In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of PHILosophY

The University of British Columbia
Vancouver, Canada

Date 27 September 1995
Abstract

This thesis is a critical examination of three influential and interrelated aspects of Donald Davidson's philosophy of action. The first issue that is considered is Davidson's account of the logical form of action-sentences. After assessing the argument in support of Davidson's account, and suggesting certain amendments to it, I show how this modified version of Davidson's account can be extended to provide for more complicated types of action-sentences. The second issue that is considered is Davidson's views concerning the individuation of actions; in particular, I examine Davidson's theory concerning the ontological implications of those sentences that assert that an agent did something by means of doing something else. The conclusion that I seek to establish in this case is essentially negative—that Davidson's theory is false. The third issue that is considered is Davidson's theory concerning the logical implications of those sentences that assert that an agent did something as a means of doing something else, which is also commonly known as the causal theory of action. Here I argue against Davidson's view by providing an alternative, and more satisfying response to the theoretical challenge that generates the causal theory. Subsequent to this I attempt to explain what motivates Davidson's commitment to the causal theory.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>ii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>iv</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter One:</td>
<td></td>
</tr>
<tr>
<td>The Idea of Logical Form</td>
<td>4</td>
</tr>
<tr>
<td>The Logical Form of Action-Sentences</td>
<td>12</td>
</tr>
<tr>
<td>The Logical Form of Attitude-Ascriptions</td>
<td>28</td>
</tr>
<tr>
<td>The Logical Form of Reason-Explanations</td>
<td>42</td>
</tr>
<tr>
<td>Chapter Two:</td>
<td></td>
</tr>
<tr>
<td>The Definition of Action</td>
<td>54</td>
</tr>
<tr>
<td>The Individuation of Action</td>
<td>75</td>
</tr>
<tr>
<td>Chapter Three:</td>
<td></td>
</tr>
<tr>
<td>The Definition of Intentional Action</td>
<td>99</td>
</tr>
<tr>
<td>The Characterization of Reasons for Acting</td>
<td>107</td>
</tr>
<tr>
<td>The Causal Analysis of Acting for a Reason</td>
<td>126</td>
</tr>
<tr>
<td>The Non-Causal Analysis of Acting for a Reason</td>
<td>138</td>
</tr>
<tr>
<td>The Analysis of Causality</td>
<td>153</td>
</tr>
<tr>
<td>Conclusion</td>
<td>164</td>
</tr>
<tr>
<td>Bibliography</td>
<td>170</td>
</tr>
<tr>
<td>Appendix to Chapter One</td>
<td>177</td>
</tr>
</tbody>
</table>
This thesis was written with, and would not have been written without, the financial support of a fellowship from the Social Sciences and Humanities Research Council of Canada. Crucial financial support also came from the Faculty of Graduate Studies at U.B.C., and from the estate of I.W. Killam.

A number of other people contributed to the production of this thesis in other ways. I would therefore like to acknowledge, and thank, my two thesis supervisors—Tom Patton and Bjørn Ramberg—for the professional advice and support that they provided throughout this project. I would also like to thank Alan Richardson for stepping in at the last moment to bring this project to completion.

Finally, I would like to thank all of those who, at one point or another along the way, provided some form of personal support. Included here are the members of my family, Janet Howey, Edward and Margaret Lang, Kirstie Lang, Marilyn Leonard, Michael McCready, Kate Talmage, and Rudy Vogt.
Introduction

Donald Davidson’s philosophy of action is embodied in a series of about a dozen, dense, interconnected articles that have been published over the course of the last three decades. Three of these articles stand out from the rest in terms of their importance to the discipline—"Actions, Reasons, and Causes," "Agency," and "The Logical Form of Action-Sentences." Corresponding to these three articles are three issues with respect to which Davidson has been influential—the explanation of action, the individuation of action, and the logical form of action-sentences. This thesis is a critical examination of Davidson’s contribution to each of these issues.

