement

- (1) The short circuit caused the fire.¹

Suppose one tries to construe the correct paraphrase of (1):

- (2) The fact that there was a short circuit caused it to be
the case that there was a fire.²

This paraphrase might suggest that the causal connective in a singular causal description really connects sentences to one another. Davidson reminds us that if the causal connective is sentential then it surely cannot be a truth functional connective defined by standard rules of inference ". . . since (2) may change from true to false if the contained sentences are switched."³

This much is uncontroversial in the debate over how to interpret the causal connective. What Davidson then goes on to show is an erroneous result that follows from interpreting the component phrases in (2) as sentential expressions with reference to no events in particular:

First Davidson notes that "substitution of singular terms in sentences like (1) and (2) does not touch their truth-value."⁴

¹. Davidson, Actions and Events, p. 151. ². Ibid.

³. Ibid., p. 152.

⁴. Ibid., "If Smith's death was caused by the fall from the ladder and Smith was the first man to land on the moon, then the fall from the ladder was the cause of the death of the first man to land on the moon." pp. 152-153.

Surely . . . we cannot change the truth value of the likes of (2) by substituting logically equivalent sentences for sentences in it. Thus (2) retains its truth if for 'there was a fire' we substitute the logically equivalent ' $\hat{x}(x=x \ \& \ \text{there was a fire}) = \hat{x}(x=x)$ '; retains it still if for the left side of this identity we write the coextensive singular term ' $\hat{x}(x=x \ \& \ \text{Nero fiddled})$ '; and still retains it if we replace ' $\hat{x}(x=x \ \& \ \text{Nero fiddled}) = \hat{x}(x=x)$ ' by the logically equivalent 'Nero fiddled.' Since the only aspect of 'there was a fire' and 'Nero fiddled' that matters to this chain of reasoning is the fact of their material equivalence, it appears that our assumed principles have led to the conclusion that the main connective of (2) is, contrary to what we supposed, truth-functional.¹

Davidson's preferred analysis of the connected phrases in a singular causal statement does not go uncontested. The subject cannot be pursued with any adequacy here although we do touch upon it again in Chapter III. Nonetheless Davidson's analysis recommends that we scrutinize reason-giving statements in the following vein.

We need to learn what it can mean to claim both (1) that sentence E describing a causal sequence is entailed by sentence A describing an act done intentionally and (2) that sentences A and E cannot be about identical events since A is about no events in particular. We will try to refine this version of the deducibility thesis in Chapter IV. For the present we may conclude this much from our considerations in this section:

Suppose it is possible to establish that an action must be referred to (or described) indirectly in terms of its

¹ Davidson, Actions and Events, p. 153.

causes and its effects. Suppose the causes and effects of an action can be referred to as nonrepeatable particulars. And suppose it is correct that a true action sentence does entail other sentences describing the particular causes and effects of a nonrepeatable act done intentionally. Then it cannot also be the case that the way a sentence describes the occurrence of an action must be indeterminate with respect to particular causal sequences. For if this last claim about the mode of reference to actions being indeterminate were true, then the former claims about determining the causes and effects of a particular action would remain undecidable.

4. Let's continue to seek the sense in which one might contrast the reference of the sentence

(A) Jack intentionally raised his arm at 3:00;

with the reference of the sentence

(C) Jack's arm rose at 3:00.

When I use the locution 'reference of a sentence' here I have in mind whatever that sentence is about.

As noted Hornsby observes the passive, or as she calls it the "intransitive"¹ form of the verb in C fails to "support inferences"² that are suggested to language participants by the "transitive"³ verb in sentence A.

¹.Hornsby, Actions, p. 2. ².Ibid.

³.Ibid.

She concludes that sentence C must therefore fail to report what the action sentence A succeeds in reporting even when the truth of C is evident by virtue of the very same occurrence that makes evident the truth of A.

This argument rests on the presupposition that two sentences that do not logically entail each other cannot describe identical things or events. By contraposition the principle implies that if two sentences do describe the same thing or event then they logically entail each other.

But generally speaking the fact that two sentences are true and are known to be true by virtue of the same nonrepeatable event or particular object, is not grounds for expecting that one of the sentences logically entails the other.

Consider the following example showing that the principle appealed to here is false, by modus tollens. The example draws again upon a parallel between sentences describing actions and sentences describing properties of entities that are not occurrences of action.

Certain evidence warrants my believing that

(X) The Morning Star is green;

I also know that

(Y) The Evening Star is green,

and that

(Z) 'The Morning Star' names the same object in the sky as does 'The Evening Star'.

But I cannot deduce any one of these statements from a conjunction of the other two, even though I may be warranted in asserting all three sentences under the same conditions. No two of these sentences entail the third because the set of truth conditions warranting one's belief in each sentence is encountered independently of the conditions that warrant one's belief in the other sentences. For example, I believe X because of evidence accumulated at a different time of day than the evidence available to warrant my believing that Y. The set of conditions that account for my correctly believing one of these sentences does not include--and may preclude--the conditions in virtue of which I am correct in believing the other two.

Like X, Y and Z, one might be warranted under certain conditions to assert both A and C of Jack. It needs to be explored whether anything about the 'logical' relation between A and C follows from the fact that they both might be warrantably asserted on the same occasion. This distinction between the truth conditions of sentences and their conditions of assertability is an important consideration for clarifying other contrasts between the classes of sentences patterned after A and C respectively. We will pursue the significance of truth vs. assertability of an action sentence later in Chapter IV.

The parallel just drawn between observing a planet

in the sky and witnessing an action done intentionally might be rejected by one who takes the description of actions to be special in some way. It remains an uncontroversial piece of the data that we appear to talk and think in a characteristic way about actions. So it behooves us to reconsider the intuitions motivating one to view sentences A and C as describing distinct sorts of thing. Why is it supposed that a causal relation might obtain between the events described by C and whatever is described by A, but never an identity relation as obtains between the referent of 'Morning Star' and 'Evening Star' as described by X, Y and Z?

One reason for assuming A and C describe different occurrences is that there are ready examples where agents' arms rise without their being raised intentionally. Conversely, A can be said to be true anytime Jack tries or attempts to raise his arm though Q may be false. Although such occasions are rare, suppose Jack did try to raise his arm and failed to make it go up, due to sudden unanticipated paralysis. One may balk at the claim that Jack raised his arm since nothing happened to his arm; yet on the other hand it's not the case that Jack did nothing at the time of his failed action: he did something and it is that which he did that failed. What turned into a failure at 3:00 was Jack's trying to raise his arm: he intended something that didn't occur.

Hornsby applies this point by saying that A and C

can be expected to be true in virtue of the same nonrepeatable causal sequence just in case D is true:

(D) Jack tried successfully to raise his arm at 3:00.

Here in D, 'tried successfully' means that he tried and succeeded in doing what he was trying to do.

The sense in which Hornsby employs this term 'trying' (to describe a type of internal causal episode of an agent's doing things) allows that trying can be effortless on the part of the agent. Trying is of course itself something one can do without trying to do it; if not we are lead into a simplistic regress. Moreover, an agent can be trying to do something according to Hornsby without being aware of it at the time he is engaged in such trying. He may be aware of the effect that he is expecting to bring about or the effect that he wants, and sometimes he is aware of the effect that he is in fact bringing about. Thus an agent may be aware that his trying is a success or a failure; or he may be in a state of falsely believing something about the state of affairs he is bringing about rather than knowing any causal facts about his trying. This point is a controversial one however, and will be taken up in Chapter V when we look at different accounts of an agent's thoughts, different things he may be presumed to know, when he acts for his own reason.

In the meantime, note that sentence D, like A, may rely on the effect or the consequence reported by

sentence C in order to describe what it was that Jack attempted or tried to do at 3:00. 'Trying' and 'doing intentionally' may count as different ways of classing an occurrence depending upon whether or not it is a causal occurrence. 'Doing intentionally' and 'trying successfully' may be names for the same sub-set of 'tryings'. Thus Hornsby alleges that A, B and D may be about the very same event at 3:00, the one that caused the event of the arm going up. The occurrence described by A, B and D must be distinct from the event that C is about, to repeat a point made earlier. Hornsby argues this on two grounds already covered: one ground is that from A or from B or from their conjunction one cannot deduce C since agents' attempts may fail, or may succeed in weird ways (Jack's trying broke the heart of the physical therapist who relented and lifted the paralyzed arm for him, thereby ruining his experiment). By this reasoning, however, A and B do not entail D either. Nonetheless Hornsby argues that D (trying successfully) indirectly describes the very same occurrence that A and B indirectly describe just as long as C is true and describes just the right causal sequence connecting Jack's trying to his arm's rising in a 'nondeviant' way.

The other ground for insisting that A, B and D can never be about what C describes (the arm's rising) is this: C may be about whatever is the causal effect of the agent's trying. So whatever C describes must be distinct from what is

described by A, B and D. We did note that this last ground is not warranted either, by appealing to the example of an electrical discharge causing a lightning strike.

It has yet to be shown why the general conclusions we derived from these examples cannot apply to causal descriptions of actions as well as they apply to causal descriptions of other sorts of occurrence.

We will try to refine the difference between describing Venus and lightning on the one hand, and acts done intentionally on the other, under the head of the 'two vocabulary thesis' in the next chapter.

So far our considerations in this chapter suggest that indeterminacy in reference to actions is unlikely to distinguish thought and talk about actions from thought and talk about causal sequences of events generally. But the arguments proposed do not prove that there can be no such distinction. We have seen that if there is some kind of indeterminacy characteristic of act descriptions, it has yet to be formulated in a way that coheres with the following claim shared by the theorists we have been quoting: acts done intentionally are regarded as identifiable members of causal sequences. Acts done intentionally are connected to causal thoughts, and to effects occurring beneath the surface of an agent's body, in his bodily movements and beyond in the occurrences around him that he makes happen in ways he intended them to happen.

5. In closing this chapter it is worth noting that two different causalist action theorists abandon the point about indeterminacy when they propose strategies for identifying the mental cause of a particular act done intentionally. In this section we will look briefly at sketches of these two strategies, reserving our critical evaluation of these methods until Chapter V.

Davidson and Searle both require that at least one thought an agent has about what he is doing intentionally must be more specific and more precise with respect to his bodily movements than are the descriptions that identify his wants and desires and other thoughts that may comprise the cause of his action.

For instance, Davidson says that Jack has intentionally flipped the switch to turn on the light just in case Jack had a "pure intending" to turn on the light and that this particular mental state caused him to flip the switch just as he did.¹ Jack's pure intending is a composite causal state of beliefs and pro-attitudes: He wanted (needed, felt obliged, was compelled) to turn on the light and believed (hoped, suspected, assumed) that putting pressure on the switch with his finger would do the trick.

Comparably, Searle says that an agent's "intention in action"² is a causal state whose content determines what

¹. Davidson, Actions and Events, p. 89.

². Searle, "Intentionality," p. 258.

type of act the agent is doing intentionally. The physical component of his acting intentionally usually involves some bodily movement caused by his being in this type of intentional state at the time the bodily movement is taking place. The agent's intention in action is the factor causing his body to move in a specific fixed way at that time. These two components constitute causal aspects (mental or "intentional" cause and "physical" effect) of a "composite event" called the agent's action.¹

The description identifying the content or the object of an agent's "intention in action" does two things: The description establishes what type of action the agent intended by making the particular movements of his body that he did make. And it also establishes which particular bodily movement(s) or other events to count as the ones he brought about intentionally. Both of these frameworks, Davidson's and Searle's, will be developed in more detail in Chapter V.

Here the point to note is this: On these two different accounts of how events are identified as a certain kind of action (and how thoughts that caused the occurrence of action are discovered) the description of what the agent thinks he's doing has a crucial structural feature. The description of what the agent thinks he's doing involves a precise and specific reference

¹ John Searle, Intentionality (New York: Cambridge University Press, 1983) p. 125.

to a particular nonrepeatable event.

Both Davidson and Searle contrast the specificity characterizing the descriptive content of an agent's thought about what he is doing, with the generality that characterizes the content or object of that agent's motivating want or desire.

In Davidson's text:

. . . the event whose occurrence makes 'I turned on the light' true can't be called the object, however intentional, of 'I wanted to turn on the light'. If I turn on the light then I must have done it at a precise moment, in a particular way --every detail is fixed. But . . . [a]ny one of an indefinitely large number of actions would satisfy the want and can be considered equally eligible as its object.¹

When Searle distinguishes the component of an action that he calls the "intention in action" or the agent's "experience of acting" from the agent's attitudinal state called a "prior intention",² he draws the contrast between these two types of thought as one of relative specificity with respect to particular occurrences:

. . . the intention in action will be more determinate than the prior intention; it will include not only that my arm goes up but that it goes up in a certain way and at a certain speed, etc.³

Searle adds in a footnote:

The relative indeterminacy of prior intentions is not obvious in the case of complex actions.

¹. Davidson, Actions and Events, p. 6.

². Searle, "Intentionality," p. 266.

³. Ibid.

In . . . carrying out my [prior] intention to drive to my office, there will be a large number of subsidiary acts that are not represented by the prior intention but are represented in my intentions in action: I . . . shift gears, pass slower-moving vans, stop at red lights, change lanes . . .¹

These quotations show that the apparent vagueness of act descriptions may be overlooked in the course of analyzing the way an agent's reason-giving statement may identify the particular mental states that caused his action. A good test for the strategies proposed by Davidson and Searle that we just sketched is to see whether these methods for tracing causal thoughts can solve for the problem of deviant causal chains. We will apply this test in Chapter V to these and also to Sellars' proposal for deducing causal thoughts from reason-giving statements.

Conclusion to Chapter II

We have been trying to pinpoint the sense in which some causalist action theorists regard act description as different from other ways of referring to particular occurrences.

First we reviewed the dispositional character of action types by contrasting the truth conditions for a sentence containing some cognate of the nominal 'arm rising' with the truth conditions for a sentence containing a cognate of 'intentionally making a legal turn'.

¹ Searle, "Intentionality," p. 280.

We then considered the implications of assuming that actions done intentionally are unseen causes of observable effects, describable only in terms of those observable effects or anticipated (intended) effects. What we found was that nothing seems to follow about occurrences of action specifically, from the general fact that descriptions of causes are distinguishable from descriptions of their effects.

We conjectured in a summary and inconclusive way about the possibilities of applying to action sentences Quine's reconstruction of certain predicates featuring "referential opacity."¹ We noted that facts about what an agent thinks he's doing as well as facts about what he is doing function jointly to determine the type of act description that applies to what he's done. Such facts determine the 'success' or 'failure' of that action. Success and failure are decidedly properties attributable to actions, or to intentions carried out in actions. And the attribution of success or failure to actions depends upon identifying an agent's intending to do something--and his carrying through with doing it--with a particular nonrepeatable causal sequence of events. Correlatively, contingent facts about each case of acting intentionally will probably determine the mode of reference to particulars

¹ Quine, Word and Object, p. 152-156.

operating in any given action sentence. We can conclude then with the following: the reference to particulars occurring in a given action sentence will depend upon facts contingent to each case of acting. The relevant facts here pertain not only to what the agent has brought about on that particular occasion, but also pertain to what he thinks he made happen and what he intended or wanted as a result.

We concluded this discussion by expecting that no uniform pattern is likely to emerge on behalf of every predicate depicting occurrences of action. In this respect 'singular action sentences' aren't distinguishable from other singular sentences describing other sorts of occurrence.

These results did not indicate conclusively whether act descriptions function in a way that other descriptions cannot. We did gain one definitive result however, from applying a central argument of Davidson's in defense of his reading of singular causal statements: If actions are identifiable members of particular causal sequences, and if causal facts about actions can be deduced from sentences containing act descriptions, then it's inconsistent to assume the reference of an action sentence can be less determinate than the references of other sentences that describe causal sequences.

Setting aside attempts to clarify the notion of indeterminacy as a distinctive characteristic of references to actions, we noted the following in two accounts of the

way thoughts causing actions can be traced. When Davidson and Searle describe the thought content identifying the causal state of an agent acting intentionally, both theorists abandon the view that act descriptions can never refer to fixed bodily movements in any determinate way.

Yet Davidson explicitly claims that no singular action sentence can describe an individual nonrepeatable event.¹ He does regard every particular action as being identical to some nonrepeatable event, as does Searle. One may wonder what is the correct way, then, to construe what a singular action sentence is about, if one believes that actions have discernible causes. This is the topic of discussion in the next chapter.

¹. Davidson, "The Individuation of Events" (1969), Actions and Events, pp. 163-180.

CHAPTER III

ACT INDIVIDUATION AND THE LOGICAL STRUCTURE
OF SINGULAR ACTION SENTENCES

Introduction

Consider the following sentence describing an action:

- (1) Jack sabotaged the bridge by triggering the dynamite at 3:00.

If sentence (1) is true, how many particular actions does it indicate occurred at 3:00, one or two?

This question can be read as a question about how to individuate actions, to which contemporary causalist action theorists provide varied answers.¹

Here is why the subject of act individuation is important for understanding the relation between actions and the reasons that explain them.

Suppose Davidson is correct in his analysis of "rationalizations" or "primary reason"-giving statements as a species of singular causal statements.² Then the set

¹Davidson, Actions and Events, p. 163; Lawrence H. Davis, Theory of Action, Foundations of Philosophy Series (Englewood Cliffs: Prentice-Hall, 1979) pp. 27-28; Goldman, A Theory of Human Action, p. 10; Hornsby, Actions, p. 6.

²Davidson, Actions and Events, p. 9.

of truth conditions for a primary reason-giving statement includes at least two distinct occurrences or "events".¹ Simplifying Davidson's claim grossly: one of these occurrences is identical to the onset of a mental state or conglomerate of states constituting the agent's primary reason for acting, while the other event is identical to some bodily movement caused by the onset of that particular mental state.

Now if a correct reason-giving statement R could give a guaranteed indication of precisely which pair of occurrences accounts for the truth of R on one occasion of use, then there should be no threat of a 'deviant causal chain' falsifying the explanation conveyed by an agent's asserting R. For to know the truth conditions of R would be to know precisely which particular thoughts caused precisely which bodily movements occurring relative to the time and place sentence R was asserted by the agent in giving his reason for acting.

According to another way that Davidson classes reason-giving statements, R may be construed as a kind of singular action sentence: R "redescribes" an action from the agent's point of view.² As will be spelled out Davidson says all singular action sentences are existentially generalizable.³ We need to study whether the existential

¹. Davidson, Actions and Events, p. 154.

². Ibid., p. 10. ³. Ibid., p. 117.

generalization of R precludes the sentence from designating any one nonrepeatable event sequence as the action R explains.

In section one we will follow for a short distance Davidson's suggestion for assimilating these two analyses: Davidson recommends "relativiz[ing] the properties of truth and reference of a sentence to a speaker and a time".¹ We haven't space here to pursue Davidson's more general remarks about the property of truth. Taking his suggestion at face value we will see how the "relativization" might determine the reference of a singular action sentence despite its feature of existential generalization, given the way Davidson construes the logical structure of this class of sentences.

It is much harder to conjure a way around the problem of deviant causal chains, so as to be certain that a reason-giving type of action sentence is correct. Relativization to a speaker and a time is apparently less helpful here: Recall that agents giving their own reasons may be ignorant about what they are doing. So a reason-giving statement might be warranted for an agent to assert yet the sentence R that he asserts might yield a false description of the event chain that occurred because he acted.

Not all action theorists subscribe to Davidson's reading of the truth conditions for a singular action

¹. Davidson, Actions and Events, p. 129, n.17.

sentence. Traditionally Goldman has been cited as holding a view contrary to Davidson's. The contrast will be spelled out in detail in section two. Here is a sketch of it in brief: On Davidson's view every particular action is identical to some nonrepeatable event. A singular action sentence is true in case there occurs at least one event of the class or type of action described by that sentence. On Goldman's view, the truth conditions for a singular action sentence may not include the occurrence of any particular event whatsoever. A true action sentence does not describe any class of events either. Rather, it describes a fact about certain properties of an agent in evidence at a specific time.

In section three I will argue that the apparent difference between these two viewpoints has to do with what sort of entities actions are, and not with how to count them. Davidson and Goldman seem to be disagreeing most directly about what sort of entities can be members of causal sequences. This is what I will try to show.

I am not suggesting that the argument between Davidson and Goldman about act individuation is superficial or merely a verbal dispute. To understand more fully the contrast between their views on individuating actions I would have to know about the difference between facts and events, and this subject goes beyond the scope of my research. There is however one point where Goldman diverges from Davidson that is open to criticism on very basic grounds.

Goldman claims that the meaning of certain names for actions can help determine the number of actions an agent performs at one time; one obvious trouble with this will be sketched briefly in section two. I mention it because Goldman's claim presses the question of how to reap causal facts from a special action phraseology. This recurrent issue will claim our entire focus in the chapter following this one.

The overall interpretation of Davidson and of Goldman offered here is suggested by the fact that both of them claim particular actions are caused by particular thoughts. And they both claim that the 'content' or the 'object' of particular action-causing thoughts can be deduced from an agent's own reason-giving description of what he thinks he's doing.

In a concluding section I will show that if one accepts these two claims about thoughts causing actions, then it appears to be inconsistent to claim that any reference to particular occurrences of action must be indeterminate. Essentially the same *reductio ad absurdum*, or so-called Fregean argument¹ applies here as it did when first borrowed from Davidson's "Causal Relations" essay, in the preceding Chapter II. The demonstration of this argument will be mentioned in this chapter rather than spelled out again.

¹. Davidson, Actions and Events, pp. 153, n.4, 169.

1. According to Davidson, sentence

(1) 'Jack sabotaged the bridge by triggering the dynamite at 3:00'

describes a class of events, not any nonrepeatable particular event. The predicate in the sentence indicates what properties of events mark the members of this class. An occurrence that can account for the truth of (1) is a member of the class of things done at 3:00 and of things done by triggering dynamite, and of bridge sabotages.

Any occurrence that can be described with respect to all of these classes conjointly is an event that satisfies the truth of sentence (1). So (1) is true if there occurs at least one such event; but it is also true if there occurs more than one such event.

Since the truth of (1) is consistent with the occurrence of more than one event, Davidson argues that the reference of (1) does not discern among the events that count in the set of truth conditions for this sentence. That is, the reference of the sentence as a whole or in any part cannot be construed correctly as being identical to any one of these particular events to the exclusion of others comprising the class described. Thus Davidson argues that a singular action sentence does not denote any particular event; its reference is "existential and general".¹

¹. Davidson, Actions and Events, p. 169; see also Hornsby, Actions, p. 17 where she appeals to Davidson's view.

Davidson contrasts singular action sentences with descriptive expressions that function as singular terms within sentences. For example the sentence 'Jack ran in the New York Marathon' contrasts with the nominals 'Jack's running in the New York Marathon' and 'Jack's run'. The nominal 'Jack's run' functions like the name 'Jack' and so it does not have a truth value. Also like a name and unlike a sentence, the nominal may function to pick out just one run on a given occasion of its use. On Marathon Day in New York, 'Jack's run' can function to designate one racer's efforts in the way 'Jack' refers to one racer, though there are many men called Jack running in the same race.

Davidson's laying stress upon the existential generalization of singular action sentences might provoke the following question: Of the many events whose occurrences satisfy the truth conditions for a singular action sentence, which event is supposed to count as the one that the sentence is "about" or "denotes"?¹ Hornsby challenges Goldman's analysis of individuation with this question, not Davidson's. She argues for Davidson's position in the pages cited below.

Davidson's diagnosis of what is wrong with this question comprises much of his essay "The Individuation

¹•Hornsby, Actions, pp. 17, 19.

of Events" (1969). There Davidson says it is "wrongheaded"¹ to assume that an indeterminacy or ambiguity characterizes the reference of a singular action sentence because it is existential and general. He says that although a singular action sentence does not describe any act token or particular, the nonrepeatable occurrence of some act particular "accounts for"² the truth of the sentence on a given occasion of its use.

Suppose sentence (1) is asserted at 3:00 p.m. on January 1, 1986 on the bank of the Bocay River in Nicaragua. Then one should construe the reference of (1) as being whichever event is most proximate to the assertion at 3:00, i.e., an event whose occurrence verifies sentence (1).³

How is a given action sentence related to one nonrepeatable occurrence and not another event occurring at the same time and describable in the same way?

Is the question posed here different from a generalizable demand for a theory of meaning applicable to all natural language participation? Perhaps the problem of understanding how act descriptions refer as a distinct class issues from the misguided attempt to isolate the function of act descriptions, as if these expressions operate separately from other sorts of expression.

¹. Davidson, Actions and Events, p. 169.

². Ibid.

³. Hornsby, Actions, p. 73.

On the other hand, intuitions have guided us all along to recognize some difficulty in locating the precise occurrence called 'Jack's running in the race'. In contrast, the same sort of difficulty doesn't seem to arise in locating the objects designated by the phrase 'Jack's sneakers' at any point throughout the time frame or duration of his running the race. From this standpoint, the question just raised about relating a singular action sentence to an individual and nonrepeatable event sequence may be read as a version of the question raised initially in Chapter I: How does one tell whether a particular bridge collapse belongs to the class of successful sabotagings or to the class of failed repair jobs? Again the question arose in a different example later in Chapter II: How does an onlooker know to count a particular clenching of Jack's fist and jaw as evidence that Jack is angry? These questions seem different from asking: Are those things on Jack's feet sneakers or are they track shoes? Questions about how to apply act descriptions are different from the application of general terms to objects, in that the type of action occurring at 3:00 depends in part on the causal thoughts occurring in an agent at 3:00. We will pursue this consideration further in Chapter V under the heading 'radical interpretation'.

For the present discussion let us shift our focus of attention to sentences containing 'action' terms and phrases, by supposing we already have an account of how

the meaning or function of those component terms and phrases fits into the overall picture of natural language participation. Now we are free to see why the existential generalization of a singular action sentence like (1)--on Davidson's reading--does not by itself dictate that the reference of (1) will have to be ambiguous for a given "use-occasion."¹

To show this, consider Davidson's reply to the question posed initially in this chapter, about how to individuate actions described by the true sentence

- (1) Jack sabotaged the bridge by triggering the dynamite at 3:00.

Davidson says that the phrases 'Jack's sabotaging the bridge' and 'Jack's triggering the dynamite' refer to exactly the same event or sequence just in case at least one event or sequence occurred that accounts for or verifies the truth of this conjunction, (2) and (3):

- (2) Jack sabotaged the bridge at 3:00; and

- (3) Jack triggered the dynamite at 3:00.

Here are grounds for Davidson's reply and the addenda that make it a fruitful one.

Davidson claims that every occurrence of an action is identical to some particular nonrepeatable occurrence. He calls this sort of nonrepeatable particular an event. He gives his reasons for this identity claim in different discussions of why an ontology should include particular

¹. Sellars, "Actions and Events," p. 197.

events.¹ As noted before I am unequipped to evaluate these discussions here, as this would presuppose familiarity that I don't have with alternative ways of comparing events and facts.

In order to follow Davidson's reasoning about action sentences suppose for now that this identity claim is correct.

Then with respect to individuating events of all sorts generally, Davidson argues that two descriptions refer to the exact same nonrepeatable event just in case their referents have "exactly the same causes and effects".²

Davidson then argues that the truth of the singular action sentence (1) and of its paraphrase (2)&(3), requires that there exists at least one event whose occurrence is identical to the act described by these three sentences.

Again, the question raised earlier might be brought to challenge Davidson here by pressing the problem of individuating actions: Was it the same event or was it distinct and simultaneous events comprising membership in the class described by sentence (1) and by the conjunction of predicates in sentences (2) and (3)?

In "The Individuation of Events" Davidson argues that if this question is not misguided then it must be a question about an empirical matter of fact regarding dated

¹. Davidson, Actions and Events, Essays 6-10.

². Ibid., p. 179.

particular events. It is not a question about the semantics or the syntax peculiar to predicates describing actions, as Goldman's analysis might suggest.

The way to tell whether (2) and (3) are true of the exact same event occurring at 3:00 is to investigate the remains of the bridge after its collapse. One needs to examine the bomb plunger for fingerprints after the explosion of the bridge, and to trace the whereabouts of Jack throughout the day of the explosion. Consider whatever empirical work is required to establish the causes and effects of the bridge collapse, of the bomb's triggering at 3:00 and of all Jack's chatter over the last two years about sabotaging the enemy. This is the same work required to establish whether this particular bridge explosion is identical to a bridge sabotaging performed by Jack, or whether these occurrences were coincidental.

To establish whether this particular explosion was an instance of Jack's sabotaging, one cannot study the meaning of the words used nor the syntax of the sentence asserted in describing the incident. This point will be brought out again, perhaps more sharply, as we study Goldman's overall view of act individuation.

On its face, Goldman's analysis involves something he calls a "synonymy test"¹ to decipher the numbers of act

¹Goldman, Theory of Human Action, p. 13.

tokens exhibited by an agent on one occasion. This strategy of relying on the meaning of phrases to individuate actions seems flawed for reasons I will show in the next section. This feature of his account aside, however, there appear to be fundamental parallels between Goldman's attribution of act properties to an agent on a given occasion, and Davidson's projection of the occurrence of events on the basis of an action sentence's truth conditions for one occasion of its use. So I will argue in section three.

2. At the outset of his analysis of sentence

- (1) Jack sabotaged the bridge by triggering the dynamite at 3:00,

Goldman observes that each descriptive component in its predicate can be reformulated as a nominal phrase:

'Jack's sabotaging the bridge' and 'Jack's triggering the dynamite'.

He notes that these participial phrases are not synonymous, and claims that therefore they name distinct "act-properties."¹ Any act-property described in the predicate of a true sentence whose subject is an agent is an act-property "exemplified" by that agent at a specific time.² Two act-properties A and B are exemplified at 3:00 by Jack just in case there occur two distinct act-tokens at 3:00, one of each type A and type B.

¹. Goldman, Theory of Human Action, p. 10.

². Ibid.

Judging from the way that Goldman posits the co-occurrence of many act-tokens at one time for one agent, it seems that his aim is to discern the different ways an agent's doing one sort of thing can be related to his doing something else at the same time. For this purpose, Goldman proposes "as a guide . . . and not more than that"¹ a taxonomy featuring four basic types of relation between act-tokens.² The list may be taken to reflect his assumption that the type of action correctly attributed to an agent at a single time is determined by a combination of factors: It is determined partly by how the incident of behavior happens to be described, partly by causal facts about the occurrence of the behavior itself, and partly by incidental facts about the agent and the circumstances in which he was disposed to do something intentionally.

For example, since Jack is an American and his bridge sabotaging occurred in Nicaragua, and since there exist rules regarding conspired destruction of municipal property by foreigners not engaged in declared combat, this particular bridge sabotaging "conventionally generated"³ Jack's breaking an international law at 3:00, January 1, 1986. Likewise, since there was a dynamite bomb rigged to the support beams of the bridge, Jack's pressing down on the bomb plunger "causally generated"⁴ his exploding the

¹ Goldman, Theory of Human Action, p. 30.

² Ibid., p. 22. ³ Ibid., p. 25. ⁴ Ibid. p. 24.

bridge. Further, his pressing down on the bomb plunger on this occasion was "augmented"¹ by his pressing down on it very quickly while kneeling with his eyes squeezed shut.

How many act tokens can be exemplified by an agent at one time after this fashion? How many relations can there be between different pairs among the act tokens that are exemplified by an agent on one occasion?

Some critics of Goldman's view complain that the answers to these questions are indeterminable.² But some limitations are obvious in determining relations among actions construed as properties exemplified on the very same occasion or at a single time.

Clearly, Daisy cannot exemplify both the property of swimming the English Channel and of winning the New York Roadrunners' Marathon at the same time. Whether she can ever exemplify either of these act properties, and what her exemplifying of one would have to do with her exemplifying of the other, both seem to be questions properly posed as research topics in athletic training.

An unlikely route to go in answering these questions would seem to be studying the meanings of the terms or the syntactic structures of the phrases comprising descriptions of these activities.

On Goldman's account it is not obvious how to go

¹ Goldman, Theory of Human Action, p. 25.

² Davis, Theory of Action, p. 30.

about empirically determining which act properties an agent is going to exemplify at a specific time. But neither is it evident on his account that this would be impossible to do. If act-properties were like the properties of being soluble, being brittle or being blue, perhaps it would be apparent what methods to use for discerning which act-property a single agent will exemplify under a given set of conditions.

One counts on molecular structure, for instance, to predict whether a particular sample of some type of substance is going to dissolve when placed in water at a certain temperature. As it stands, if there are such properties as act-properties they seem likely to be very complex, like allergic reactions. If this were an empirical fact it might vastly complicate, or perhaps even preclude, any practical method of predicting in detail the exemplification of a certain type of act-property like bridge-sabotaging.

Apart from such conjectures, Hornsby has rejected Goldman's analysis of act properties on the grounds that it neglects the ready fact that one occurrence can be classed or described according to more than one type of action. Here is Hornsby's argument:

If Goldman's theory of act individuation were correct, then it would be impossible for onlookers to witness an agent's doing one sort of thing by means of

his doing something else.¹ But one can observe, for example, an agent's turning on a light by flipping a light switch with his finger. Many counterinstances show that a sentence containing non-synonymous act descriptions flanking the 'by' connective may be true in virtue of a single occurrence.² So Hornsby concludes, *modus tollens*, that there is something wrong with a theory implying that two descriptive components in a complex predicate can't have the very same referent if they are non-synonymous.³

To illustrate her argument with an example: on Goldman's view a stabbing could never be identical to a killing because 'stabbing' is not synonymous with 'killing'. Hornsby points out that Goldman is committed to the conclusion that after stabbing Caesar, Brutus had yet a further action to perform, since it is true that Brutus killed Caesar. Hornsby argues that the truth of 'Brutus killed Caesar by stabbing him' can be accounted for more simply and readily by the thesis that his killing Caesar is identical to his stabbing Caesar on the occasion so described.

This objection by Hornsby seems to run parallel to a criticism that Davidson has lodged against Kim's criterion

¹.Hornsby, Actions, p. 18.

².Ibid., pp. 6-8, 38. ³.Ibid., p. 19.

of event individuation.¹ Here I briefly sketch this parallel. However without making a more extensive comparison of 'act-tokens' and event particulars, I cannot tell whether Goldman's view of actions as exemplified properties is vulnerable to the difficulty that Davidson finds on Kim's view of events as particulars. Drawing the parallel is fruitful in other ways, even so:

Davidson objects that Kim's method of individuating events precludes the possibility of formulating knowledge claims after the pattern $[a=b]$ where a and b are singular referring expressions that designate nonrepeatable events. Kim's account implies that no two sentences true under the same set of conditions can refer to exactly the same event. For two sentences true under the same set of conditions are distinct only if they contain distinct particulars counting as their referents. And two descriptions are distinct whenever they attribute different properties to some particular. But particulars with different properties attributed correctly to them cannot be identical to one another. So two distinct statements true at t cannot designate the numerically same event occurring at t . (This capsulation of Kim's view comes from Davidson's reading of it in "The Individuation of Events.")²

Davidson cites counterexamples to Kim's analysis:

¹. Davidson, Actions and Events, p. 170.

². Ibid.

Some stabbings are (identical to) killings because they happen to induce death; and some bridge explosions are sabotages because they were induced deliberately.

Goldman might argue with these identity claims thus: A particular stabbing cannot be identical to a killing if the reason and the cause for one is not the reason and cause for the other's occurrence. For example, the reason for Brutus' stabbing Caesar is not the reason for Brutus' killing Caesar. The cause of Brutus' killing Caesar was partly his puncturing the aortic valve in Caesar's heart with a knife; whereas the cause of Brutus' stabbing Caesar with a knife had nothing to do with Caesar's aortic valve. The cause of the stabbing was in part Brutus' choice to use a knife instead of poison in order to avenge his injured honour or whatever.

In reply to this objection, Davidson concedes that different reasons apply for stabbing Caesar and for killing Caesar. But Davidson claims the difference derives from the nature of explanation, not from the identity of the episode so explained. Thus: reasons explain the truth of "sentences or propositions, not . . . what the sentences are about."¹

For example, different reasons may apply in explaining the stabbing and the killing, while 'stabbing' and

¹. Davidson, Actions and Events, p. 171.

'killing' abbreviate two descriptions of the very same sequence of events.

Davidson attributes Goldman's objection to this identity claim (i.e. regarding this stabbing as identical to a killing), as the result of confusing cause with effect in the event sequence constituting Brutus' taking Caesar's life. Davidson says¹ that if the reason for an action is its cause, then it is incorrect to refer to the reason for Brutus' killing Caesar by using the phrase 'Brutus' stabbing Caesar'. The stabbing that Brutus did was not the cause of Brutus' killing Caesar (contrary to Goldman's doctrine of causal generation of act-properties). What his stabbing caused was Caesar's death. 'Brutus' killing Caesar' likewise describes the cause of Caesar's having died as he did. Death was one effect of Brutus' action, called a 'stabbing', 'killing' or 'assassination'. And Brutus' action against Caesar's life can be explained by citing different reasons, depending upon whether his action is described as a killing, a stabbing, a political turning point, or a puncturing of the aortic valve.

Davidson argues in another place ("Causal Relations") that the explanation for an event's occurrence may be entailed by a descriptive sentence referring to that

¹. Actions and Events, p. 14.

particular event if the description is "sufficiently rich" to entail causal facts about the event.¹ Now suppose we accept this claim just at face value for the moment; it may suggest how to apply Davidson's criticism about confusing cause and effect in a way that is favorable to Goldman's method of contrasting act descriptions.

Perhaps one can modify the nominalizations 'Brutus' stabbing Caesar' and 'Brutus' killing Caesar' in a way that exhibits the synonymy relation Goldman posits between two descriptions that can refer to the exact same act token. Perhaps, for example, 'Brutus fatally stabbed Caesar' would be regarded by Goldman as synonymous with the description of what happened in 'Brutus killed Caesar with a knife'. Then Goldman's view would resolve with Davidson's, to the verdict that this particular fatal stabbing is identical to the killing of Caesar.

But the weak point in Goldman's account remains his "synonymy test."² On this test if two descriptive phrases "express" the same act property then they can refer to the same act token. But the test is not meant to be applied to phrases that only "indirectly refer" to properties rather than "expressing" properties.³ 'Being blue' counts as a phrase that "expresses" the same act property that the phrases 'the color of the sky' and

¹. Davidson, Actions and Events, p. 156.

². Goldman, Theory of Human Action, p. 13.

³. Ibid.

'my favorite color' refer to only "indirectly". What about the phrase 'the color of this skirt'; does this "express" or "indirectly" refer to blue? It is unclear how to tell which phrases are eligible for the synonymy test to identify the number of act-properties exemplified at a certain time.

The upshot here may echo a conclusion arrived at earlier in reviewing Davidson's analysis of act descriptions in relation to particular occurrences.

In order to tell whether the occurrence or the exemplifying of Brutus' stabbing Caesar had the "same causes and effects"¹ as his killing Caesar, one has to examine Caesar's corpse for stab wounds, and one has to search for Brutus' fingerprints on the weapon that was used to inflict this very set of stab wounds. To establish whether this particular instance of killing Caesar is identical to the stabbing done by Brutus, one cannot compare the meaning of component phrases used to describe the incident.

One might still resist identifying any stabbing with a killing on the basis of this empirical consideration: Since Brutus' stabbing Caesar might not have been fatal, his act of stabbing cannot be identified with his act of killing the victim. Likewise, Jack's plan to sabotage the bridge might be carried out completely yet the bomb

¹. Davidson, Actions and Events, p. 179.

might fail to go off; meanwhile suppose a truck carrying gasoline overturned accidentally and fuel ignited the bridge to the point of explosion. The sabotage was carried out at 3:00 and the bridge exploded at 3:00 but the sabotage of the bridge was not identical to the explosion of the bridge. Hence in the instance where the sabotage is carried through without a hitch so that the bomb explodes the bridge according to plan, it is still the case that Jack's act of sabotage cannot be identified correctly with his act of exploding the bridge.

Here as in other discussions of action identification, it is important to avoid conflating different senses of the cognates 'identity' and 'identification' as occur in the following phrases:

- 'identity of an X-ing',
- 'being identical to an X-ing',
- 'conditions for identifying an X-ing',
- 'being identified with an X-ing',
- 'identifying an X-ing by a Y-ing'.

One may confuse these different senses when one appeals to the essential or 'necessary' features of an occurrence in order to establish which distinct act descriptions may refer to it.

For example, Hornsby argues that Brutus' stabbing Caesar was a necessary feature of his killing Caesar in just the way he did; so the stabbing and the killing were identical

on the occasion of Caesar's death.¹

According to Hornsby's argument, if A and B describe actions and if an act of type A "needs to have been done if" an agent is to have done an act of type B, then A and B refer to the exact same occurrence of his action on at least one occasion.²

In the current example, Hornsby notes that there "was nothing else [Brutus] . . . needs to have been doing" to kill Caesar other than stab him in just the way he did.³ Further, had Brutus not stabbed Caesar on the occasion in question in just the way he did, then Caesar would not have died in just the way he did, by the hand of Brutus. So, she concludes, that particular stabbing of Caesar is identical to the one and only killing of Caesar.

There is a trivial sense of being a 'necessary component', in which every feature of an episode or sequence of events counts in order for that occurrence to be the very occurrence that it is and not another.

But in this sense there is nothing about the necessary components of an event that indicates what its identity is with respect to one type of action and not some other type. To establish whether Brutus' stabbing

¹.Hornsby, Actions, p. 9. I am transposing Hornsby's points to apply to this example, to keep the illustrations in our discussion here consistent. In her text Hornsby refers to Anscombe's example of someone poisoning the inhabitants of a house by pumping water from a contaminated well.

².Ibid.

³.Ibid.

Caesar counts as a murder and not manslaughter, one must determine the presence or absence of certain relevant intentions, attitudes and beliefs on the part of Brutus. This is so just as one must learn how Caesar's heart fared the puncture with the knife in order to know if Brutus' stabbing him counts as his killing Caesar.

Suppose Brutus' stabbing Caesar was a murder. It's not possible for that particular sequence of events including Brutus' bodily movements and their effects on Caesar's body to have been a case of manslaughter or severe injury, if indeed they constituted a murder. Whereas it is possible for us to classify the sequence in some other way besides calling it a 'murder' (e.g. 'assassination', 'fated error' and so on).

What about Brutus' name for his action?

The occurrence of the thoughts that Brutus had must figure in some way to determine the type of action he was doing intentionally.

Yet note it is the occurrence of these thoughts, or the fact of their occurrence, whatever their content may be, that figures into the particular nonrepeatable causal sequence counting as his reason leading to his action.

Hence what does not appear crucial is the meaning or the syntax or the grammar of the expressions that occur in descriptions of that particular sequence, to establish what type of action occurred. But then what significance can the agent's reason-giving statement have in determining

the cause of his own action?

Descriptions and syntactic features of sentences describing an action's occurrence are indeed crucial but only insofar as they may represent somehow the 'content' or the 'object' of an agent's thoughts about what he is doing. In this capacity a description may help to track or otherwise identify the particular thoughts or attitudinal states that causally contributed to the agent's doing what he did intentionally. Thoughts causally contribute to action by virtue of their occurrence or by virtue of their obtaining some other way at the time the agent acted. Or so the causalist action theorist assumes.

But since agents can be wrong about what they are doing, it is not obvious how the descriptive content of an agent's beliefs does indicate which subsequent events to count as the ones caused by his thoughts. So again our considerations point to the need for scrutinizing the role that an agent's beliefs about his action must play, if what he says in earnest about his action is to yield causal facts about some nonrepeatable sequence of events. We will take up this concern again in Chapter V.