The first chapter in what follows is reserved for the issue of logical form. Davidson’s main contribution here has been to defend the idea that ordinary action-sentences like "Brutus stabbed Caesar" have the logical form of an existential generalization. After examining the justification that Davidson provides for this idea, and for his specific proposal for implementing this idea, I will argue that a modified version of Davidson’s proposal is better justified. Subsequent to this, we will consider the logical form of more complicated action-sentences, such as "Brutus intentionally stabbed Caesar," and "Brutus stabbed Caesar because he wanted to end the tyranny." The former sort of sentence is what one might call an ascription of intentional action, and the latter is what is called a reason-explanation. Davidson has not explicitly dealt with
the question of the logical form of reason-explanations, but I will show how his treatment of ascriptions of intentional action can be extended to accommodate reason-explanations as well. There are problems, in both cases, but I will argue that the proposed analysis of reason-explanations is no more problematic than Davidson's proposal for ascriptions of intentional action.

Once we are clear on the logical features of action-sentences we will be in a better position to sensibly discuss the main issue of the second chapter—the individuation of action. There are, in fact, a number of questions associated with the general issue of individuation; one is the question of what distinguishes one action from another, a second is the question of what distinguishes the class of actions from the class of events that are not actions, and yet a third is the question of the relation between two actions when an agent performs one by means of performing the other. Insofar as actions are events, the first question should be answered in terms of whatever criterion we use for distinguishing one event from another. In response to the second question, Davidson has advanced the idea that an event is an action if and only if there is a description under which it is intentional. And with respect to the third question, Davidson has defended the view that when an agent performs one action by means of performing another what appears to be two distinct actions is in fact one action under two different descriptions. After examining the justification that Davidson provides for each of these two claims, I will argue that both ought to be rejected.
The third chapter in what follows is dedicated to the central thesis that Davidson defends in "Actions, Reasons and Causes," which is not only his most influential work, but also one of the most important publications within the Analytic philosophy of action. The historical importance of that paper is due, in the first instance, to the fact that it upset a consensus that had already formed around a particular understanding of what it means to perform an action for a reason. Associated with this consensus was the view that reason-explanations are irreducibly different from the sort of causal explanations given in the physical sciences. In "Actions, Reasons, and Causes," Davidson argues that when an agent performs an action for a reason, the agent’s reason is the cause of his action. In advancing this claim, Davidson sought to show that the popular view was mistaken, and that there is an important respect in which reason-explanations are no different than causal explanations. Clearly, Davidson’s argument was successful in shifting the consensus, which has remained intact to the present day.

In the third chapter I will attempt to show that the central thesis of "Actions, Reasons, and Causes" is flawed, and that there is good reason to reject the claim that when an agent acts for a reason, the agent’s reason is the cause of his action. Correspondingly, I will argue that reason-explanations are, in the relevant sense, irreducibly different from causal explanations.
Students of logic are taught that an argument consists of premises and a conclusion, and that a valid argument is one in which the conclusion is true if the premises are true; they are also taught how to prove the validity of at least some valid arguments by recourse to a deductive theory. The first step in any such proof is to symbolize the premises and conclusion in such a way as to allow for the application of the rules of whatever deductive theory one uses. The logical theory that is typically taught today is divided into two distinct but related levels. The first level concerns arguments whose validity depends only the meaning of certain sentential connectives, such as "and," "or," "if," "only," and "not"; the second level concerns arguments whose validity depends essentially upon the meaning of certain quantificational terms, such as "some," "all," "any," "none," and "one." As far as the first level of logic is concerned, it is sufficient to treat the sentences that are joined by the sentential connectives as logically inert wholes; as far as the second level is concerned, it is necessary to find logical structure within these wholes. By way of illustration, consider the following argument:

A1 (1a) If the man in the brown hat is a spy, someone is a spy.  
(1b) The man in the brown hat is a spy.  
-----------------------------------------------  
(1c) Someone is a spy.
The argument in A1 is simple, and valid, and though the sentences within the argument contain quantificational terms, the validity of the argument does not depend upon the meaning of these terms. Thus, the argument in A1 can be symbolized as follows:

\[
\text{A1* (la*) If } P \text{ then } Q \\
\text{(lb*) } P \\
\text{(lc*) } Q
\]

Given this symbolization, A1 is revealed as an instance of the rule of \textit{modus ponens}, and the validity of A1 can be demonstrated by the sort of proof-procedures one is typically taught in sentential logic. In systems of natural deduction, \textit{modus ponens} is itself treated as a valid rule of the system.