3. Next we can compare similarities between Davidson's and Goldman's methods of individuating actions.

If we cut away acts done intentionally from their effects in a way that Davidson, Goldman and Hornsby all seem content to do, it is not obvious how the distinction is drawn

between, for example: a particular sabotage that was successful and a particular attempt at bridge sabotage that failed. Nor is it obvious how to tell when a stabbing counts as a killing. Here's the problem.

Recall that Goldman's analysis gives no notice of what events to count as the particular "exemplifying" of a type of act-property in evidence by an agent at a fixed time. On his view, no event in particular accounts for the truth of an action sentence that attributes the property of 'bridge sabotaging' to Jack at 3:00. It appears to follow that the sentence 'Jack sabotaged the bridge at 3:00' can be true though nothing discernible happens to any particular bridge. Stated differently this result seems less queer: By 3:04 it was apparent that Jack had failed in his undertaking--perhaps the bomb he set was a dud. Nonetheless it remains correct to say that it was his undertaking to sabotage the bridge at 3:00 that failed. 'Sabotaging the bridge' indeed describes what Jack was doing intentionally when he was depressing the bomb plunger at 3:00 with his eyes squeezed shut.

It is this consequence of Goldman's view of act properties and their "exemplifications" that appears consistent with the consequences of Davidson's way of locating a particular event identical to an act done intentionally.

Davidson says that every action occurs just at the surface of some agent's body. Hornsby argues that it is

more consistent with Davidson's initial assumptions to locate acts somewhere inside the agent's body.¹ Just as Goldman's view seems to imply, the action described in (1) is done when Jack's attempt at sabotage is done, or at the moment he depresses the bomb plunger; according to Davidson and to Hornsby: "the rest is up to nature."²

But surely in order to describe adequately what it was Jack was doing intentionally or trying to do when he pressed the plunger at 3:00--let alone to identify what it was he succeeded or failed doing by depressing the plunger just so--some mention must be made of a bridge. It was a particular bridge after all, and not some lamp post, that Jack meant to explode.

It is not controversial to note that the effects of bodily movements do indeed matter to the correct description of a causal sequence as a failure or as a success of an action of a certain kind. And yet again, the agent's own description of what he intends to do or to bring about can neither entail nor otherwise guarantee a description of what he brings about in fact.

Now we can observe a similarity between the proposals for organizing these considerations that Davidson and Goldman each offer independently. The result they are both aiming for is a way to establish the right connection

¹. Davidson, Actions and Events, p. 59; Hornsby, Actions, pp. 9, 14.

². Ibid.

between a nonrepeatable occurrence and the warranted assertion of a reason-giving statement. This is actually an extension of the agenda that Davidson and Goldman have set themselves: more precisely, their respective concerns address the connection between a nonrepeatable occurrence and the warranted assertion of a singular action statement. But if reason-giving statements count among singular action sentences, then the extension seems tenable.

At any rate, a correct analysis of this connection should reveal how the reason-giving statement manages to causally explain ("redescribe")¹ that nonrepeatable occurrence. So a good test for the analyses offered by Davidson and Goldman is this: do their accounts of this connection avoid the possibility of a reason-giving statement 'explaining' a deviant causal chain, thereby falsely attributing properties of being intended or attempted to an event that is not, in fact, an act done intentionally for an agent's own reason?

Demonstrating the similarity between Davidson's and Goldman's analyses of particular act descriptions requires appeal to the following parallel. Suppose Goldman's

¹Davidson, Actions and Events, p. 10. Although Goldman does not like the locution 'redescribing' issuing from Melden's use in Free Action (1961) p. 88, Goldman does regard reason-giving as a causal description of an occurrence "in terms of the agent's action plan. . ." see A Theory of Human Action, pp. 78-79.

attribution of act properties to agents is parallel in some ways to ascribing beliefs to speakers. This supposition will be tested later in Chapter V when we try to use models of 'radical interpretation' of utterances to analyze the interpretation of certain events as actions 'rationalized' by the content of certain causal thoughts.

To illustrate the suggested parallel consider the fact about Jack that he believes that grass is green. This fact is true of Jack whether or not grass is green. Suppose the fact that Jack believes grass is green entails that Jack has the property of being a believer that grass is green.

Now consider the fact about Jack that he is intentionally sabotaging a bridge at 3:00. Again it may be true of Jack that he is intentionally sabotaging a particular bridge whether his attempt succeeds or fails.

It will be useful to formalize this parallel. Let Jack's believing P be stated simply by this sentence:

(4) Jack believes that P.

Sentence (4) appears to report a fact about Jack without referring directly to any unique occurrence or event. Yet considering the set of conditions in virtue of which (4) is true, it may also be true under that same set of conditions that

(5) There is a particular dispositional state of belief, identifiable as having the content 'P', and this particular state (or the event of its onset) is attributable to Jack or locatable in Jack's neuro-physiological structure at a specific time.

Does the truth of (4) entail sentence (5)?

Suppose for argument's sake that it does. Then the parallel between attributing propositional attitudes and attributing actions may suggest the following about Goldman's and Davidson's analyses of singular action sentences.

A 'Goldman' action sentence can designate at least one particular "exemplification"¹ or act token when and only when there occurs at least one event that "accounts for"² the truth of a 'Davidson' action sentence.³

4. If this interpretation of Davidson's and of Goldman's analyses of action sentences is tenable then it carries further implications about the correct way to construe the reference to a particular sequence constituting an act done intentionally.

As mentioned already, Davidson and Goldman are among those action theorists who believe that a certain reason-giving statement about an act can yield a singular description of a nonrepeatable causal sequence occurring at a specifiable time and place. So it may be correct to assume that they believe there is a specifiable fixed number of particular acts that a true reason-giving statement causally explains. Of course these theorists

¹. Goldman, Theory of Human Action, p. 10.

². Davidson, Actions and Events, p. 169.

³. These name-labels are dubs, obviously, for each theorist's respective way of construing an action sentence.

may agree that there is just one correct number of actions explained by a particular reason-giving statement yet disagree about what that number is. Can Davidson and Goldman also maintain coherently that the reference to particular actions is indeterminate?

In Chapter II we already noted that if the deducibility of causal thoughts thesis is correct, then the unique indeterminacy thesis may lead to this absurdity: The truth of Jack's reason-giving statement entails that Jack's doing something intentionally at 3:00 is causally related to anything that happened at 3:00. Under ordinary circumstances, we can expect that an agent's report of what he was trying or intending to do will block any absurd causal inferences: Jack himself can tell us what sort of thing he had his mind set on doing, and this is enough to offset the erroneous supposition that his trying at 3:00 caused anything at all that happened at 3:00. Yet recall again that Jack may be ignorant of what else his trying at 3:00 is causally connected to. Just this fact, in principle, can preclude our drawing any warranted causal inferences if indeed these inferences are based on truth functional rules alone.

To show this, consider some of our sentences from Chapter II again:

- (A) Jack raised his arm intentionally at 3:00; and
- (E) Jack caused his arm to rise at 3:00.

On causalist analyses that Davidson and Goldman subscribe to, A and E both describe some causal sequence constituting what Jack did intentionally at 3:00. As the analysis we have been scrutinizing seems to imply, A refers only 'indirectly' to that nonrepeatable sequence, by 'entailing' a causal statement that refers to the sequence directly. On one standard reading, the subject term 'Jack' in E is construed as standing elliptically for specific attitudinal states or events occurring inside Jack's body, causally connected to his arm's rising.¹

For the moment let us suppose that the thesis of 'unique indeterminacy in reference to actions' applies to sentence A and not to sentence E, as suggested above. A is an action sentence; E is about a causal sequence. Let this causal sequence include Jack's arm rising (event b) and also a particular mental state or conglomerate of states (a) whose occurrence is noted by the ellipsis 'Jack' in sentence E. Suppose a caused b. Though a might be monitored (perhaps not isolated) by neurotechnology it might be supposed that a cannot be observed in the direct way that b's occurrence is observed. I think this

¹. Davidson, Actions and Events, p. 19; Goldman, Theory of Human Action, p. 18; Hornsby, Actions, pp. 98-101. Here is another natural place to diverge and introduce the controversy about whether there is such a function as 'agent causality' distinguishable from 'ordinary event causality'. Again, I haven't the space here nor enough background in the literature about causation to canvass this subject adequately.

is a shaky contrast, but its suggestion is one possible motivation for the tendency to contrast action sentence A with this sentence:

(C) Jack's arm rose at 3:00.

C refers to event b in a way that may be regarded as more 'direct' than the way action sentence A refers to the causal sequence constituting Jack's raising his arm intentionally. At any rate, here's our model of the two theses--'deducibility of causal thoughts' and 'indeterminacy in reference to actions'--side by side.

Now according to the deducibility thesis, an onlooker who is warranted in believing A and C is warranted in assuming E. If this thesis is correct, can it also be true that A does not refer in a determinate way to any nonrepeatable causal sequence or causal network of which events a and b are members?

In order to go further than we already have in answering this question, it seems necessary to study what the grounds are that warrant one's assuming E if one believes warrantably that A and C.

Already we have established that the grounds cannot be truth functional. We noted in Chapter II that if the deducibility of E from A is truth functional then the ascription of indeterminacy in the reference of A leads to an absurdity.

And we have already concluded from this argument and other considerations that the thesis about indeterminacy

in reference to actions requires reformulation or rejection. But we can conclude more than this. We can also conclude that if one is warranted in assuming E whenever one believes that A and C, it cannot be in the standard sense that one is warranted in drawing inferences on the basis of truth functional rules applying to sentences. To show this we need one counterexample, and here it is.

We noted already that there may be circumstances where not only A is true but C is true (Jack intentionally raises his arm and his arm rises)--yet it is not the case that E is thereby true. For the arm's rising might occur in some 'deviant' way other than Jack's causing it as he intended to or as he thought himself to be doing at the time. There is the odd case where A may be true though E is false.

Of course, it seems awkward if not outright false to say something was done intentionally if it didn't happen according to plan; one might object that the circumstances of this odd case will cancel out someone's saying A, or render A false. Then the question that reappears is:

What grounds determine the success or failure of a trying or an attempt, such that the truth or falsity of sentences like A can be established? And once again we find ourselves tracking causal sequences of events, to account for the warrantability of asserting A.

The upshot here is that we need to consider exactly what there can be about correctly believing sentence A

that warrants one's also believing E is true. And we need to consider what conclusions can be drawn exclusively about actions (as distinct from other sorts of occurrence) from the fact that causal assumptions generally are warranted on non-truth functional grounds. These will be our concerns in the next chapter.

Conclusion to Chapter III

We have studied two standardly opposed views in the controversy about how to number the acts done intentionally by an agent in a single episode of his acting for his own reason. First we examined Davidson's characterization of singular action sentences as cases of existential generalization. We found no basis on his view to support the claim that a reason-giving statement cannot refer to a specifiable event sequence constituting an act done intentionally. Then we compared Goldman's construal of singular action sentences. We found that his identification of actions as properties bore the same fundamental constraints as Davidson's identification of actions with events. This seemed to be so since they both view the cause of a particular action to be particular mental events or attitudinal states.

We found that when a causalist action theorist asks how many actions a particular agent performed at one time, the issue may be best construed as an empirical matter of fact. We came to no conclusion about the conditions that

warrant an onlooker's drawing specific causal assumptions about an action based on what the agent gives as his reason for what he thinks he is doing. This matter will be the subject of the next chapter. There we will also continue to press for clarification of the claim that actions done intentionally are described in a vocabulary distinct from the terminology and grammar comprising descriptions of all sorts of causal sequences.

CHAPTER IV
GROUNDING CAUSAL ASSUMPTIONS IN THE
SEMANTICS OF SINGULAR ACTION
SENTENCES

Introduction

In Chapter I we noted that reason-giving statements do not entail causal descriptions of actions in the standard sense that 'entailment' gets defined by truth functional rules of inference. In Chapters II and III we searched for a way to clarify the sense in which the notion of 'indeterminate reference' might be applied to referring expressions and to sentences describing actions as a special class of occurrences among different sorts of events.

In this chapter we continue to pursue these two claims: (i) that there exist two vocabularies, one for describing actions and the other for describing occurrences of all kinds including occurrences of actions as members of causal sequences. And (ii) belief in a statement composed of terms from the first vocabulary warrants belief in a statement of the second category. Our first job will be to refine the first thesis about these "two vocabularies"¹

¹.Hornsby, Actions, p. 106.

and then we will consider the basis for making the second claim.

Underlying the view that there are two distinct ways of describing bodily movements may be the belief that there are two kinds of bodily movement.

Hornsby is one theorist who classes bodily movement into two types.¹ One type includes bodily movements that are done intentionally. The other includes all bodily movements that are incidental, inadvertent, accidental or can otherwise occur without anyone doing anything intentionally, for instance during sleep.²

The distinction between these two types of movement is apparent to Hornsby given considerations we have noted earlier: Sometimes arms rise without being moved intentionally; and sometimes agents try to move their arms and nothing happens to their arms. Hornsby believes that an episode of someone's raising his arm intentionally involves an internal event she calls a 'trying', while the other type of movement never involves this sort of internal event. Thus there is a difference between the type of causal sequence constituting these two types of bodily movement.

¹.Hornsby, pp. 3, 5.

².If you make a fist while sleeping in order to punch someone in your dream, I don't know whether this counts as a bodily movement done intentionally. It doesn't matter to this discussion. I do not touch the literature on 'basic' actions at all in this essay.

Hornsby stresses that some event sequences can be classified or described as belonging to both types of bodily movement. That is, a particular occurrence of an 'arm rising' is identical to an occurrence of 'arm raising' just in case the arm's movement was caused by an agent's trying to raise his arm. Then she warns: the fact that a token occurrence may be describable in two ways should not be confused with a conflation of the two ways of describing token occurrences.¹

Hornsby's formulation of the two vocabulary thesis is useful to focus on since she shares many assumptions with the other causalist theorists we are studying. She takes as her starting point certain shared intuitions about the meaning of key words and syntactic characteristics of act descriptions and sentences describing actions. Hornsby expects that an analysis of act descriptions will disclose nonlinguistic facts about the constitution of actions themselves.

She presents her views about acts and their descriptions as a "logical continuation" of Davidson's thesis that particular acts are identical to nonrepeatable events.²

¹.Hornsby, Actions, p. 5.

².Ibid., p. 20. Note that Davidson expressly rejects reliance upon either the meaning of particular verbs or upon "simple grammatical" patterns as fruitful in accounting for "the entailment patterns that seem peculiar to simple action sentences." In Actions and Events see his "Agency" p. 44, and "The Logical Form of Action Sentences" p. 142. With respect to her divergence from Davidson's "conception of the

Hornsby shares with Goldman some assumptions about the way that an agent's knowledge of his own actions contrasts with his knowledge about other agents' actions and of any agent's physiology including his own. And she shares with Searle a reliance on causalist models of visual perception in her analysis of acts done for reasons. We will not pursue all these themes here.

Instead we will focus on the questions that emerge when a causalist action theorist asserts the two claims set out initially--about two vocabularies for describing individual occurrences of action and about the relation between statements comprised of terms from these two vocabularies.

Suppose then that an agent's raising his arm intentionally is identical to his trying to raise his arm and being successful in his attempt. An agent's attempt is 'successful' just in case his attempt is part of whatever directly causes his arm to rise on a particular occasion, matching or fulfilling his expectations for how he would

subject", Hornsby's approach bears a notable similarity to Melden's method of analyzing action statements in two respects: (i) Melden argues that an act done intentionally is wrongly construed if it is regarded as identical to any event or causal sequence, and (ii) the meanings of key terms like 'action', 'willing' and 'intention' bear central importance in Melden's characterization of action sentences as a special class of description, contrasted with causal descriptions. But on the basis of his meaning analysis, Melden denies that reasons can be causes, and here Hornsby parts company with him.

do this arm raising **in case** he had any such expectations.

Thus it appears that in order to determine the success or the failure of a particular type of action occurring on a particular occasion, one must inspect the causal sequences of events occurring on that occasion. How does one know which of the particular events and bodily movements to count as evidence that an agent's attempt has been successful or has failed? How does one know which events are not causally related to the agent's attempt?

The agent's own reason-giving description can only go part of the way here, by telling what type of occurrence to look for as the desired or intended or anticipated effect. But this strategy will not determine whether a particular nonrepeatable occurrence counts as evidence of success or failure at what the agent was attempting or doing or trying to make happen. As already noted, often agents will not know whether a particular event counts among the effects of their attempted endeavours. An agent may endeavour to do something without registering a belief about the specific bodily movements constituting what he would regard as a happy outcome. In cases where the welcomed outcome is just a certain bodily movement like a raised arm, the agent may be ignorant about his own efficacy. Generally agents carry a 'general idea' of the sorts of thing they want or try to bring about. Less often is it the case that an agent will have a vivid expectation worked up about exactly what movements will be involved in

getting his legal left turn accomplished, for instance.

This attention to particular events counting as evidence of success or failure is not just a function of over-fastidiousness. Recall that our concern here is how reason-giving statements convey causal descriptions, if they do. Indeed, the explanatory adequacy of reason-giving statements has never been advertised as the sort of causal story from which broad generalizations can be concluded. It is uncontroversial to note that the explanation afforded by an agent's own reason-giving statement is a one-shot deal. So we are obliged to understand how an individual action, comprised of a nonrepeatable sequence, hooks up to the statement purported to explain it-- hence all this fuss about the particular causal chain constituting success or failure of an action on a particular occasion for a single agent. We will reserve for the next chapter a closer examination of what must be involved in identifying a particular nonrepeatable occurrence as a certain type of failed or successful action.

For now we will consider this generally received view (ii) that descriptions of what an agent has wanted or tried or intended to do bear some non-contingent relation like entailment to sentences describing what that agent **actually** has done.

Other theorists besides Hornsby recognize and label a linguistic connection of this sort: Goldman calls it a

"logical" connection between an "occurrent want" and an action.¹ Searle calls the connection both "intrinsic" and "internal", between an agent's causal intention and the act he performs because he has that intention.² Sellars calls such a connection "conceptual."³

Suppose these theorists are also correct to claim that there is a contingent relation between an agent's action-causing thought, e.g. his trying to raise his arm, and the particular effects that he brings about in his success or failure to do something intentionally, e.g. his arm's rising. Then it's not obvious why these theorists regard a description of the thought causing the effect as a description "entailing" a description of that effect.⁴

There is a certain degree to which it seems accurate to generalize, as Hornsby does, that all actions are described in terms of their effects, intended or realized. It may be that this view of actions as describable only in terms of some effect or other yields the supposition that

¹ Goldman, Theory of Human Action, p. 112.

² Searle, Intentionality, p. 11; "Intentionality and Its Place in Nature," Synthese 61 (1984): 3-16, p.4; IIA, p. 263.

³ Sellars, "Actions and Events," p. 194.

⁴ Hornsby, Actions, "On Some Causative Transitive Verbs" Appendix A, pp. 124-132.

a sentence describing an action somehow "entails" a sentence describing its effects.¹ This supposition needs a clearer formulation. In what sense can an action sentence entail a causal statement that is not its truth functional equivalent?

In section one we will try to tighten Hornsby's formulation of the two vocabularies that are not truth functionally equivalent ways of describing sequences of events. Then in section two we will demonstrate the failure to derive causal facts about actions from action sentences on truth functional grounds. In section three we will explore non-truth functional possibilities for warranting generally accepted assumptions about the causal features of actions done for agents' own reasons. We will do this by drawing a distinction between the conditions that establish the truth of a sentence and the conditions that warrant its assertion.

1. If one does a survey of the conditions warranting one's assent to 'arm raising' and assent to 'arm rising', one finds that there is no exceptionless connection between incidents where these two descriptive phrases apply. From this linguistic observation Hornsby concludes that there exist two type of bodily movement,² and that what it is to be an act done intentionally "must be different than what it is to be a [mere] bodily movement".³ Sometimes, of course,

¹.Hornsby, Actions, p. 124.

².Ibid., p. 5. ³.Ibid.

the two descriptions apply to the same nonrepeatable occurrence of a limb moving, that is, just in case the movement is caused by an agent's trying to move his own limb. Hornsby argues that the agent's trying to move his limb is identical with his doing so intentionally and successfully, just in case the limb's movement follows as a causal effect of the internal event called the agent's 'trying' or his acting intentionally.

Let's suppose actions done intentionally are causal sequences beginning inside the body. And suppose that the acting 'itself' or the 'trying' member of this sequence must occur causally prior to the muscle contractions and other changes that in turn cause the limb movements usually executed whenever anything is done intentionally by a human agent.

According to Hornsby the difficulty in locating precisely where Jack's action is inside his body is not due to a lack of technical sophistication: Through neurochemical tracings on a screen, for instance, a neurophysiologist might have access to

a change of physiologically articulated causally related events which culminates in a bodily movement that occurred . . . only because someone did something.¹

In this way Hornsby suggests we can imagine learning the causes of actions with refined technology. But establishing its cause would not resolve the question of "exactly where

¹•Hornsby, Actions, p. 105.

in the chain to find the action itself . . . occurring after its cause."¹

Hornsby warns against confusing the agent's "causing things and the things he caused."²

The neurophysiologist can only become more adept at observing and describing

details about the events we cause to happen, about the immediate effects of our actions . . . for instance records of it on a screen wired up to the brain.³

Hornsby says there is no reason to presume that what the neurophysiologist "singles out is the same as what we single out using our everyday vocabulary."⁴

Why is this? Why can't the neurophysiologist describe the action anyone else describes and just as explicitly if he can describe its cause and can observe its effects?

Hornsby contrasts the neurochemical terminology used to describe the cause and certain of the effects of an act done intentionally with "everyday talk" about actions. "Everyday talk" involves cognates of the terms 'intention', 'motivation', 'desire' along with other key words. In some predicates the "everyday" verbs that name different kinds of action occur in their transitive

¹.Hornsby, Actions, p. 105. Note that in one passage it appears that actions may be some sort of relation or a function of other sorts of occurrence: "Actions will be located between mental states and the events of which the subject of those states is the agent;. . . actions happen where desires are translated into desired effects." p. 123.

².Ibid., p. 106. ³.Ibid. ⁴.Ibid.

form with the agents as the subjects of the sentences describing actions: 'Jack is making a legal turn'; 'Jack is sabotaging the bridge'; 'Jack is going to raise his arm'. Hornsby says such sentences reflect a "concept of action" that may not be

invested with such determinacy that we can draw any sharp line . . . between causing things and the things caused.¹

These considerations do not explain so much as they restate and underscore an appearance of contrast between "two vocabularies" composing descriptions of actions and descriptions of other kinds of occurrence including causal sequences. Since the appearance is one that many besides Hornsby recognize and acknowledge as being significant, it is worth asking what this apparent contrast amounts to.

Here may be a way of formulating more precisely Hornsby's contrast between sentences that describe bodily movements that may be counted as acts done intentionally, and sentences that describe movements that may only be the effects of acting intentionally.

For purposes of clarifying the distinction Hornsby has in mind consider the following sentences:

- (1) Jack intentionally moved his arm at 3:00;
- (2) Jack's arm moved at 3:00;
- (3) Jack successfully tried to move his arm at 3:00.

¹ Hornsby, Actions, p. 106.

We will examine in detail the sort of assumptions, expectations and beliefs generated by these sentences in a later section. For now, recall that Hornsby claims (1) and (3) on certain occasions may refer to precisely the same occurrence inside an agent's body, although that occurrence is not locatable and reference to it is always 'indirect', always couched in terms describing its effects. She argues that (2) must refer to another event caused by the one described, for instance, in sentence (1) and in (3).

Suppose 'NMS' labels a particular and nonrepeatable causal sequence of neuromuscular events and chemical changes of state occurring inside an agent. And let 'S' stand for a singular sentence that describes an occurrence of some type. Sentences (1), (2) and (3) can count as examples of S. Now consider the following relation R:

(R) The fact that NMS occurred at 3:00 entails that S is true.

Hornsby's claim about R is what I formulate as her "two vocabulary" thesis;¹ using the sample sentences (1)-(3) to state her claim:

(TV) Substituting (1) or (3) for S in R will not alter the truth value of R. But if for S one substitutes (2) then the truth value of R does change.

In her TV thesis Hornsby appears to be applying a standard rule for substitutivity of truth functionally

¹•Hornsby, Actions, p. 106.

equivalent component expressions in a truth functional compound statement. TV presupposes that the term 'entails' in R as a main connective links 'the fact that NMS' and 'S' in truth functional relation to each other.

But is R a truth functional compound statement? How could one go about determining whether TV is correct? By addressing the second of these questions one may see the difficulties involved in answering the first.

Suppose there is no evidence available to warrant an onlooker's asserting (1) or (3) other than evidence that also warrants his belief and assertion that (2) is true as well. For example, suppose that the onlooker is watching Jack drive and is monitoring the neural impulses in Jack's arm while Jack is driving.¹ Next suppose the only evidence that warrants this onlooker's believing (2) that Jack's arm is rising, also warrants his believing (1) that Jack is raising his arm, or (3) that Jack is trying to raise his arm. This evidence might be afforded the onlooker as he sees Jack approaching a corner and turning his car left through heavy traffic. Should the onlooker, with evidence warranting his belief that (1)-(3) are true, thus regard (1), (2) and (3) as substitutionally equivalent in R? This substitution of (1)-(3) would render TV false if it were tenable. Does

¹•Hornsby, Actions, p. 106.

this onlooker have evidence that TV is false?

One can argue that the answer here is no; for the conditions warranting one onlooker's assertion of (1), (2) and (3) can be the same although these sentences are not truth functionally equivalent. We saw this to be the case in Chapter II, where the example involved observable truths and identity statements about the planet Venus. In the present example, one may have evidence at 3:00 that warrants the assertion of any pair of sentences (1)-(3) without this evidence indicating whether the pair are truth functional equivalents. For an obvious example, compare evidence warranting the belief that Jack's arm is rising at 3:00 with evidence that he is raising his arm intentionally at 3:00. There are afternoons where the evidence supports one of these beliefs though no evidence occurs to support the other.

So far our considerations run contrary to Hornsby's TV claim: if the main connective term 'entails' in R is functioning as a standard truth functional connective, then it is false to say that (1) and (3) can be regarded as substitutionally equivalent in R, whatever may be the case about (2). Statement (1) may be logically equivalent to (3) but empirical evidence does not demonstrate that it is. Theory may propose or reveal the logical equivalence of (1) and (3), as Hornsby's theory of action attempts to do. But it is the tenability of this very theory that we have been considering here in its abbreviated form, the

TV thesis.

Suppose, on the other hand, 'entails' in R is being used in a non-standard way. Perhaps Hornsby's claim about the relation between two vocabularies is meant to be qualified by the limits of leading edge empirical observations. Suppose the onlooker is Jane who is reknowned for advanced neurophysiological investigation. Let us say Jane is warranted in asserting R on the basis of her accumulated observations of Jack and other agents under optimal conditions.

Now suppose we apply Jane's assertion of R in a given situation where the truth conditions for (1) and (3) are clearly discernible from the set of conditions in virtue of which (2) is true. To make such a case very plausible, suppose we revise (1) and (3), respectively, to: 'Jack made a legal turn at 3:00' and 'Jack successfully tried to make a legal turn at 3:00'; while (2) makes the same claim as before about Jack's arm rising at 3:00. At 3:00 Jane may be able to assert that the occurrence of NMS entails certain descriptions of Jack's movements and not others. But the assertability of R cannot show whether the converse of R is true--at 3:00 or under other conditions.

Jane's evidence cannot establish whether Jack's doing something intentionally or trying successfully to do something must ever or always entail (involve) a certain type of neuromuscular sequence that warranted Jane's asserting R on one occasion. That is, evidence that confirms R doesn't

reveal in what sense of 'entailment' sentences of one vocabulary are supposed to stand in relation to sentences of another vocabulary. And until the sense of entailment occurring in R is clarified, the TV thesis is undecidable or otiose.

How does one tell whether the term 'entails' in R is operating in a standard truth functional sense?

2. Let us suppose that 'the fact that NMS occurred' clause in R reports a causal fact. In this section we will canvass Hornsby's reasons and some independent intuitions for positing a special sense in which true action sentences share an 'entailment' relation with causal facts.

Hornsby notices the following linguistic generality holds true for a certain class of verbs that often occur in sentences reporting acts done intentionally: Language participants are propelled by certain conventions to reason from the sentence form:

$$aV_T b \text{ to } bV_I.^1$$

The pattern of reasoning here is not based on the notion of standard logical entailment, according to Hornsby. She stresses that 'John moved his arm intentionally' and 'John's arm moved' are not truth functionally equivalent. This is the substance of her "two vocabulary" thesis

¹.Hornsby, Actions, p. 2; the subscripts 'T' and 'I' stand for the transitive and intransitive forms of the verb 'V', respectively.

introduced in section one.

Nonetheless speakers are prepared to assent to 'The chocolate melted' if they believe 'Jack melted the chocolate'; and they are prepared to assent to 'The bridge was sabotaged' if they believe 'Jack sabotaged the bridge'. Hornsby observes further, that speakers are likely to assume it was Jack that caused the chocolate to melt and the bridge to be sabotaged, granted the speakers believe that Jack was melting and sabotaging intentionally on the occasions in question.

Obviously many sentences with transitive verbs carry no suggestion that an action has been done intentionally: bacteria break down glucose; the comet follows its trajectory; heat melts wax.

But note here that 'Heat caused the wax to melt' seems to be entailed by 'Heat melted the wax'. This is just what Hornsby observes the relation to be between Jack's melting the chocolate and his causing the chocolate to melt. No special form of agent causation is implied in her thesis here.² Hornsby sees an advantage in assimilating certain sentences describing actions into a broader class of sentences whose transitive verb constructions carry causal implications. Our job is to try for clarification of the nature of these causal implications or to find

¹.Again, no belief in a special sort of agent causation is being attributed to speakers here; see previous footnote in Chapter III, p.107.

².Hornsby, Actions, p. 101.

what the basis is for assuming such causal facts, since it is apparent that Hornsby cannot be talking about logical entailment.

Hornsby is not alone in recognizing the causal "implicature"¹ of certain phrases. One may be reminded of a distinction cast by Sellars in "Actions and Events" (1973) in terms of "functional and non-functional" descriptions.² Sellars contrasts the non-functional mode of description in the sentence 'The blister was caused by fire' with the functional subject term in 'The burn was caused by fire'.

The contrast here hinges on the difference between the meaning of 'burn' and 'blister' whereby 'burn' "carries causal commitments"³ that 'blister' does not.

If the project of this essay were cast on a broader scheme, we might ask the question: How does one tell which terms in a language (or specifically in our language) to count as the key components distinguishing 'functional' descriptions from other kinds?⁴ This global question cannot be pursued here, and so we pose a narrow version of it with respect to Hornsby's view of two classes of sentences represented by the pair introduced earlier in section one;

¹. This use of the term I lifted from H.P. Grice, "Logic and Conversation," in Speech Acts, ed. Peter Cole and Jerry L. Morgan, Syntax and Semantics Series no. 3 (New York: Academic Press, 1975) pp. 41-58.

². Sellars, "Actions and Events," p. 189. ³. Ibid.

⁴. I owe this question to Bill Child, in conversation at his rooms in All Souls College, Oxford, England, 5 January 1985.

recall

- (1) Jack moved his arm intentionally at 3:00;
- (2) Jack's arm moved at 3:00.

Hornsby regards the "surface form" of certain sentences as "some sort of guide to their logical form,"¹ By 'surface' grammar Hornsby refers to patterns of utterance that do not reflect unexceptional rules or law-like uniformities in speech and thought.

Hornsby observes, for instance, that generally if one is asked what Jack did one doesn't reply after the pattern of (2) but rather after the pattern of (1). She notes different connotations are evoked by the transitive and the intransitive forms of the main verbs in these sentences, as well as the occurrence of key terms like 'intentionally'.

A sentence describing an action seems to provoke in language participants certain associations, assumptions and expectations about the episode described. The sort of associations Hornsby has in mind are not unlike the ones we noted casually in Chapter I, when we were studying why the D-N model of explanation fails to account for the way reason-giving statements explain actions.

Thus, the following claims are apt to be believed more by one who believes sentence (1) above than by someone

¹ Hornsby, Actions, p. 8, n.5.

who discredits the applicability of (1) and believes (2) instead.

- (3) Jack is an agent.
- (4) Ordinarily Jack regards himself as able to move his arm.
- (5) Jack's beliefs, desires, or related attitudes can contribute causally to his making his arm move.
- (6) At 3:00 Jack had his own reason to move his arm.

Note again that none of statements (3)-(6) entail (1) any more than they entail (2). The inference likely to be drawn from sentence (1) to any of (3)-(6) is not based on their truth functional equivalence. It may not be correct to call the relation between (1) and any of (3)-(6) an 'inference'. But what is that relation?

Notice that assumption (5) above may be regarded as an informal correlate of a fact about Jack's NMS occurrences at a particular time. This is so if our considerations in Chapter III were tenable, when we were comparing Goldman's claims about attitudinal and act properties of agents with the truth conditions for Davidson's primary reason-giving statements, where both indicate the onset of certain attitudes and belief states inside an agent.

To the list of (3)-(6), then, one might be tempted to add

- (5') A certain belief and attitudinal state occurred in Jack, identifiable by the content of his thoughts about raising his arm, and this mental state or event causally contributed to his making his arm move.

Hornsby points out that the conviction a speaker has that (1) is true does seem likely to encourage him to hold these other beliefs (3)-(6) about Jack, whereas believing (2) provides no ground or incentive for believing any one of (3)-(6).

In the remainder of this section and in the following one, we will be trying to locate the basis for sentences like (1) conveying sentences like (3)-(6), including causal statements like (5) and (5').

Examples will show why it is not obvious how causal facts can be deduced or otherwise gleaned on the basis of the 'surface' grammatical pattern of an action sentence.

Consider that from

(6) Jack opened the car door by pressing the handle,
one is likely to assume or to expect that

(7) The car door opened.¹

But the assumption or expectation may be due in part to a lack of knowledge about the circumstances that account for the truth of (6). Consider for instance that on one occasion when (6) is true it may be correct to assume or expect that

(8) The car exploded.

Under such circumstances the tendency to suppose

¹•Hornsby, Actions, pp. 2, 124.

(7) on the basis of the grammatical form and the meaning of the terms composing (6) is otiose.

Perhaps like rhyming, the linguistic uniformities noted by Hornsby occur persistently yet independently of evidence that can warrant one's claim to know other facts about occurrences, including causal facts about acts done intentionally.

Perhaps the example I have selected goes too far along the causal chain of events constituting a reason and an action and its effects to capture the occurrences whose descriptions share a special sort of conceptual or intrinsic or internal or quasi-logical connection. For the example I selected contains descriptions of occurrences that do not read like acts done intentionally by the agent in the example. The car's exploding may not have occurred except as a result of the agent's doing something intentionally, though from his own point of view the car's exploding was something he did inadvertently--quite by surprise.

Perhaps we have to narrow the scope of our examination to sentences describing what is done intentionally, ignoring all contingent effects that may unexpectedly interfere or follow from something an agent does for his own reason.

We noted earlier in Chapter III that an agent's beliefs together with causal facts about an episode jointly determine the type of action that has occurred intentionally.

For example suppose Jack killed the captain partly because he falsely believed that the loaded gun in his own hand was a water pistol and that he was aiming at a life-sized paste-up. His killing the captain would be correctly called a case of manslaughter rather than a murder. Last time we used this example to show how causal facts about a particular occurrence can help determine the different sorts of act properties Goldman would correctly attribute to an agent at a specific time. Now we need to consider a different question: Does the converse relation hold between causal facts and act descriptions? That is, from a linguistic fact about the correct way to describe what Jack did intentionally can one deduce any causal facts about what he did (or about his doing it)?

On first sight it seems obvious that knowing the difference between the meanings of the terms 'manslaughter' and 'murder' cannot be sufficient to determine whether any episode of killing counts as a case of manslaughter or a murder. To establish the kind of killing that took place on any occasion one needs to investigate facts about the incident and facts about the agent involved. This ground was covered in Chapter III.

But it might be argued that gathering facts about the killing and about Jack--his whereabouts before the killing, how he happened to be pointing a gun at the captain and so on--is necessary in order to speculate about what description Jack himself was applying to the killing when

he pulled the trigger and shot the captain. It is just this suggestion that we will explore next.

Can it be tenable to suggest that the difference in kind between Jack's murdering the captain and Jack's committing manslaughter is wholly a linguistic matter 'relativized' to the agent at the time he acted?¹ Can we determine the type of action done intentionally that occurred solely by deducing from a sentence containing terms which Jack himself applied to his shooting the captain? Perhaps it would be correct then to claim more generally that for every occurrence of acting intentionally, there is some description--that of the agent's own viewpoint--that is crucial to determining what type of action it is.

In order to find out what Jack did and why, obviously it is not enough just to understand the sentences comprising Jack's story of the shooting. For obviously what we are interested in is not how Jack applies different terms to the different things he does. We want to know instead facts about his state of mind when he does the things he does intentionally. If Jack meant to kill the captain then, and if his shooting the captain wasn't in self-defense then it is correct and useful for our purposes to call his shooting the captain on that occasion a murder. Here it is less pertinent that Jack himself could use another word

¹. I am borrowing loosely here from Davidson's notion of truth relativized to a speaker and a time. See his "True to the Facts" (1969), Inquiries into Truth and Interpretation, pp. 43-44.

for what he did, granted his idiolect.

Of course, Jack's own word selection is crucial to establish precisely his intention and whether he succeeded or failed in his attempt to fulfill it; recall for instance Jack's sabotaging the bridge in order (he says) to undermine the communist-terrorist plot.

Much here seems to depend upon facts specific to each case. Perhaps this waffling about the role of the agent's chosen vocabulary in describing his own action could be abandoned if we stick to examples where the agent is neither deceived nor deceptive about what he is doing. Suppose Jack gives us his primary reason and describes the occurrence of his action accurately. In so doing does his statement entail causal facts about the particular episode in question?

When we query Jack about what he thinks he's doing we're looking for evidence of the particular thoughts whose occurrence caused him to make certain movements at the time he acted. Suppose we just focus on the meaning of certain terms and grammatical patterns that Jack correctly applies to a set of circumstances. Then we fail to locate what it is about those circumstances that indicates precisely which sequence constitutes Jack's particular thoughts and the subsequent events rendering his action a success or a failure.

The fact that it is Jack who describes an episode of his own action is not a guarantee that we will be able

to tell which events occurring at a certain time and place constitute his reason and his action, successful or failed.

Perhaps a better way of stating the problem at hand is to ask how it does happen that we are able to tell which event sequence to count as the one correctly called Jack's action done intentionally. It is apparent that language participants do make this connection; what is not obvious is how.

Our results from Chapter II suggested that speakers may establish which events to call Jack's action in the same way they determine which hostages are reported as being safe, or which of Jack's ties is described by a particular sentence as being lost.

Yet the problem seems particularly complicated for describing sequences as actions, in contrast with describing properties of things. For an agent may not know exactly all that has happened, though he is supposed to be an authority about some of what happened.

Jack may be presumed to know, for instance, that at this moment he wants to be irreverent and intends to shoot water at a cardboard paste-up of the captain. But we cannot guarantee that Jack knows any causal facts about his action. (If we could make this guarantee, there would be little or no trouble in fitting Jack's reason-giving statement to the D-N model; but there would also be no need for the model.)

Here we face the familiar problem of 'deviant' causal chains again: Jack may think he fired a harmless water pistol at a life-sized paste-up though in fact he fatally shot the captain with a real bullet. Obviously the course of events that took place has something to do with securing a correct description of what Jack did, just as the subsequent course of events has to do with determining whether his attempt was a success or a failure. And yet, knowing that a particular description is the correct one to apply to a particular episode does not, in and of itself, entail knowing what type of action was done intentionally. For instance: it is certainly correct, but unilluminating, to report that Jack shot the captain on this occasion.

Thus consider the question: What knowledge or beliefs warrant someone's saying that Jack meant to shoot the captain with a water pistol and that he blundered tragically? This question is not about what follows from the meaning of the description Jack gives to what he did. Rather, this question is about what makes Jack's statement--along with any statement describing what he did--true or false. But this must be shown.

Notice that no causal facts follow from the meaning of action sentences alone. Consider this example: One who believes the sentence

- (9) Jack shot the captain
- is likely to believe the sentence
- (10) The captain was shot.

But the correctness of these two beliefs does not help one to determine whether (9) describes what Jack did intentionally on an occasion when both (9) and (10) are true. For (9) alone does not indicate whether the captain was wounded or killed, nor by whose shot nor intent. This is what was meant earlier when the point was made that the meaning of words correctly describing a situation as a certain kind of action does not afford any indication of which particular sequence of events occurring in that 'situation' constitutes the act intentionally done, successful or failed.

That is, suppose one knows (9) and also claims correctly that

(11) Jack caused the captain to be shot.

Still, (11) alone does not indicate whose shot it was that Jack caused to hit the captain. Nor does (11) suggest what effect the shot had on the captain's overall well being.

Suppose linguistic convention dictated the following: in the absence of evidence to the contrary, if one accepts the report that Jack caused a shot to hit the captain as a true report, then one is prone to believe that Jack's shooting the captain was fatal. This is not in fact our convention, but it is easy to imagine that we might use certain cognates of 'shot' in some sentences as synonymous with 'killed'. That is, perhaps one who believes (11) is

likely to assume

(12) The captain was killed by Jack's shot.

Still one cannot infer from sentences (9)-(12) whether Jack's killing the captain was murder or accidental homicide.

In the reconstruction of Jack's thoughts about shooting the captain, causalist action theorists stress that the particular thoughts constituting Jack's causal state of mind leading to his pulling the trigger need to be identified. Davidson, Goldman, Hornsby and Searle each recommend a different label for the component of an event (or for the prior event, or for the concurrent mental state) that marks an event as being something done intentionally.¹ It is the occurrence of this type of thought that can determine whether Jack's pulling the trigger counts as murder or accidental homicide.

Yet to know the type of thought that is causal is different from knowing which particular thought in evidence at the time of an action's occurrence should be counted as one of this causal type.

We have noted before that it is not obvious what sort of evidence to look for to determine when the right sort of causal connection between thought and an event (like a bodily movement) has occurred, rendering that event

¹. We already have noted the different labels given to this special type of causal thought on each theorist's account; see Chapter I, section five, p. 28.

something that was done intentionally. Nor is it obvious how such evidence, if we had it, could select for us which particular thoughts to count as members of the causal sequence constituting the primary reason for the type of action done intentionally on that particular occasion.

We have been considering the thesis that certain sentences describing actions can provide such evidence. Yet in this section we have noted that the truth conditions of sentences in pairs like (1) and (2) cannot support the plausibility of inferring one sentence of the pair whenever the other is believed true (recall that (1) described an act; (2) described a causal episode).

So one may ask what it is that warrants an onlooker's assuming one causal fact rather than another about an act done intentionally. For instance what is it that warrants the obvious assumption that it was Jack himself--that is, Jack's own thoughts or attitudinal states--that caused the captain to be shot, and not some other agency, just in case it is believed that Jack did shoot the captain intentionally?

How do we isolate the particular thoughts an agent has at a specific time as being causal ones? In the next chapter we will explore the possibility of identifying action-causing thoughts in clusters rather than as discrete proposition-sized units.

3. It will be useful at this point to contrast the set of conditions that verifies the truth of a sentence with the conditions that warrant its sincere assertion. This distinction between truth and assertability of a sentence may serve to resolve the discrepancy we noted initially in Chapter I, and again in the preceding section. That discrepancy is between the truth functional entailments of a sentence *S* and assumptions that seem readily obvious to one who believes *S* and who does not believe that these assumptions are truth functionally equivalent to *S*.

To recap where we are, recall the pair of sentences (1) and (2), where (1) describes an act done intentionally and (2) describes the same episode in another (causal) way.

Suppose now there is a difference in the assertability of these two sentences by a single observer, on an occasion when they are both true. Devising such a set of conditions is very simple. Suppose Jane sees Jack's arm rise but fails to hear Jack when he says that he did it on purpose. Or she hears Jack say he is going to raise his arm now but her view is obstructed so she doesn't see his arm go up.

Is there anything that follows exclusively about actions from studying the contrast between the assertability of these two sentences given an occasion when both are true? Based on the contrast between truth and assertability of certain statements about what is being done intentionally, can we build a satisfactory model of how an agent's asserting

his own reason divulges causal facts about his action?

We are now in a position to search for an alternative to the D-N model (or for a substantive amendment to that model) to account for reason-giving statements. This proposal is based on the premise that there exists a mode of discourse--evidenced by certain patterns of statement and a distinct vocabulary--exclusively for talking about actions.

Again it will be useful to employ a provisional example involving no acts done intentionally, where we may discover what there is in the contrast between assertability and truth conditions generally. Then we can see whether ~~this~~ contrast is applicable to the peculiarities observed in statements about actions.

Consider some set of circumstances whereby all four of these sentences are true:

- (13) Lightning struck at 3:00;
- (14) I saw a flash of light at 3:00;
- (15) There occurred a flash of light at 3:00;
- (16) There occurred a flow of electrons at 3:00.

Note that although (13) and (14) are both true at 3:00 by virtue of the same nonrepeatable occurrence, some of the conditions that account for the truth of one do not account for the truth of the other; so their respective sets of truth conditions are different. For example, (14) presupposes that I am not totally blind, while the truth of sentence (13) clearly does not depend

on this condition.

In this vein also note that the conditions warranting someone's assertion of one of these sentences may differ from the set of conditions that warrants his asserting some other sentence on the list. For an obvious example, the assertion of (14) would be false if the speaker were totally and irreparably blind. Sentence (15), on the other hand, would be a warranted inference for a blind man to draw if he believed correctly (13) or (16) and if he had beliefs about the relation between light phenomena and electrical discharge.

A physicist might know certain covering laws regarding photon emission and electromagnetic radiation. Someone else might have the belief that such laws exist. In either case, the relevant assumptions might give someone "good grounds for believing"¹ that (15) is true when he believes that (13) or that (16) is true.

Suppose that generally whenever there is good reason for believing that (15) is true there is at least some reason for asserting (15) under those conditions. I hold tacitly this assumption about belief and assertion, though I should acknowledge that I am leaving untapped a vast literature on the subject of propositional attitudes associable with assertion under different conditions.

What this example is meant to show is that the

¹. Hempel, Aspects, p. 368.

set of assertability conditions for a sentence includes relevant facts about the speaker who asserts that sentence. Facts that are relevant about the speaker include what the speaker knows or believes regarding the subject matter of his assertion. But other facts about the speaker besides his beliefs may be relevant to his asserting a particular sentence warrantably.

For example, the blind physicist who hears thunder under certain conditions is likely to be warranted in asserting (15) but he is not thereby warranted in asserting (14). Someone else less informed than the physicist yet sighted might be warranted in asserting both (15) and (14) by virtue of the same conditions prevailing in the sky. Someone with no beliefs about electrons, for instance, might be warranted in asserting both (14) and (15) if weather conditions were right.

Now suppose one might argue that this observer is warranted in asserting that a flash occurred whenever he sees a flash, just because the truth of his report of seeing a flash entails that a flash occurred. (Note this entailment postulate is different than claiming one's report of seeing a flash entails that lightning has struck; this latter claim is obviously false.) One might then say that this observer with no knowledge of electrons is likely to infer (15) from the fact that he saw a flash occur, that is, from his belief that (14) is true.

What then of the sighted physicist who is warranted

in asserting both (14) and (15)? Recall that his blind colleague was warranted in asserting (15) but not (14), while the sighted physicist might be warranted in asserting both (15) and (14) just as the ignorant observer may be warranted in asserting both (15) and (14). How are we to judge what sort of ground warrants the sighted physicist's assertions?

It may seem obvious that (15), the statement reporting the occurrence of the flash, follows from (16) which reports the electron flow, together with statements about the relation of light phenomena and electrical discharges. Whereas (14), reporting someone's seeing the flash, obviously does not follow from the same set of statements. And so it may also seem obvious that if the ordinary layman knew as much about electrons and light as the physicist, the layman's belief that (15) might be something he could have inferred from his belief in (16) and related covering laws. As it is, our hypothesizing about the layman's inference is confined to the 'simpler' move in reasoning from what he observes, i.e. his belief in (14), since we are assuming he is ignorant of the relevant photon radiation theory.

But again, consider the case of the sighted physicist. It is not obvious how to tell whether his assertion of (15) is warranted on the basis of his knowledge of 'simple' entailment principles known to the layman,

in particular the principle¹ that 'whenever (14) then (15) also.' Perhaps the sighted physicist's warranted belief that (15) is based instead on his knowledge of photons and electrical discharge conjoined with his belief that (13), and quite independent of his belief that he saw a flash in the sky at 3:00.

These considerations suggest that some causal assumptions are warranted on non-truth functional grounds and that it is not clear which assumptions these are. In the case we drew up for inspection, it was not clear whether the sighted physicist arrived at his belief (15) as a logical inference based on other beliefs and knowledge he had that is applicable to lots of different occurrences supposedly involving electronic behavior. For his belief that (15) might be one that he became acquainted with 'directly', as the untrained observer might also become acquainted with this belief--perhaps just by virtue of having his eyes open under certain conditions.

Suppose the latter case holds for this physicist: he is sighted and catches a glimpse of the flash, thereby believing (15). Is it correct to say that he inferred (15) on the basis of other beliefs at all? Did the sighted physicist just assume or associate (15) as a matter of fact

¹My use of the term 'principle' in this sentence is lifted from Sellars, "Actions and Events," p. 188. Sellars posits that certain sentences represent "normative statements" a speaker learns from his "linguistic environment." These "principles" formulate observable uniformities and regularities in speech and thought, according to Sellars.

relevant to a particular experience he had on a single occasion of seeing a particular flash at 3:00, the sort of flash he has learned through language participation to call 'lightning'?

Generally speaking, it may be that assumptions based 'directly on experience' are not systematically distinguishable from assumptions based solely upon the truth functional entailments of an individual's accumulated set of beliefs. One may phrase this point in a slightly different way paved by Grice and Strawson: the sought after distinction is recognized well enough in a general sort of way, as evidenced through speakers' assent patterns to individual sentences, although no systematic or precise formulation of the distinction has yet been established.¹ Obviously much needs to be pursued in this suggestion that cannot be pursued here. But let us suppose this basic point carries: the distinction between belief acquisition via logical entailment and belief acquisition via direct acquaintance is not clear cut. We have so far found no reason to suppose that what appears to be generally true about speaking observers' assumptions does not also apply to the assumptions made by speakers about particular occurrences called acts done intentionally.

¹. I am alluding to the response of H.P. Grice and P.F. Strawson to Quine's analysis of the analytic/synthetic distinction, "In Defense of a Dogma," Philosophical Review 65 (April 1956): 141-158.

Conclusion to Chapter IV

Jane as an observer may readily accept certain statements as obvious since she assents to a report about what Jack did intentionally at 3:00. Nonetheless the statements she accepts may not all be truth functional entailments of facts that she believes about the episode in question.

In this chapter we have observed this to be the case as well for beliefs about occurrences that are not acts done intentionally. What we have found is that generally when assumptions are drawn about a reported occurrence, it is difficult to demarcate clearly which ones are drawn upon 'logical' grounds subject to an individual's acuity in applying truth functional rules to his own beliefs. Strictly 'logical' inferences are difficult to disassociate from assumptions based solely upon the individual's familiarity with contingent matters of fact.

We did find, though, that by distinguishing a set of 'assertability conditions' for a sentence from its set of truth conditions, it's possible to account for the discrepancy between logical entailments and certain causal assumptions that appear obvious to an individual who believes it is correct to describe an incident in a certain way.

But nothing unique about action discourse seems to follow from this result. In particular, we noted that it doesn't follow whether or not a particular action sentence makes reference to events in a way that is any less

determinate than other ways of describing occurrences. In general, we found the prospects discouraging for our original project of clarifying the basis for supposing there exist two distinct vocabularies that describe actions, having a special set of meaning entailments between them.

A related concern that was mentioned in this chapter has to do with identifying a particular causal sequence as a certain type of action that has succeeded or has failed. There appears a tension that highlights the need to refine or dissolve the view that act descriptions are indeterminate in the way they refer to nonrepeatable sequences. We noted that inspection of particular causal sequences partly determines types of actions as they occur and that this in turn wholly determines whether the specified attempts of agents have succeeded or have failed. In order to tell whether an agent has succeeded or has failed at what he intended, we need to have familiarity with particular facts about nonrepeatable events in causal sequence. We noted that these facts cannot be deduced from descriptions the agent offers of the episode in question.

We noted that even when an agent knows what he is doing and knows his primary reason, it is rare for the agent to focus his thoughts on particular nonrepeatable events constituting what he is doing. And yet it is commonplace to recognize that what the agent says or thinks he's doing plays a key role in the interpretation of event sequences

as acts done intentionally occurring on specific nonrepeatable occasions.

Thus it appears that this recalcitrant vagueness in reference to particular events separates reason-giving statements off from other sorts of causal explanation. Granted this vagueness of reference, it remains unclear how specific features of a nonrepeatable causal sequence can be deduced from an agent's own reason-giving statement about his own action. We pursue this tension in the next chapter, as a facet of 'radical interpretation.'

CHAPTER V

HOLISTIC APPROACHES TO INTERPRETING
CAUSAL SEQUENCES AS ACTIONSIntroduction

We have been exploring the idea that statements about actions can convey causal facts in their own distinctive way by virtue of the special vocabulary and grammatical form that marks them off from other ways of describing events. But we have found this suggestion provides no ground for building an alternative to the D-N model in order to show how an agent's reason-giving statement counts as a distinct species of causal explanation.

In this final chapter we will shift our attention from the language couching agents' thoughts about their own actions, to look instead at the causal aspect of these thoughts themselves. We continue of course on the assumption that an agent's thoughts are somehow causally connected to what he does intentionally for his own reasons. But it is by no means clear what this causal connection amounts to, nor is it clear how particular instances of the connection can be identified.

Recent attempts to provide an alternative to the D-N standard (or substantive amendments to this standard)

for causal explanation of particular actions might be dubbed 'holistic' models of the relation between reason and action. By using the adjective 'holistic' I mean to focus specifically upon two generally received theses in current action theory and natural language theory. These two theses have been forwarded by Davidson under the label of "holism."¹

One thesis of holism is that thoughts causing an occurrence of action can only be identified in clusters or pairs; the paradigm of this interdependence is the way belief and desire are supposed to function together. This was illustrated perfunctorily in Chapter I, section five.

Another way of stating this interrelatedness of distinct thoughts comprising a primary reason is to claim the following: An onlooker has to attribute batches of attitudes and belief to an agent in order to interpret any bodily movement or other event as a certain kind of action done intentionally.

An obvious parallel to this assumption is widely held in current theories of meaning: that is the interdependence between a speaker's beliefs and the meaning of his utterances. We will be pursuing some of the implications of this parallel in later sections four and five.

¹.An example of Davidson's use of this term that I have in mind appears throughout "Mental Events" (1970), especially p. 222. The quotations throughout this chapter are culled from Essays 9-12 in Inquiries into Truth and Interpretation, especially from "Radical Interpretation" (1973): 125-140; and "Thought and Talk" (1975): 155-170, (hereafter cited as Davidson, Truth and Interpretation).

But there is another presupposition found in current theories about talk and action that will need our attention first. The thesis is that you have to assume a speaking agent knows something or other about what he's saying or doing in order to interpret a sound string as his speech or an event as his action. In section one we will see why it's not obvious what the precise role of an agent's particular beliefs is supposed to be on a holistic account of the relation between reason and action. Likewise, in cases where the holistic approach has guided theories of meaning, it has not been established how to trace a distinct causal route back from a particular utterance to the specific set of attitudes or mental states that caused the speaker to make that particular statement at that particular time. We will consider whether holistic theories of interpreting sounds and events as speech and action can solve the recalcitrant problem of deviant causal chains. It was noted in Chapter I that this problem can act as a good test for alternative accounts of reason-giving explanation. For it is a major count against the standard D-N formulation that it permits reasons as explanations that are not primary reasons. Also it may count events as acts done intentionally that are not done in the way nor for the reason that the agent himself had in mind.

We also noted in Chapter I that the holistic model is faced with a tension that every account of reasons as causal explanations is faced with: that is the ambivalence

with which we treat an agent's own thoughts about what he is doing. In what follows we will find that the clusters of thought conveyed by a reason-giving statement are not transparent clusters; it is hard to tell exactly what in the mind of the agent constitutes the particular reason that caused him to act on a particular occasion. This is partly due to the fact that agents can act on erroneous beliefs and on the basis of propositional attitudes that an onlooker may not fathom. It may be that what an agent says and does underdetermines the different constellations of thought that an onlooker can reasonably hypothesize as the cause of what he observes the agent do and say. But these conjectures need to be followed through by studying examples in detail before they can lead us anywhere. A final preliminary conjecture follows here.

Perhaps we have been going about the formulation of the indeterminacy thesis in reference to actions in the wrong way. Maybe the indeterminacy characteristic of reference to actions has nothing to do with the way certain sentences refer to events identified as individual actions, nor with the descriptive content of an agent's beliefs about what he's doing. Rather, perhaps indeterminacy characterizes the very structure or nature of the mental state itself that constitutes a primary reason. Perhaps indeterminacy is encountered by trying to discriminate certain types of attitude or distinct kinds of thought as essential components or constants of a primary reason. This way of interpreting

the indeterminacy that seems endemic to reason explanations referring to particular sequences as actions done intentionally, may be suggested by certain problems incurred on the holistic models of Davidson and of Searle. In this chapter we will examine these specific problems.

1. In sections one through three we will contrast three views about the structure of an intention. By 'structure' here I mean the component thoughts or propositional mental attitudes constituting an intention, or whatever type of thought is construed as linking the desires of an agent (and his other motivating attitudes) to the occurrence of his action.

Traditionally action theorists have regarded desire as one type of attitude in a whole class of thoughts that can be part of what causes agents to do things intentionally. But the description of a particular desire, obligation, resolution or need may not indicate that any particular event called an action is even likely to occur. Suppose I have a desire to eat ice cream or to stop the war. There's really no telling what I will do at a particular time and place because I have such desires: will I eat one scoop of vanilla or two of chocolate or will I chew gum instead? Will I sit in front of a government building or write a letter to a congressman? Which of my bodily movements are the ones caused by the presence or the functioning of these particular desires in my system at a given time?

So much depends upon other contingent facts about each case where some movement of mine counts as a causal effect of one of my functional desires.

In this respect, having a desire or other mental attitude seems quite a different sort of attribute from being soluble, for example. The contrast suggested here is a compelling one to pursue in detail, though I can't do so without researching the literature on desire. Let us settle with this one superficial consideration: Granted one believes salt is soluble, it doesn't seem as though contingent facts will affect one's predictions about what salt will do in water on individual occasions. At least, contingent facts don't seem to intervene in the same way or to the same degree for the case of predicting what the salt will do in the water today, as they do affect the case of predicting what Jane will do in the water today.

Can the holistic approach help us account for the difference in consideration of causal factors between these two cases?

In their respective efforts to analyze the way thoughts causally connect to actions, Davidson and Sellars claim that beliefs play an essential role in the structure of an intention that bears fruit in acting intentionally. But they envision quite differently the sort of belief that is crucial to acting for one's own reason. We will compare their analyses first.

Then we will examine Searle's way of construing intentions that can cause actions. Searle doesn't think that propositional beliefs are uneliminable components of causal intentions. He posits a different sort of mental awareness to do the job of linking an agent's own reason with the particular thing he does because he has that reason.

It will be important to notice that despite the difference among these three approaches to analyzing the structure of an intention, each one presupposes an agent has knowledge about what he is doing. We noted this in passing about Searle and Davidson at the close of Chapter II, when we were judging the status of the claim about the indeterminate way that act descriptions refer to individual sequences.

I want to study whether one must know something in order to act for one's own reason. But the question is too difficult to tackle here. Instead we will take it as given that knowledge of some sort is a requisite feature of all practical reasoning. Then we can see whether this widely received tenet helps us understand how agents' own reasons can be causes of particular actions.

In section five of Chapter II, it was noted that Davidson relies on membership of at least one factual

belief state in the causal sequence of thoughts leading up to an occurrence called an act done intentionally.

Suppose Davidson is correct that an agent's having a factual belief about an event marks the step in reasoning where his thought contents shift from the general terms of desire, need and approval to the particular terms designating a nonrepeatable occurrence of action. Then the content of his factual belief might be construed as having this pattern or structure: The subject clause of his belief content refers to a particular event and the predicate describes that event as having a characteristic or property approved or desired by the agent.

But false beliefs may be part of what causes an action. False beliefs may function differently than Davidson depicts the role of factual beliefs in primary reasons. For it seems that a false belief cannot fulfill this pivotal function between generalized dispositions to act and particular occurrences of action. Here's why.

Suppose Jack blew up a bridge because he believed he would be pleasing his uncle and he wanted to please his uncle. And suppose Jack's belief is false: Jack has no uncle.

Then there is no event Jack could bring about intentionally or otherwise that would have the property of pleasing his uncle. Each of the events that Jack causes to happen (inadvertently, unwittingly, under duress, or intentionally) when he acts for this reason is an event that

would fail to have the desired uncle-pleasing characteristic since there is no uncle to please. So how can Jack's false belief refer to any event or sequence in particular as the one event or sequence identical to his action?

Of course it is by asking Jack what he thinks he's doing that one may learn the sort of event Jack would count as his action--as what he intended to be doing to please the uncle he thinks he has. But suppose Jack's attempt fails altogether without his knowing it.

Now notice how the problem with deducing a causal sequence from an act description is compounded by the fact that agents may be wrong about what they are doing in more than one respect.

For example suppose this sentence correctly describes what Jack did when he acted for his own reason at 3:00:

(Q) Jack inadvertently exploded the Golden Gate Bridge.

And suppose this sentence is one Jack himself would offer to describe what he thinks he is doing at 3:00:

(P) There is a bridge explosion in Nicaragua that I am bringing about and this event is pleasing my uncle.

Now in the course of failing his sabotage attempt, Jack may have brought about many different events. Along with the event described by Q; the San Andreas fault was weakened slightly; air molecules on the West Coast were moved, the United States' rationale for waging full scale war was inspired; a premature birth was induced. Which of these events count as the one identical to Jack's inadvertent action or his failed attempt at sabotage?

As noted, one can ask Jack and learn that he believed P at a certain time. But one cannot deduce Q from P to the exclusion of other sentences describing these other events on the list of what he caused to happen at 3:00. Conversely: from the true sentence Q it is not clear how to deduce as a causal fact that Jack believed P. He may well have had other erroneous beliefs and they might have played a causal role in his acting intentionally at 3:00. Perhaps Jack believed that inducing premature births in California constitutes a successful sabotage anywhere in the world and that the exploding of the bridge in Nicaragua was just a handy means to his desired end.

A more classic construal of the trouble encountered because of agents' ignorance is illustrated by the next example. Suppose Jack inadvertently succeeded in pleasing a man he never heard of who happens to be his uncle defected to the enemy's camp. Suppose the uncle is pleased by a report that his nephew's sabotage backfired. Jack caused his uncle's pleasure but the causal sequence does not constitute what Jack intended nor believed himself to be doing.

These examples show that one cannot deduce from an agent's beliefs all the facts relevant to explaining how his reason causally related to what he has done intentionally on a specific occasion.

2. Sellars¹ regards it as fundamentally mistaken to treat a sentence representing the content of a "belief in matters of fact"² as the pivotal member in a set of reason-giving premises leading to a practical conclusion. Here 'practical conclusion' means an act terminating a causal sequence of "practical" thoughts.³

Here are some examples of Sellars' beliefs in matters of fact or "first level" thoughts: 'This bomb will explode this bridge'; 'Lo, there is the enemy's bridge'; and 'I see the enemy's bridge.'⁴ Sentences representing the content of first level thoughts divulge certain patterns of a language participant's "response to sensory stimulation" and "associations" reinforced by his "linguistic environment."⁵

Sellars contrasts the "function" of first level thoughts about the world with "higher level" thoughts about the relations between factual beliefs, desires and other propositional attitudes identified by assent to certain sentential representations of the world.⁶

Sellars says that by appreciating the way different linguistic functions operate together in one complex belief structure one can understand how an agent's reason can cause his action.

¹.These remarks are culled from Wilfrid Sellars, "Actions and Events," NOUS 7 (1973): 179-202.

².Ibid., p. 190. ³.Ibid., p. 187.

⁴.Ibid., p. 187-188. ⁵.Ibid. ⁶.Ibid., p.188.

Sellars contributes two key features to Davidson's account of primary reasons as causes: Sellars attributes to reasoning agents a special causal attitude he calls "acceptance of an entailment statement as a principle of inference."¹ A principle of inference is a relational thought about other thoughts, e.g. 'that-P entails that-Q', where P and Q stand for first level beliefs.

The mental state of acceptance of an inference principle is marked by thoughts occurring to an agent in a sequence at a particular time. Such a sequence may lead an agent to do something. Sellars calls this the distinctive "language departure" function of certain sentence contours, patterns or types. An example of a language departure type of sentence is 'I shall do _____ now.'²

Thus Sellars contrasts two functions of the term 'entails' here which we will return to in a moment. One function is when 'entails' operates as the main connective in an inference principle, and the other function is apparent when 'entails' occurs as a component in a first level thought about conditions in the world.

When Sellars speaks of the "function" or the "structure" of certain sentences he seems to be regarding a sentence in one sense as a representation of a distinct

¹. Sellars, "Actions and Events," p. 184.

². Ibid., p. 192.

type of thought. Any single occurrence of thought might be classified or typed by the sound pattern sometimes produced by the individual having that thought. Thoughts may also cause speakers to produce movements with other parts of their bodies besides the larynx. For example, the function of the sentence pattern 'I shall now raise my arm' can be established by observing the regularities in limb movement that participants in the same language exhibit when thoughts of this type occur to them.

This is roughshod, but it serves to show how the structure of a statement may represent the causal role played by a certain type of thought. According to Sellars, the formal pattern "that-P entails that I shall now do Q"¹ represents a type of thought that must be part of what causes an occurrence of doing Q intentionally for reason P. The sentence pattern just described has several paraphrases: 'doing Q because of P', 'Q on account of P' and 'Q for Jack's own reason that-P.'

To make sense of Sellars' amendment to Davidson's model of primary reasons it is crucial to appreciate the difference he posits between the propositional belief that 'P entails Q' and the attitude of "accepting" 'that-P entails that-Q' as a "principle of inference."² Here we make use of the two-fold function of 'entails' that Sellars

¹. Sellars, "Actions and Events," p. 184.

². Ibid., p. 183.

emphasizes.

One is warranted in claiming that Jack inferred Q from P just in case one is warranted in claiming that Jack is a language participant who both knows the meaning or function of the statement 'that-P entails that-Q' and who "accepts" that statement as a principle in some episode of his reasoning. Jack "accepts" this principle if his thinking that P on most occasions causes him subsequently to think that Q. By attributing this attitude of acceptance to an agent under certain conditions, Sellars distinguishes episodes of "merely associating" the thoughts of P and of Q, from episodes where someone is "reasoning" from P to Q.¹

Here is how the "acceptance" of a principle fits into Sellars' account of the relation between an agent's reason and his action when his reason causes his action. Someone who "accepts" the principle 'that-P entails that I shall now do Q' is someone whose thinking that P causes him to do Q.

By asking an agent what he is doing one can hypothesize that the agent has "accepted" certain entailments and not others. Suppose Jack says something that suggests he has the thought 'that I please my uncle entails that I shall sabotage the enemy's bridge now.' Then one is warranted in ruling out certain events as candidates for the type of act

¹. Sellars, "Actions and Events," p. 185.

that Jack did intentionally. The catastrophe in California, for example, would count as 'inadvertent'. This is so on the grounds that the consequent of Jack's accepted entailment principle:

'I shall sabotage the enemy's bridge now'
is not replaceable in Jack's thinking by:

'I shall blow up the Golden Gate now.'
We know this since Jack will reject the identity claim:
enemy's bridge = Golden Gate Bridge,
if we propose it to him.

Likewise, if we examine the antecedent of Jack's practical inference principle:

'I will please my uncle'
we can learn what Jack's reason was for inadvertently blowing up the Golden Gate. His attempt at enemy sabotage was caused by his thinking that he was pleasing his uncle.

We run into a familiar trouble here if we regard Jack's practical inference principle as a belief about the relation between other beliefs, as Sellars recommends we do. The trouble arises because Jack's principle of inference, on Sellars' construal, is a compound statement containing two sentential phrases flanking the 'entails' connective. Granted that our considerations in Chapters II and III were well founded, we may ask: How can any particular sequence of events be deduced from Jack's inference principle?

Sellars offers a solution to this apparent problem, that we borrowed in Chapter III in our discussion of the

purported existential generalization structure of singular action sentences (on Davidson's construal).

Sellars analyzes the component sentential phrases in the statement

'that-P entails that-I shall do Q'

as referring to particular "use occasions" of the simple or first level beliefs 'P' and 'I shall do Q now.'¹ Sellars says that these two sentential expressions are true if and only if their counterpart singular referring expressions succeed in designating specifiable states of affairs or events on dated occasions. So for example Jack's principle 'that I please my uncle entails that I sabotage the enemy's bridge' is true just in case

'pleasing my uncle' and 'sabotaging the enemy's bridge' both refer to a particular event sequence or state of affairs on a single occasion of their joint use.

The fact remains, however, that an agent may act on principles that are false or misleading. We may reserve a careful review of what can go wrong with "accepted" entailment statements until Appendix A. For it is not obvious how the principles an agent follows in his reasoning, even if true, can yield descriptions of particular causal sequences.

In brief summary, the problem we are left with given Sellars' model of the structure of causally functioning beliefs

¹. Sellars, "Actions and Events," p. 197.

is this: A false entailment statement may operate causally yet fail to yield the necessary reference to particular events to tell what an agent has done or whether he has succeeded or failed to do what he intended to be doing. On the other hand, an "accepted" entailment statement might be vacuously true so long as it both fails to refer to the thoughts that did cause the agent to act and fails to refer to any of the effects actually brought about by the agent's having those thoughts. (If 'that-P' and 'I am doing Q now' are both false then the statement 'that-P entails that I am doing Q now' in one sense is vacuously true.)

How do we tell 'a' reason for Jack to do something from the primary reason that did in fact cause him to act?

3. Though Searle is in agreement with Davidson and Sellars that an agent's having his own reason can cause his action, Searle denies that beliefs can fit into the causal structure of a primary reason. Searle claims that a certain kind of intention does cause action. But he says the "logical properties" of a causal intention must be more complex than figured on Davidson's and Sellars' different accounts of how beliefs function conjointly with pro-attitudes to comprise a primary reason.

Two features stand out on Searle's proposal for solving the problem of deviant causal chains that we have been using as a test for an adequate model of reason-giving causal explanation.

Searle characterizes an action done intentionally as a "composite event"¹ containing both the mental cause and some physical effect of acting intentionally. Here, cause and effect figure as two "aspects" or properties of a "composite entity."²

Correlatively Searle characterizes an agent as the sort of entity that can directly experience itself being in a causal state of "making something happen."³

To ease the flow of this discussion we will use these phrases interchangeably:

'experiencing oneself being in a causal state',

'being conscious of one's mental state', and

'being in a conscious state.'

This usage is probably too facile but we can't pursue here the clarification needed to learn whether there is a relevant sense in which the meanings of 'having an experience' and 'being in a conscious state' do not overlap.

Searle clearly recognizes that agents are neither conscious of everything they bring about nor are they always correct in what they think they have brought about. I may think I am causing my arm to rise a certain way though it doesn't move that way at all.⁴ This is a regular

1. John Searle, Intentionality, pp. 108, 125. These remarks also generate from Searle's article "The Intentionality of Intention and Action," cited previously as "Intentionality."

2. Ibid., p. 117; "Intentionality," p. 271.

3. Ibid., p. 123; "Intentionality," p. 262.

4. Ibid., Intentionality, p. 125.

occurrence among swimmers. Yet while lecturing or swimming I may raise my arm intentionally and not be conscious of doing so. Instead I may be thinking of the verbal point I want to make or the work I have to do when I get out of the pool.

But sometimes agents who do experience themselves in a state of making something happen are correct about what they think they are doing. Sometimes agents are conscious that they are causing other things to change and happen.

Presumably one can become aware of being in such a causal state by drawing one's immediate attention to what one is doing. Searle counts these cases as the relevant data for building a causal analysis of doing something intentionally.¹

In veridical or "successful" cases of causal experience, Searle says that "the experience itself does the causing."²

In action our experiences cause bodily movements and other physical events.³

Our question here is: how much can one know about a particular mental state or thought by being aware of its occurrence?

Suppose that if I draw attention to what I am doing

¹. Searle, Intentionality, p. 130.

². Ibid., p. 125; "Intentionality," p. 262.

³. Ibid., Intentionality, p. 125.

right now I can have an experience of making something happen. That is, I can gain the experience or become aware of being in a certain causal state. It is in this sense that Searle says an agent always "knows what he is doing," that is, he always knows "what he is trying to do."¹

Thus if I am having this experience of making my arm rise right now, then surely I cannot be wrong in thinking that right now I am having this thought or awareness of attempting to raise my arm or of making my arm rise in just this way. I can always claim this much certainty whether I actually succeed or fail in making the arm rise, and whether or not I construe this experience I'm having as one of succeeding or failing to make anything happen at all. I can know what it is I experience myself as trying to do or as attempting, without formulating a belief regarding what it is I am actually bringing about. Davidson makes this controversial point with the example of trying to type ten carbon copies.²

Thus from the fact that one is conscious of being in a causal state it does not seem to follow that one must be conscious of any particular effect or event that one's conscious state in fact may cause to occur. But then the following also seems to hold. By having the awareness or experience of being in a causal state right now, I can not

¹. Searle, "Intentionality," p. 263.

². Davidson, "Intending" (1978), Actions and Events, p. 92.

know whether the state that I am experiencing as causal is in fact a causal state. For I may be wrong in thinking that right now I am causing my arm to rise.

Searle's appeal to certain experiences of making things happen may indeed provide evidence that our "concept" of causation or "notion of causation" is accounted for by "primitive experience."¹ But appeal to a particular experience of causation, that is, appeal to the experience of oneself being in a specifiable causal state cannot serve as evidence that any particular conscious state is in fact a causal one. The most evidence that one can gain from the experience of making something happen is evidence that there exists some conscious mental state that may or may not also be a causal state. But this is just to say one is having an experience that may be wrong or misleading.

Searle meets an objection related to this one with respect to veridical perceptions. He says that my veridical perception of a car cannot count as evidence warranting my belief that the car is there. It cannot count as evidence, on Searle's view, since whenever I see a car veridically, my perceptual experience or awareness of the car just is my knowing that the car is there.² The fact that I sometimes hallucinate cars does not count against the fact that at other

¹. Searle, Intentionality, pp. 126, 132.

². Ibid., pp. 73-74, 130-131.

times I see cars. Sometimes I am right in thinking there is a car in front of me.

But the fact that I sometimes hallucinate does give credence to my wondering on some occasion when I think I see a car whether I am right on that occasion in also thinking that there is a car to see.

This latter concern is bothersome when considering how much Jack can know about what he is doing on a given occasion when he intends to be doing something, given his experience of acting at that time. The fallibility of agents' experiences, like the fallibility of their beliefs, leaves open the possibility that a 'deviant' causal chain can occur unbeknownst to the agent.

The problem here is not solved by calling the agent's own conscious state and its physical effect components or aspects of the same complex event. For an agent may be conscious of himself as doing something intentionally or of attempting to do something and yet be unsure of which event, if any, to count his "experience of acting" as a causal component or "causal aspect."¹

We noted before that whether or not I am conscious of having the particular thoughts I have as they occur to me, presumably I could become aware of such thoughts by drawing

¹ Searle, "Intentionality," pp. 257, 262, 273-277; Intentionality, pp. 102, 108, 125.

attention to what I am doing in a certain way. But this may not be the case for all my thoughts. Searle makes the claim that I may have beliefs of which I am not now conscious "and may never have thought of."¹ And perhaps there are certain desires or hopes whose object or content I cannot become aware of for one reason or another, though these pro-attitudes may now be causing me to do something intentionally of which I am conscious, and for a reason of which I am partially aware.

How can my being conscious of a certain mental state indicate whether or not my unconscious beliefs partly comprise my being in this conscious state?

More generally, how can my experience of being in a particular causal state inform me that the presence of this conscious state is related to my having certain propositional attitudes but not other attitudes about what I am aware of doing? If I am conscious of being in a causal state, does it follow that I must be conscious of every "causal aspect" of that state?²

I lack the research background in the study of consciousness needed to pursue these questions about knowledge of one's own conscious states.

Another reason Searle thinks that beliefs cannot be causal components of actions is because beliefs have the

¹. Searle, Intentionality, p. 3.

². Ibid., pp. 116-117.

wrong "logical properties."¹

On Searle's view each non-repeatable occurrence of thought has distinctive logical properties that determine the type of thought that it is. He calls these logical properties the thought's "conditions of satisfaction."² It's not obvious in what sense Searle is using the term 'satisfaction'. On my interpretation it seems to have something to do with the identity of a particular, nonrepeatable thought. But on my interpretation Searle seems to be stating an unhelpful truism about the identity of particulars that we addressed in Chapter III, when we were considering Hornsby's appeal to the 'necessary' features of a particular action as a criterion of identity. I have set my questions about "conditions of satisfaction" and also about Searle's notion of "direction of fit"³ into Appendix C.

It is also not clear to me how to construe the peculiar "logical connection"⁴ that Searle posits between the experience I have of making something happen and my own "specification" of what it is I think I am making happen.⁵ Searle says that only certain types of thought can share this special logical connection, due to

¹. Searle, "Intentionality," p. 262.

². Ibid., pp. 263-264, 256, 264, 271; Intentionality, pp. 101, 122.

³. Ibid., "Intentionality," p. 278.

⁴. Ibid., Intentionality, p. 121. ⁵. Ibid., p. 13.

the decisive "direction of fit" that catalogs thought into logical types.¹

Searle says this logical connection is "much weaker than" and "different from entailment between statements."²

Certainly Searle seems right in saying that there are things I experience myself doing or making happen that I can't formulate into sentential representations. For instance I am at a loss as to how truth functional expressions can represent exactly what I am making happen with my arms, torso and legs when I feel sure I am doing the butterfly stroke correctly; one finds this if one tries to teach the stroke.

But without resorting to rules that apply solely to truth functional expressions I am equally at a loss as to what sort of "logical or internal relation"³ can obtain between an experience of trying something and the experience of getting it just right. The failure to articulate the nature of this "internal"⁴ relation is no proof that such relation does not exist between thought and action. But here the problem we are wrestling with is just how to get from articulations of reasons to the identification of event sequences containing actions and the thoughts that caused them.

¹•See Appendix C.

²•Searle, Intentionality, pp. 121, 126.

³•Ibid., p. 121. ⁴•Ibid., p. 126.

Let us take stock of the preceding remarks overall in sections one through three, regarding the structure or "logical properties" of a causal intention. The following conclusion about beliefs related to actions may be the most plausible.

Searle seems to be right when he claims the logical structure of any sized propositional belief is too simple to accomodate the vast wealth of considerations that an agent may be aware of or think about when he does something for his own reason.

Perhaps no two causal intentional states would appear alike if the intricately detailed complex of considerations causing an action could be brought into focus just the way an agent experiences his own reasons, on reflection or at the moment he is taking action.

Yet from this consideration it's not clear how it follows that certain distinct types of thought must be or cannot be components of the reason causing an agent to do something intentionally.

4. Some action theorists assume there are important similarities between an episode of someone doing something intentionally and an episode of someone saying something.

To acknowledge such similarities one may not be committed to the view that saying something is an instance of doing something intentionally. It is controversial whether an adequate account of linguistic behavior can rely wholly upon a prior understanding of acting intentionally. For

instance Davidson observes that in order to make sense of "the attribution of finely discriminated intentions to an agent"¹ one must be able to interpret the agent's speech. So Davidson expects that any theory giving an adequate account of how we make "fine distinctions" between types of thought or attitude must also be a theory that gives an adequate account of how utterances are interpreted.²

At another place it would be valuable to explore the plausibility of applying primacy to either thought or to language when trying to understand the relation between them.

Here we are exploring whether Davidson's or Searle's models of the "interdependence of belief and meaning"³ have any bearing on the relation between an agent's reason and a particular causal sequence called his action done for that reason. Can the batch of thoughts causing an action be traced in the way that the meaning of an utterance is deciphered?

We will try adopting Davidson's theoretical strategy for "radical interpretation":

Consider a vast array of disparate, nonrepeatable sequences of bodily movements and facial expressions as an open set of 'uninterpreted' occurrences. Then ask how one might go about correctly describing certain ones of the occurrences as different kinds of action done for specifiable reasons

¹. Davidson, Truth and Interpretation, p. 127.

². Ibid., pp. 127, 164.

³. Ibid., p. 134.

the agent would regard as his 'own'.

This approach is suggested by Davidson's proposed general account of what we might know that would enable us to redescribe uninterpreted utterances as the right interpreted ones.¹

In Appendix B I outline Davidson's theory of preconditions for participation in a natural language.

Searle's holistic model of interpretation bears promising similarities for us. Searle regards any physical event brought about intentionally as something that can be construed as a speech act or as a nonverbal "utterance."² These could be marks made on a page or bodily movements or any other events. The key is that the occurrence must be construable as a representation that the agent intended to produce. That is to say, specific marks or movements or light beeps have meaning whenever an agent intentionally produces such occurrences with the auxiliary intention of representing something else by doing so.³

This thesis needs elaboration which we cannot pursue adequately here. One question about the agent's intending to use sound sequences to "express" his thoughts is raised at the end of Appendix C.

In any case, Searle claims that a causal "Background"

¹.Davidson, Truth and Interpretation, p. 126.

².Searle, Intentionality, p. 176.

³.Ibid., p. 169.

of thoughts¹ must be postulated in order to make sense of what an agent is doing intentionally or what he is saying. Hornsby refers to an agent's "know-how" in a way reminiscent of Searle's Background: an agent's skills and capacities that jointly determine the scope of things he can feasibly set himself out to do.² Propositional beliefs have no place in the causal Background of thought. Some of Searle's reasons for this point are also reviewed in Appendix C. More will be said about Background thoughts in section five of this chapter.

Davidson also claims that in order to interpret what someone is saying, one cannot attribute beliefs to that speaker in isolated, singular propositional units. One must hypothesize whole sets or batches of belief and other propositional attitudes to the speaker in order to make sense of anything he says. If Jack says earnestly and sincerely, 'Grass is green' it is likely that he believes grass is green; but this is not likely unless he also believes 'grass' and 'green' have certain meanings, and that this sentence as a whole is appropriate to assert under certain conditions and not others.

This interdependence hypothesis extends to all cases of interpreting verbal and nonverbal behavior that

¹. Searle, Intentionality, p. 143.

². Ibid., p. 151. Both Searle and Hornsby posit two kinds of knowledge, reminiscent in some ways of Gilbert Ryle's 'knowing how to' and 'knowing that' distinction in The Concept of Mind (New York: Barnes and Noble Books, 1949), pp. 27-61. See also Hornsby, Actions, pp. 83-84 and Searle, Intentionality, pp. 143-144.

exhibits preference, choice or decision, according to Davidson.¹

Making sense of what someone says or does also involves presuming "optimal agreement" between the interpreter's version of the facts and the sentences that the speaker generally appears to "accept" or to "hold true", according to Davidson.² In Appendix B we will take a closer look at this "optimal agreement" precondition and Davidson's reasons for postulating it.

Earlier in sections one through three of this chapter we concluded that the role of an agent's knowledge in the structure of action-causing thought is not well defined. In sections four and five we will look for clarification from parallel holistic models of sentence interpretation, where knowledge is also assumed to be necessary in order to produce interpretable speech.

Again the crucial test will be whether these models can provide a means of detecting 'deviancy' in a sequence of sounds or events, in order to avoid misinterpreting the sequence.

Let's suppose Jack utters the sound string 'There's a yellow rose' and Jane has evidence that he is making a sincere and also a correct avowal. It's not controversial to go ahead and assume that Jack is having

¹. Davidson, Truth and Interpretation, p. 163; Actions and Events, p. 99.

². Ibid., Truth and Interpretation, pp. 135, 161, 169.

a thought occur to him whose content is somehow represented by the sound string produced because he had that thought. Jane makes this assumption, and thereby understands Jack's statement, and to this limited extent she also understands Jack.

Which thought content is it that caused Jack to produce this particular sound string? This is the question Davidson asks in "Radical Interpretation" (1973).¹

Suppose Jane knows the truth conditions for 'There's a yellow rose' in English. If Jack is an English speaker sincerely avowing 'There's a yellow rose', then why can't Jane just read off the correct interpretation of Jack's utterance from the set of conditions that obtain whenever it is correct to hold this sentence true? Why can't the sentence count as the content of Jack's thought?

The "obvious objection" to this suggestion is that Jack may be making an incorrect avowal.²

Suppose Jane knows that in the vicinity there is no rose but there is a camelia. Why can't she assume that Jack incorrectly believes the nearby flower is a yellow rose and that this false belief partly caused him to make the incorrect assertion?

Jane can't assume this with warranted confidence because Jack can be wrong in ways that Jane may not fathom.

¹. In Truth and Interpretation, pp. 125-139; also see "Thought and Talk" (1975) in the same volume, pp. 155-170, and "Reply to Foster" (1976), pp. 171-179.

². Ibid., p. 136.

Jane's thinking that she knows the truth conditions for Jack's utterance can't help her discern which of the many possible conjunctions of wayward belief and meaning postulates causes Jack to report the presence of a yellow rose while looking straight at a camelia. Perhaps Jack believes that he is looking at a peonia and that all peonias shaped like camelias constitute a species of flower called 'yellowrose'.

But Jane could rule out many of these possibilities by further query and systematic presentation of white camelias, camelias in bud, roses of other colors and so on. This would require her to think she knows many sets of truth conditions from which she can judge where Jack's responses indicate he has gone wrong in his taxonomy of flowers.¹

Still the potential for ascribing false beliefs to Jack will be endless if all manner and degree of error is regarded as attributable to him. For all Jane knows, Jack may be attending to flower stages or to all yellow patches against any blue spatial expanse. Thus Davidson claims that Jane must adopt the strategy of optimizing agreement between her own version of the facts and the interpretation she assigns to sentences that Jack appears to hold true.²

¹. Davidson, Truth and Interpretation, p. 169.

². Ibid., p. 137.

. . . widespread agreement is the only possible background against which disputes can be interpreted . . . ¹

But now a different difficulty emerges in the attribution of beliefs using this strategy.

Suppose Jane's empirical generalizations yield that the conditions for asserting two sentences are usually interchangeable, though their truth conditions are different. An example might be the pair of open sentences 'x has a heart' and 'x has a kidney.'

Would Jane be correct in assuming that the same clusters of belief caused a speaker to assert either of these sentences at a time when they both happen to be true and she believes that they are true?

For instance in anatomy class suppose Jack fails to assent to or deny 'x has a heart' but he does exhibit the expected and appropriate uniformity in his response to 'x has a kidney.' This evidence indicates he knows from kidneys but has no anatomical knowledge about having a heart.

Suppose now that Jack utters both sentences during lab, under conditions where they both happen to be true. Jane should discern the difference between Jack's felicitously parroting 'x has a heart' and his correctly asserting 'x has a kidney', for she has the evidence from previous occasions that Jack is fudging in one case but not

¹. Davidson, Truth and Interpretation, p. 153.

in the other. But on the model we are considering now, the "right way" to interpret any utterance offers no distinction between correct assertion and fortuitous error or parroting. For according to this model, correct interpretation wholly depends upon "external circumstances"¹ together with presumed knowledge of the relevant facts on the part of the speaker and presumed optimal agreement between speaker's beliefs and the interpreter's beliefs. The model offers no way to mark off an aberrant case of taking a lucky guess from the ordinary run of cases where someone asserts something true because he has reason to believe it's true.

Cases of taking a lucky guess and thereby making a correct assertion, and cases where a speaker makes a certain amount of sense without knowing altogether what he is talking about or whether he's right, are commonplace in language participation. Participants can find each other out in such cases, so a distinction is made in the project of interpretation which Davidson's model of interpretation cannot seem to account for. Davidson's strategy of "optimizing agreement" seems vulnerable to these speech analogs of the classic "inadvertent success" story plaguing action theorists, who are likewise searching for a model to illustrate how we manage to track causal thought sequences correctly, back from the things agents say and do.

¹. Davidson, Truth and Interpretation, p. 125.

5. Searle argues that no appeal to representational thoughts like propositional beliefs can be sufficient to account for the occurrence of something that amounts essentially to being a representation itself.

In Appendix C we can cover some worries that emerge in trying to understand Searle's distinction of thoughts into "logical" types by use of the demarcation between "representational" and "nonrepresentational" contents.²

Without pursuing this matter, we can suppose that a belief with a propositional content is an obvious sort of representational thought. Searle shows that beliefs are either insufficient or unnecessary for interpreting an agent's doing something intentionally with sounds or limb movements.

Searle argues that factual beliefs do not really assist in the interpretation project as it is radically defined. In the case of Jack's uttering 'There's a yellow rose', treating the sentence as representing the content of a true belief that caused it does not help interpret the utterance. For now Jane is obliged to ask what is the right interpretation of the belief content. The propositional belief in itself gives no good reason to suspect that the causal connection between utterance and thought can be narrowed down in one way rather than some other.

¹•Searle, Intentionality, p. 148.

²•Ibid., p. 157.

To push the account in the direction that Searle has in mind for its completion, one must ask how it is that the agent had this belief or dispositional thought occur to him rather than some other, thus causing the utterance to be produced? Why was the agent disposed to "intentionally impose" certain constraints that make this sound string a certain type of sentence and speech act rather than another type?

Searle does not see anything gained by positing "principles of inference" to explain the causal connection. For the content of the principles of inference then beg for interpretation, since they are representations too.¹

Searle argues that even an "algorithmic procedure" for tracing the entire repertoire of Jack's speech dispositions will not address the question he is raising here about interpreting one of the utterances Jack does in fact produce at a certain time.² This is because no rules are "self-applying,"³ Searle argues that an infinite regress of practical principles would have to occur in a single episode of reasoning to get Jack to conclude anything in thought or deed.

We can display this point briefly. Searle challenges the appeal to principles of inference with the question, what is the relation that is revealed by the conjunction of these three representations:

¹. Searle, Intentionality, p. 153.

². Ibid., p. 148. ³. Ibid., p. 153.

'P' and 'Q' and 'that-P entails that-Q'?

One is obliged to construct a longer representation that orchestrates the relation of these three:

'Whenever 'P' and 'that-P entails that-Q' then 'Q'.'

But this new representation constitutes another pattern of "lifeless marks" standing for some other thought content in the projected system of beliefs.

Searle claims that 'nonrepresentational' mental capacities make rule-following possible. These are "social practices, biological capacities, stances and habits."¹ They are the sort of thoughts constituting the agent's mental Background.

Searle seems to be making an important point when he notes that many diverse phenomena (he illustrates three) are covered by the notion of 'rule following', and that it is wrong to take this bewildering array as a self-evident² or primitive component of agency. What is hard to follow about Searle's position is how he demarcates a special class of thoughts that can cause speech and action by virtue of its prior or primary functioning in relation to another class of thoughts.

Like Searle, Hornsby also posits a special "kind of knowledge"³ essential to acting intentionally, founded on the contrast between "knowing how" and "knowing that."

¹. Searle, Intentionality, pp. 144, 147.

². Arnold Koslow, in discussion of prospectus for this essay, CUNY Graduate Center, New York City, July 1982.

³. Hornsby, Actions, p. 83. See also above, p. 188.

On both of these holistic models, know-how or non-representational skills and capacities "enable" all other intentional states and causal thinking to function.¹ It is not obvious how a certain set of dispositions and capacities relegated to the 'background' of thought can claim primacy or priority over the causal capacity of other thoughts more accessible through language, like beliefs.

Here is why the background is supposed to be primary, in a crude sketch: Searle argues that representational, derivative thoughts like factual beliefs (Hornsby's "theoretical knowledge") are eliminable from accounts of many instances of acting intentionally. Background thoughts, however, must always be presumed to exist in agents who do things for their own reasons. Examples illustrate this.

For instance, Searle says you can ski without following rules in your head about how to ski, if you have the know-how. He says you ski very fast with the experts because you are skiing "in a different way altogether" than when you practice or learn to ski.²

He says practicing or learning to ski does involve rule-following via repetition of representations in your head that dictate what to do with your poles, your ankles, your weight distribution and so on. Whereas winning races involves the cessation of representational thoughts like these.

¹ Searle, Intentionality, p. 157.

² Ibid., p. 151.

Actually he suggests that the representations of body position may get replaced in the racer's mind by his thoughts of winning the race, representations of the prize money, the glory of triumph or the shame of losing, and so on.¹

Suppose this is so. If expertise constitutes an agent's doing an altogether different kind of thing than what he does to acquire his skill, then the anatomy of expertise may not be relevant to understanding what it is that makes rule-following possible in the course of training and pursuing excellence.²

Searle stresses that the nonrepresentative Background must be understood as a mental capacity peculiar to the individual agent, not as a feature of his environment beyond his skin and persona boundaries. That is to say, the causal antecedent of acting intentionally must be understood as originating in the mental life of the particular agent involved, not in the environment--physical or social. Jack acts for his 'own' reasons sometimes. This fact is most evident to Searle when agents' intentional endeavours "break down."

For example, if one intends to enter one's office and finds a chasm behind the door, it seems that one's intention to enter the office is frustrated because there is no office to enter. Searle points out that one formulates the intention to enter the office initially, on

¹. Searle, Intentionality, p. 151.

². Ibid., p. 157.

the basis of nondescriptive expectations that may be but need not be articulated in speech. Perhaps one's initial formulation of expectations about doors and what lies beyond them dates back to one's earliest encounters with doors as a nonspeaking toddler. In any case it is this nondescriptive stance or background expectation that is offended first and foremost by the sudden absence of one's office behind one's office door.

But surely the breakdown of one's current expectation originates in the current state of environmental affairs via the uncanny alteration of this bit of the world beyond the office door. And the affect of this breakdown is likely to be articulated in speech at the very first opportunity. It seems that such a change in the 'nonmental' world would disrupt both 'representational' and 'nonrepresentational' thinking alike. And such a gross disruption would surely engage the foreground of thought to the exclusion of much else, at least for the moments after one encounters this unexpected change in the world.

The example leaves one uncertain whether factors influencing the most "fundamental" stances and expectations of agents can be reserved as an exclusive feature of their "mental" life.¹

Searle argues that Background must be understood as an exclusively "mental phenomenon" since I would still have

¹ Searle, Intentionality, p. 142.

the thoughts that I have even if everything constituting the objects of my thoughts--the office door, the ski slope--was done up with holograms while my brain resided in a vat.¹ This illustrates that every one of my thoughts is essential to my having just the "mind-brain" that I have and not someone else's.

What the vat illustration doesn't show is which thought contents (and skills, capacities, etc.) I would be likely to produce or exhibit whether I was in the vat or out of it. So it's not clear from the illustration how to mark off into distinct types all the phenomena that my mind-brain apparently does produce, vat or no vat. The truism that my thought contents--in order to be mine--would be mine no matter what my changing circumstances were like, does not indicate how my thought contents come to be identified as "mental" vs. "physical" and "social" types of phenomena.

Searle's hypothesis of the Background seems to throw into clear relief the difficulty of demarcating the mental from the nonmental members of causal sequences constituting acts done for agents' own reasons. As compelling as this matter seems, we must leave it unexamined here.

The outstanding difficulty with lending a priority or primacy to nonrepresentational thoughts in the Background is this: How can we account for the success or the failure

¹. Searle, Intentionality, p. 154.

of a particular capacity to enable a certain desired outcome on a given occasion, unless one identifies (describes) that capacity with respect to the type of outcome desired or intended on the occasion in question?

To show the problem recall the antics on the ski slope. As 'uninterpreted' phenomena picture the flurry of snow, the different gyrations of the colorful skiers manipulating themselves down the slope at different speeds, and the slower movements they make gliding back up the slope on the ski lift. Which background capacities should be construed as causally connected to these phenomena and events?

Suppose one says that the whole batch of background capacities attributable to an agent is causally connected to one thing that a skier does on a slope. Then one's causal account may yield a result parallel to the unsatisfactory consequence of appealing to all the skier's propositional beliefs or practical principles of inference. The fact that I am capable of doing all the things I might intend to do on a ski slope does not explain what it is that I am in fact doing intentionally when I make one particular set of movements rather than some other set I am also capable of making. Nor does the whole set of capacities attributable to me explain why I am doing one thing rather than another. This of course was the problem that prompted Searle to invoke the background thought and capacities as nonrepresentational to begin with.

What to count as a causal capacity to act skillfully on the slope depends upon which effects are counted as evidence of skill on the slope--which effects count as successes or failures when skiing. Suppose skiing were all about staving off the inevitable dominance of gravity, such that falling down on one's skis midway on the slope is regarded as a partial compromise and relative victory when compared to the ultimate comedown of reaching the bottom of the hill first. Imagine the goal in this activity is reaching the top of the hill and staying near it the longest, like an inverse of scuba diving. What counts as success and failure requires reading the initial assumptions and expectations involved in producing certain effects called 'excellent skiing' in English. This much is obvious.

What is hard to see here is why Searle's solution involves the claim that representational thought is eliminable from the considerations necessary to interpret causal sequences in the right way.

For instance someone who doesn't know how to ski might accurately make out what's doing on the slope in virtue of his propositional beliefs or 'theoretical knowledge' that these people are in a race, that races require an attempt to win against competitive odds, that winning a ski race involves reaching the bottom of the hill first, whereby the ski lift counts as an incidental portion of the proceedings. Likewise, it seems that entering a ski robot into the race, one capable of winning on every

slope, somehow misses or spoils the point of producing the intended 60 mile an hour effect since the ski robot has no attitudes towards being a winner or loser, no complicating feelings about the risks or thrill involved. It is a classic difficulty to articulate precisely what point it is that competitive robots may miss and spoil.

In sum of these considerations, it's not clear why the know-how that constitutes the essential part of a causal network of occurrences called action must comprise mental states and events prior to other thoughts characterized by their sentential content. If indeed this is how acts done intentionally occur, then it's not clear how the specific capacities of an agent's mental background can be traced and identified as members of causal sequences on the basis of the agent's reason-giving 'representation' of what he's done.

Conclusion to Chapter V

We have been examining whether it is plausible to apply results from holistic models of interpretation theory to the problems incurred by tracing causal sequences of events that constitute an agent's reason and his action. First we looked at how the accepted models for describing thoughts leading to a reasoned action fail to designate a nonrepeatable causal sequence. We found that the agent's capacity to do things for his own reasons based on erroneous beliefs and principles complicates the account as it is

standardly proposed. That is, on standard accounts it is supposed that agents must know something about what they are doing in order to act for their own reasons. This supposition operates also in holistic models of action, despite the recognition that agents are regularly ignorant and fallible.

Since a parallel supposition obtains on theories of natural language interpretation, we next examined two holistic strategies for interpreting the meaning of utterances. We found that an analog to the problem of deviant causal chains (occurring in cases of action) also plagues theories of interpretation that presuppose a speaker knows when he is making a correct or appropriate statement every time he does so.

In the course of these considerations our formulation of the indeterminacy thesis about action has undergone a shift.

We started out hypothesizing that there must be something idiosyncratic about the way actions are described that makes their causal explanations so anomalous. We supposed that there was something characteristically vague in the class of referring expressions ordinarily occurring in action statements, including reason-giving statements offered by agents.

In previous chapters it emerged that act description seems to carry on in some fundamental ways like other sorts

of description, including causal descriptions of nonrepeatable occurrences. Despite this however, it seems that there remains an indeterminacy with respect to the causal network of states conveyed by reason-giving statements.

We have found no ground for positing a certain kind of thought or type of mental attitude as essential to every case of acting for one's own reason. This doesn't prove that there isn't any fundamental mental core of every intention, comprised of an exclusive cluster of propositional attitudes or other types of mental state. What we have found, however, is that one cannot deduce a specific set of attitudinal states or fixed types of thought from the descriptive content of an agent's reason for what he has done intentionally. Reasons may indeed cause actions, but it has yet to be shown how reason-giving statements designate which particular types of thought are operating causally whenever an agent's reason causes his action.

APPENDIX A

SELLARS' "PRINCIPLES OF INFERENCE"

FUNCTIONING IN REASONS

It would be incorrect to conflate Sellars' entailment principles functioning as normative statements with the "means-end beliefs" that Hornsby, Searle and others see as comprising standard models of practical reasoning. One crucial difference is that the conditional beliefs Hornsby speaks of invite an infinite regress of thoughts leading up to an action; whereas Sellars is careful to distinguish statements about statements from first level implication statements about matters of fact.¹ Nonetheless on many models of practical reasoning, sentences following the pattern 'that-P entails that-I do Q' are taken to represent in one way or another thoughts that are crucial to causing action. So it is worth seeing how causal facts about action are not readily deduced from sentences following this pattern.

An agent may assent to an entailment statement whose consequent fails to refer to any events that he caused to occur. For example, Jack thinks he has managed

¹ Sellars, "Actions and Events," p. 180. See Hornsby, *Actions*, pp. 79-80 and also Searle, *Intentionality*, p. 148 regarding this regress argument.

to destroy the enemy's bridge and is unaware that the Golden Gate exploded instead. His primary reason might still be formulated as the practical principle:

(E) That I please my uncle entails that I bomb the enemy's bridge now.

It isn't clear how E or its truth functional equivalent can refer to the particular sequence of events including the catastrophe that Jack unwittingly wrought, since Jack would assent to E and be surprised by news from California.

The same kind of problem recurs in the case of Jack's inadvertent success. Jack inadvertently made his treasonous uncle happy by virtue of his failing to sabotage the bridge in Nicaragua. The fact that Jack believed or accepted E caused him to act; yet this remains so whether the antecedent of E states a truth or a falsehood about the secretive uncle. Furthermore, from the fact that Jack accepts E it may follow that the whole of E states a truth or a falsehood. So knowing the principles that guide an agent's practical reasoning cannot determine which events to count as evidence that the agent's attempt has succeeded or failed, even if the agent has some of the initial facts straight (i.e., the antecedent of his accepted principle is true).

To show this, suppose as just stated that E is false, though its antecedent is true and Jack both accepts E as a practical inference principle and believes, correctly as it happens, that he is pleasing his uncle by what he is doing.

Now one can deduce from Jack's reason-giving statement (whose content can be boiled down to E) either that Jack did or that he did not blow up the bridge in Nicaragua. Here we see that knowing the agent's reason and knowing he has some of the facts about the outcome of his action--as is usually the case with most outcomes--tells us nothing about what the agent actually did do because he had that reason. Since one can't deduce the relevant causal facts about what occurred from Jack's beliefs about what occurred, one can't tell whether Jack's attempt succeeded or failed.

In the foregoing example, it is controversial to claim that Jack knew he made his uncle happy since his uncle was made happy and Jack believed he was making his uncle happy. But we have no resources here to pursue this controversy about the criterion for knowing that P. I think the example can survive without the controversial knowledge attribution, but I will forego finagling with it here.

Cases of unwittingly failed action may always occur. So how can the descriptive content of a principle guiding the agent's practical reasoning establish which causal sequence is the one being explained by Jack's reason-giving assertion of E?

On the one hand, consider a case where an agent accepts an entailment principle in his practical reasoning that happens to be true: He knows what he is actually doing

and he knows why.

Precisely what is it that the agent is supposed to know when he knows his own reason for doing Q? Sellars' model is ambiguous on this point.

Does Sellars' model imply that the agent's reason for doing Q is the fact that he had a particular attitude (desire, fear, hope, belief) toward the alleged fact that P? Does Jack's assenting to the statement 'that-P entails that-I shall do Q' imply that he also knew what type of attitude caused him to accept that statement as a principle of inference? To assume this might be to impute more self-knowledge to agents than they appear to have about their own motivations for adopting the reasons and principles by which they act. Yet perhaps this assumption is required if the contents of entailments accepted by an agent are to divulge the particular attitudinal states that caused him to act on a particular occasion. We have addressed this matter of the opaqueness of intention in later sections four and five of Chapter V.

Alternatively, perhaps Sellars' model implies that the matter of fact represented by 'that-P' in the agent's thought sequence is what caused him to do Q. This seems plausible until one recalls that the descriptive content of that-P might be false and still believed and acted upon by the agent. Or the agent might have been wrong about what he supposed P would entail on the particular occasion when he acted because he thought that-P and accepted principle E.

Again the problem recurs that acting on a false principle of entailment about a true belief can cause an agent to do Q, or to fail to do Q or to succeed inadvertently in doing Q. Yet if the entailment containing that-P as its true antecedent is a false entailment, how can it yield a correct description of the outcome that actually occurred when the agent thought he was doing Q because of P?

Another complication is even more provocative. Is it possible for an agent to believe or suspect or fear that-P¹ is erroneous yet act anyway "in accordance with"¹ the principle 'that-P entails that-I shall do Q'? This question invites us to study the causal role that principles play in episodes of doing something intentionally 'on blind faith' or 'irrationally'. Although such study may be crucial to understanding the relation between causal intentions and actions deemed 'rational', the question moves us beyond the resources of research done for this inquiry.

¹. Sellars, "Actions and Events," p. 188.

APPENDIX B

DAVIDSON'S CONDITIONS FOR "RADICAL INTERPRETATION"

In this appendix I try to spell out Davidson's view of how a language participant may discern the constellation of thoughts causally related to something someone else has said.

Suppose one had resources enough to compile the data of an entire speech community's utterances produced under diverse circumstances over a long period of time. Suppose one then classified all these tokens of utterance according to the sets of conditions, according to the greatest number of language participants' assent to distinct sound patterns. Suppose that the uniformities and generalities emerging from this way of summarizing one's data warranted one's supposing this: Each speaker favors consistency in the sound patterns that he produces under like circumstances over time; further, each speaker favors agreement between his own sound patterns and those uttered by other speakers under like circumstances. Suppose one calls the set of favored conditions for uttering a certain sound pattern the 'truth conditions' for the sentence type as well as for the belief content represented by that sentence type. One might abbreviate this summation of data by formulating

an open set of biconditionals called "T-sentences"¹ for the language. Let the left hand term of each biconditional name a type of sentence while the right hand term of each biconditional represents its corresponding set of 'truth conditions'. Call this entire open set of biconditionals or T-sentences one's "theory of truth" for the natural language one has been canvassing. A theory of truth for L is thus a conjunction of law-like generalizations about the assertability conditions for token utterances of different sentence types occurring in L.²

It probably needs to be discussed how to distinguish community-wide conventions that determine assertability conditions from truth conditions in cases where the two sets of conditions for a sentence appear to be congruent.³ But this matter leads far afield from our present concern with statements asserted just about actions.

Davidson hypothesizes that if one knew all the truth conditions for any sentence type that speakers of a natural language might utter, then one could interpret correctly any utterance of L that was produced under known conditions. Presumably one could do this because one's theory of truth for L indicates under what conditions

¹. Davidson, Truth and Interpretation, pp. 132, 134.

². Ibid., "Truth and Meaning" (1967), 17-36, p. 26, n.11.

³. No one is likely to assert 'X needs a heart' without also being ready to assert 'X needs a kidney' at our stage of medical technology. Yet we regard the truth conditions for these two assertions as distinct sets.

a speaker is likely to hold a type of utterance true. Suppose one assumes that a speaker asserts a sentence just when he holds it as true unless conditions strongly suggest he is lying or being facetious. Then a correct theory of truth for L can likewise identify the propositional beliefs and the meaning postulates that jointly caused a speaker of L to assert a particular sentence of L at a particular time, if his circumstances were known. More generally, if interpreting an utterance or nonverbal movement "must go hand in hand"¹ with identifying the thoughts that caused it, then knowing a theory of truth for L should suffice to identify an agent's causal attitudes at any time based on what he says and does under known conditions.

It should be noted that Davidson is not engaged here in an empirical reconstruction of the actual situations in which one finds natural language participants acquiring the joint skills of interpreting and producing meaningful utterances.² Clearly each one of us can be assumed to manage correct interpretations of some of the marks and sounds our English-speaking neighbors produce without it being presupposed that each of us knows what it would take for any sentence utterable in English to be true. Rather, Davidson is interested in spelling out what it is that makes natural language participation possible in the first place. Thus in "Radical Interpretation" (1973),

¹. Davidson, Truth and Interpretation, p. 163.

². Ibid., p. 125.

Davidson is engaged in the construction of an 'original position'-type of hypothesis for a speech community.¹ He is seeking the "necessary and sufficient" conditions² for producing representative sounds and marks that are correctly interpretable in only one way (or in specifiably distinct ways, for cases of ambiguity).

One feature of Davidson's "holistic constraint" that may give rise to questions is when he speaks of a theory of truth capable of yielding the right "location" of a sentence's place in the "logical or epistemic space" mapped by "all other sentences . . . in the language as a whole."³

I am somewhat vague about exactly how to find this location of a sentence in thought.

For instance I do understand that if Jack sincerely asserts 'I see a yellow rose' then ordinarily it is correct to assume that he believes he sees a yellow rose, in the absence of overriding evidence to the contrary.

But if Jack says correctly that he sees a yellow rose it doesn't follow that he believes he sees a rose, nor does it follow that he is disposed to say 'I see a rose.' For Jack may take it that all his yellow rose percepts are hallucinations or that they are evidence that there is a peonia in his vicinity.

Obviously we cannot pursue here the incongruities

¹. Truth and Interpretation, pp. 125-139.

². Ibid., p. 125. ³. Ibid., pp. 138, 139, 157, 158.

between the logical 'grid' regimented by entailment and the elusive design etched by an individual's system of beliefs. Nor are we prepared to pursue further the distinction between truth conditions of a sentence and its conditions of assertability when the latter standards are set by community-wide conventions of conversation. But it does seem important for our purpose to recognize that such a distinction may readily be made for some sentences and their so-called entailments.

APPENDIX C

SEARLE'S "LOGICAL PROPERTIES" DISTINGUISHING
TYPES OF THOUGHT AND ATTITUDINAL STATES

Searle claims that an agent's own awareness of what he is doing or making happen is a key indicator for determining the particular intentional state that causes him to act. Searle says the experience of being in a certain mental state

couldn't be that experience of awareness if it did not have just those conditions of satisfaction, and there is no way to have that experience without knowing it has those conditions of satisfaction.¹

This may be a way of saying that whenever I am in a particular conscious state like having a desire for ice cream, the particular circumstances for gratifying that desire will be unique and distinguishable from any other occurrence of the same type of attitude towards ice cream.

For instance suppose I have a desire for ice cream on the beach at noon; contrast this with another occurrence of my craving for ice cream in the bathroom at midnight. The following may strike someone who has never experienced any desire for ice cream as an obvious fact about my desire. It may appear as obvious that the conditions required to

¹. Searle, "Intentionality," p. 263.

satisfy the two distinctly described desires will differ in some ways. One difference that may seem obvious even to the observer unacquainted with the desire himself, is that the time and place where these two desires can be gratified looks different. It may be impossible to satisfy the first desire, for instance, anywhere but on a particular beach at noon.

But is such a constraint a "logical" feature of this particular desire for ice cream? Perhaps if you promise me at noon to give me a dish of Haagen-Dazs later tonight in the bathroom then my hankering on the beach at noon will be thoroughly quelled by your promising. Or perhaps a dip in the cold ocean will satisfy that craving for ice cream if only I am sufficiently determined to be rid of the desire at the time it hits me. Might it be the 'same' type of desire that recurs half an hour later? Or is any desire for ice cream that strikes after noon a different logical type than the desire striking at midday, since conditions for satisfying these two instances of desire appear to be distinct?

The conditions that determine the 'satisfaction' of specific desires and other related dispositional attitudes seem to depend upon empirical features of each case where agents act intentionally because they have these attitudes. It may be objected that these considerations are buffeted by the vagaries of desire--a particularly difficult attitude to assess for general characterization. For perhaps only the

individual agent knows these conditions for satisfying his desire, by simple virtue of his own experience of that desire in all its immediacy and vividness. Still, the factors determining what type of desire it is come out looking like empirical knowledge that the agent alone can have about himself--until he chooses to tell someone else what he is experiencing. It is unclear how the access an agent has to his desires through experiencing them as they occur can yield knowledge of their 'logical' type.

Generally I have a rough time deciding how to distinguish between types of mental attitude in a systematic way, though it is patently clear that such distinctions are made on a regular basis, whether systematically or no. Particularly I find it difficult to recognize how a distinct type of attitude can be identified as potentially causal or not, depending solely upon its "direction of fit."¹

Searle says that beliefs have absolutely the wrong direction of fit to play a causal role in action. Beliefs can play secondarily or indirectly into the cause of an action by directing the formation of secondary desires. Davidson agrees that this is how he comes to want a plane ticket, since initially he wants to be in London and believes that having

¹. Searle, "Intentionality," p. 278. The logical connection between intentional cause and physical effect is something that an agent can be aware of sometimes in his "experience of acting" according to Searle. See also p. 262; and Intentionality, pp. 65-66 where the connection is called "internal."

a plane ticket is part of getting himself to be in London. However, it is not clear how this direction of fit feature of a mental state distinguishes particular occurrences of thought into "logically" distinct types, only some of which can have a 'direct' causal aspect.

For example: much like causal intentions and contrary to beliefs, commands are supposed to have a "world-to-word" or "world-to-mind" direction of fit.¹

If the order is disobeyed it is not the order . . . which is at fault but the world in the person of the disobeyer of the order ²

This strikes me, with considerable urgency, as contingent upon the content of the command.

Contrarily, if I propose some phenomenon to you as empirical evidence that doesn't "fit" your belief in the law of gravitation or in your partner's fidelity, it is not at all obvious whether you will (or should) presume straightaway that your belief is "at fault" rather than blame my evidence for the disparity.

Another basis for distinguishing causal states or attitudes from those that cannot be causal is to demarcate "representational" content of thoughts like beliefs, from mental states that Searle calls "nonrepresentational." The distinction between representational and nonrepresentational thought that Searle has in mind is not obvious or clear cut. There is a controversy about which

¹. Searle, "Intentionality," p. 255.

². Searle, Intentionality, p. 7.

psychological modes or mental attitudes unlike belief can have content that is represented by marks or sounds in the way that belief contents are described.

Jack and Jane, for instance, may both hate war; and suppose it is correct to say that the object of their hatred is represented clearly by 'war'. Yet their attitudes toward war may be different significantly. Suppose it is correct to reflect this difference by contrasting Jack's hatred with Jane's using different propositional that-clauses to identify the object of their respective hatreds. Suppose Jack's hatred of war issues from the fact that wars are destructive and degrading to human life, whereas Jane's hatred of war focusses on the fact that wars rarely bolster real estate profits. Here someone might regard 'war' as being not the best choice for representing the content or object of these attitudes, since 'war' blurs even this crude distinction suggested by the two that-clauses above. Will other that-clauses reveal finer distinctions between the objects of these two attitudes? Perhaps the attitudes are not different--hatred prevails in each case. But what would decide when the mental attitude towards war is different between Jack and Jane other than a description of what it is they claim to hate?

Consider a different example. If it is a fact that Jack loves Jane, it might also be true that Jack loves that Jane exists. For surely if Jack loves Jane then he would feel mighty bad if she no longer existed. Yet perhaps if Jack loves Jane then he still loves her after

he learns who she really was, or even if she is a fiction pictured in a magazine. Are we considering one or more types of psychological mode or attitude dubbed 'love' in this example? Clearly, further inquiry is required before we can clarify the way a distinction between attitudes is established in light of the role description plays in deciding what sort of attitude someone has, and then in deciding its object or content. This work is only **part of the project** of deciding which attitudes have objects that can be represented sententially, and which attitudinal states cannot be represented sententially, or causally in other ways--like through bodily movements.

Searle models the identification of all sorts of "Intentional states" on his method of deciphering the type of speech act that a speaker is performing.

Let us say that an agent may produce a sound string with the intention of "imposing" certain functional constraints on that sound string.¹ He might be using the sound string to represent his belief that it is raining. Then the string 'It is raining' counts as the agent's "expressing" his own belief that it is raining.² Of course he might otherwise be using the sound string in a different "psychological mode" to lie, humiliate, joke,

¹.Searle, Intentionality, p. 169.

².Ibid.

confuse or practice English. The correct way to interpret any sound string depends in part upon

the conditions of satisfaction imposed on the physical event by the agent.¹

The same applies to interpreting an arm rising as a signal that the enemy approaches, or as a stretching.

This model leaves many questions open. It's not obvious that when I say something I can "intentionally impose" anything on the sound strings I produce since I'm usually unaware of the sounds I make as I talk.²

Searle seems right to treat it as a derivative concern what the phenomenon happens to be that an agent regards as his "expression" or representation of what he intends to say or do. I would prefer to examine more fully Searle's separation of one's intended representation from one's producing or "expressing" it, but this is a subject for a different essay.

¹.Searle, Intentionality, p. 169.

².David M. Rosenthal, comment in his seminar "Propositional Attitudes," CUNY Graduate Center, New York City, Spring 1983.

SOURCES CONSULTED

- Anscombe, G.E.M. "Under a Description." NOUS 13 (1979): 219-233.
- Armstrong, David Malet. "Acting and Trying," and "The Causal Theory of Mind." The Nature of Mind and Other Essays. Ithaca: Cornell University Press, 1980, 16-31; 68-88.
- Austin, John L. How to Do Things With Words. Edited by Marina Sbisa and J.O. Urmson. Cambridge: Harvard University Press, 1962.
- _____. "Performative-Constative." Philosophy and Ordinary Language. Edited by Charles Edwin Caton. Urbana: University of Illinois Press, 1963, 22-54.
- _____. Philosophical Papers. Edited by J.O Urmson and G.J. Warnock. New York: Oxford University Press, 1979.
- Beardsley, Monroe C. "Actions and Events: The Problem of Individuation." American Philosophical Quarterly 12 (October 1975): 263-276.
- Bennett, Daniel. "Action, Reason and Purpose." Journal of Philosophy 62 (February 1965): 85-95.
- Berofsky, Bernard. "Causality and General Laws." Journal of Philosophy 62 (December 1965): 148-157.
- Brand, Myles, and Walton, Douglas, eds. Action Theory: Proceedings of the Winnipeg Conference in Manitoba, Canada, May 1975. Dordrecht: Reidel, 1976.
- Brandt, Richard, and Kim, Jaegwom. "Wants as Explanations of Actions." Journal of Philosophy 60 (July 1963): 425-434.
- Brodbeck, May. "Meaning and Action." Philosophy of Science 30 (October 1963): 309-324.
- Chappell, V.C. "Comments on Davidson." Journal of Philosophy 60 (November 1963): 700-701.
- Child, Bill. Conversation at All Souls College, Oxford, England, 5 January 1985.

- Chisholm, Roderick. "The Agent as Cause." Action Theory. Edited by Myles Brand and Douglas Walton. op.cit. 199-211.
- _____. "Converse Intentional Properties." Journal of Philosophy 79 (October 1982): 537-545.
- _____. "The Descriptive Element in the Concept of Action." Journal of Philosophy 61 (October 1964): 613-624.
- _____. "Freedom and Action." Freedom and Determinism. Edited by Keith Lehrer. New York: Random House, 1966, 28-44.
- _____. "Sentences About Believing." Aristotelian Society Proceedings (February 1956): 125-148.
- _____. "The Structure of Intention." Journal of Philosophy 67 (October 1970): 633-647.
- _____. "What It is to Act on a Proposition." Analysis 22 (October 1961): 1-6.
- Clark, Romane. "Concerning the Logic of Predicate Modifiers." NOUS 4 (1970): 311-335.
- Collingwood, R.G. The Idea of History. New York: Oxford University Press, 1946.
- Davidson, Donald. Essays on Actions and Events. Oxford: Clarendon Press, 1980.
- _____. Inquiries into Truth and Interpretation. Oxford: Clarendon Press, 1984.
- Davis, Lawrence H. Theory of Action. Foundations of Philosophy Series. Englewood Cliffs: Prentice-Hall, 1979.
- Dennett, Daniel C. "Intentional Systems." Mind Design. Edited by John Haugeland. Montgomey, Vermont: Bradford Books, 1981, 220-242.
- Donagan, Alan. "Philosophical Progress and the Theory of Action." Proceedings of the American Philosophical Association. Presidential Address (24 April 1981): 22-53.
- _____. The Theory of Morality. Chicago: University of Chicago Press, 1977.
- Donnellan, Keith. "Knowing What I'm Doing." Journal of Philosophy 60 (July 1963): 401-409.

- Dray, William. "R.G. Collingwood on Reflective Thought." Journal of Philosophy 57 (1960): 157-163.
- Feigl, Herbert. "De Principiis Non Disputandum....?" Philosophical Analysis: A Collection of Essays. Edited by Max Black. Freeport, New York: Books for Libraries Press, 1950, 113-147.
- Frankfurt, Harry G. "The Problem of Action." American Philosophical Quarterly 15 (April 1978): 157-162.
- Gallagher, Kenneth T. "Some Recent Anglo-American Views on Perception." International Philosophical Quarterly 4 (1964): 122-141.
- Gardner, Patrick. "The 'Objects' of Historical Knowledge." Philosophy 27 (1952): 211-220.
- Goldman, Alvin. A Theory of Human Action. Princeton: Princeton University Press, 1970.
- Gorovitz, Samuel. "Causal Judgments and Causal Explanations." Journal of Philosophy 62 (December 1965): 695-711.
- Grice, H.P. "The Causal Theory of Perception." Aristotelian Society Proceedings Supplement 35 (1961): 121-152.
- _____. "Logic and Conversation." in Speech Acts, Syntax and Semantics Series no. 3. Edited by Peter Cole and Jerry L. Morgan. (New York: Academic Press, 1975) 41-58.
- _____. "Meaning." Philosophical Review 66 (1957): 377-388.
- _____, and Strawson, P.F. "In Defense of a Dogma." Philosophical Review 65 (April 1956): 141-158.
- Gustafson, Donald. "A Critical Survey of the Reasons vs. Causes Arguments in Recent Philosophy of Action." Metaphilosophy 4 (October 1973): 269-297.
- Hamlyn, D.W. "Analytic Truths." MIND 65 (1956): 359-367.
- _____. "Behavior." Philosophy 28 (1953): 132-145.
- _____. "Causality and Human Behavior." Aristotelian Society Proceedings Supplement 38 (1964): 120-142.
- _____. "Conditioning and Behavior." Explanation in the Behavioral Sciences. Edited by Robert Borger and Frank Cioffi. Cambridgeshire: Cambridge University Press, 1970, 139-166.

- Hamlyn, D.W. "Psychological Explanation and the Gestalt Hypothesis." MIND 63 (October 1951): 506-522.
- _____. Sensation and Perception. New York: Humanities Press, 1961.
- _____. "The Stream of Thought." Aristotelian Society Proceedings 30 (1957): 63-82.
- _____. "Unconscious Intentions." Philosophy 46 (1971): 12-22.
- _____. "The Visual Field and Perception." Aristotelian Society Proceedings 31 (July 1957): 107-150.
- Hampshire, Stuart. Freedom of the Individual. Princeton: Princeton University Press, 1965.
- _____, and Hart, H.L.A. "Decision, Intention and Certainty." MIND 67 (January 1958): 1-13.
- Hempel, Carl G. Aspects of Scientific Explanation. New York: Macmillan, 1965.
- _____. "Rational Action." Proceedings of the American Philosophical Association (1962): 5-24.
- Hornsby, Jennifer. Actions. London: Routledge, 1980.
- _____. "Actions and Identities." Analysis 39 (October 1975): 195-201.
- _____. "Arm Raising and Arm Rising." Philosophy 55 (January 1980): 73-84.
- _____. "Singular Terms in Contexts of Propositional Attitude." Mind 86 (January 1977): 31-48.
- _____. "Which Physical Events are Mental Events?" Aristotelian Society Proceedings 81 (December 1980): 195-201.
- Kenny, Anthony. Action, Emotion and Will. London: Routledge, 1963.
- Koslow, Arnold. In discussion at prospectus defense, CUNY Graduate Center, July 1982.
- Landesman, Charles. "Conduct and Rational Causation." Journal of Theory in Social Behavior 11 (1981): 241-252.

MacIntyre, Alasdair, and Nowell-Smith, P.H. "Purpose and Intelligent Action." Aristotelian Society Proceedings 34 (1960): 79-112.

_____, and Bell, D.R. "The Idea of Social Science." Aristotelian Society Proceedings 61 (1967): 94-132.

Malcolm, Norman. "Scientific Materialism and the Identity Theory." Dialogue 3 (1964): 115-125.

Melden, A.I. "Action." Philosophical Review 6 (October 1956): 523-541.

_____. Free Action. London: Routledge, 1961.

_____. "Reasons for Action and Matters of Fact." American Philosophical Association Proceedings 35 (October 1962): 45-60.

_____. "Thought and Its Objects." Philosophy of Science 7 (October 1940): 434-441.

_____. "Willing." Philosophical Review 69 (1960): 475-484.

O'Shaugnessey, Brian. "Trying (As the Mental 'Pineal Gland')." Journal of Philosophy 70 (July 1973): 365-386.

_____. "Human Action." Journal of Philosophy 60 (July 1963): 365-386.

Peacocke, Christopher. Holistic Explanation. New York: Oxford University Press, 1979.

_____. "Holistic Explanation: An Outline of a Theory." Rational Action. Edited by Ross Harrison, New York: Cambridge University Press, 1979, 61-74.

Prichard, H.A. Moral Obligation. New York: Oxford University Press, 1949.

Putnam, Hilary. "Psychological Properties." Philosophical Papers vol.1: Mathematics, Matter and Method. New York: Cambridge University Press, 1975, 313-322.

Quine, Willard Van Orman. Mathematical Logic. Cambridge: Harvard University Press, 1940.

_____. "Methodological Reflections on Current Linguistic Theory." Synthese 21 (1970): 386-398.

_____. "Quantifiers and Propositional Attitudes." Reference and Modality. Edited by Leonard Linsky. New York: Oxford University Press, 1971, 101-111.

- Quine, Willard Van Orman. "Reference and Modality." From A Logical Point of View. Cambridge: Harvard University Press, 1953, 139-159.
- _____. "Three Grades of Modal Involvement." The Ways of Paradox. Cambridge: Harvard University Press, 1966, 158-176.
- _____. Word and Object. Cambridge: MIT Press, 1960.
- Rosenthal, David M. "Keeping Matter in Mind." (typewritten) City University of New York, 13 July 1977.
- _____. "Mentality and Neutrality." Journal of Philosophy 73 (July 1976): 386-415.
- _____. Seminar on 'Propositional Attitudes,' The City University of New York, Spring term 1983.
- _____, and Sellars, Wilfrid. "The Rosenthal-Sellars Correspondence on Intentionality." Intentionality, Mind and Language. Edited by Ausonio Marras. Urbana: University of Illinois Press, 1972, 461-503.
- Ryle, Gilbert. Concept of Mind. New York: Barnes and Noble Books, 1949.
- Scheffler, Israel. "The New Dualism: Psychological and Physical Terms." Journal of Philosophy 47 (December 1950): 737-753.
- Searle, John. Intentionality. New York: Cambridge University Press, 1983.
- _____. "Intentionality and Its Place in Nature." Synthese, 61, (1984) 3-16.
- _____. "The Intentionality of Intention and Action." Inquiry 22 (1979): 253-280.
- _____. "Minds, Brains, and Programs." The Behavioral and Brain Sciences 3 (1980): 417-457.
- Sellars, Wilfrid. "Actions and Events." NOUS 7 (1973): 179-202.
- _____. "Counterfactuals, Dispositions and Causal Modalities." Minnesota Studies in the Philosophy of Science vol.2. Edited by Herbert Feigl, Grover Maxwell, and Michael Scriven. Minneapolis, University of Minnesota Press, 1958, 225-308.

- Sellars, Wilfrid. "Form and Content in Ethical Theory." The Lindley Lecture. Lawrence, Kansas: University of Kansas, 1967.
- _____. "Language as Thought and Communication." Philosophical and Phenomenological Review 29 (June 1969): 506-527.
- _____. "Obligation and Motivation." Readings in Ethical Theory. New York: Humanities Press, 1952, 511-517.
- _____. "Volitions Revisited." Action Theory. Edited by Miles Brand and Douglas Walton. op.cit. 47-66.
- Smart, J.J.C. "Causality and Human Behavior." Aristotelian Society Proceedings Supplement 38 (1964): 143-148.
- Suppes, Patrick. "The Philosophical Relevance of Decision Theory." Journal of Philosophy 58 (October 1961): 605-614.
- Thalberg, Irving. "Emotion and Thought." American Philosophical Quarterly 1 (January 1964): 45-55.
- _____. Enigmas of Agency. New York: Humanities Press, 1972.
- _____. Perception, Emotion and Will. New Haven: Yale University Press, 1977.
- _____. "The Socratic Paradox and Reasons for Actions." Theoria 32 (1966) 242-254.
- Thomson, Judith Jarvis. Acts and Other Events. Ithaca: Cornell University Press, 1977.
- Urmson, J.O. "Motives and Causes." Aristotelian Society Proceedings Supplement 26 (1952): 179-194.
- Vendler, Zeno. Linguistics in Philosophy. Ithaca: Cornell University Press, 1967.
- _____. Res Cogitans. Ithaca: Cornell University Press, 1972.
- Warnock, G.J. "Every Event Has a Cause." Journal of Philosophy 64 (November 1967): 704-713.
- Watson, A.J. "Comments on Hamlyn's 'Conditioning and Behavior'." Explanations in the Behavioral Sciences. Edited by Robert Borger and Frank Cioffi. op.cit. 153-166.

- Williams, Bernard. "Internal and External Reasons."
Rational Action. Edited by Ross Harrison. op.cit.
17-28.
- Yolton, John W. "Ascriptions, Descriptions, and Action
Sentences." Ethics 67 (1957): 307-316.
- Ziff, Paul. "About Behaviorism." Analysis 21 (1958): 132-136.
- _____. "On H.P. Grice's Account of Meaning." Analysis 28
(1967): 1-8.