On the other hand, consider the argument in A2, which is also clearly valid.

\[
\text{A2 (lb) The man in the brown hat is a spy.} \\
\text{(lc) Someone is a spy.}
\]

The conclusion in A2 is identical to the conclusion in A1, but in order to prove the validity of A2, it is necessary to find logical structure within the conclusion; in this case, the validity of the argument does depend upon the meaning of the quantificational terms. The solution that recommends itself is to interpret each of the sentences in A2 as consisting of the one-place predicate "x is a spy," where, in (lb), the argument-place is occupied by the singular term "the man in the brown hat," and, in (lc), it is occupied by a variable bound to an existential quantifier. Letting "Spy" stand for the predicate "x is a spy," (lb) can therefore be represented as follows:
(1b**) Spy(the man in the brown hat).

Correspondingly, (1c) can be symbolized thus:

(1c**) (∃x)(Spy(x)).

Given this symbolization, the argument in A2 is portrayed as an instance of the rule of existential generalization, and the validity of A2 is easily demonstrated by the proof-procedures that are typically taught in quantificational logic. Once again, in systems of natural deduction, existential generalization is itself usually treated as a valid rule of the system.

When we symbolize sentences of the natural language in terms of a syntax that allows for the application of the rules of a deductive theory, we credit those sentences of the natural language with a certain logical form. The logical form of a sentence is a theoretical construct at least in the sense that it is relative to a logical theory. For most of the last two millennia, sentences such as those in A2 would have been attributed a different logical form than that just given; correspondingly, the validity of the argument in A2 would have received a different explanation. Nevertheless, there are reasons for believing that the sort of theory that is typically taught today in quantificational logic is better than its Aristotelian predecessor.

The goal of any logical theory is to enable one to test and explain the validity of arguments given in the natural language in terms of a set of rules that apply to many other arguments as well. Thus, the most basic criteria for adjudicating amongst competing theories is the number or variety of arguments whose validity can
be explained, and the number or complexity of the rules that are employed in the explanations. The syllogistic theory that derives from Aristotle is inferior to what is taught today insofar as it contains more rules and explains the validity of fewer valid arguments. Thus, insofar as it is possible to adjudicate amongst competing logical theories, we can make sense of the idea that a certain sentence in the natural language, such as (1c), has a certain logical form, such as (1c**).

Any sentence in the natural language can play the role of either premise or conclusion in an indefinitely large number of valid arguments. In attributing logical form to a sentence one seeks to attribute whatever logical structure is needed in order to demonstrate the validity of the various valid arguments in which that sentence may be involved. This, clearly, is one reason for attributing logical form to a sentence. However, according to Davidson, the attribution of logical form to sentences in the natural language is guided by a second, and even more important, goal. Thus, he writes that:

What should we ask of an adequate account of the logical form of a sentence? Above all, I would say, such an account must lead us to see the semantic character of the sentence—its truth or falsity—as owed to how it is composed, by a finite number of applications of some finite number of devices that suffice for the language as a whole, out of elements drawn from a finite stock (the vocabulary) that suffices for the language as a whole. To see the sentence in this light is to see it in the light of a theory for its language, a theory that gives the form of every sentence in the language. A way to provide such a theory is by recursively characterizing a truth-predicate, along the lines suggested by Tarski. (Davidson 1968, p.94)

The foregoing passage gives the essence of the sort of theory of
meaning that Davidson seeks to provide for the natural language. According to Davidson, a theory of meaning for the natural language is a theory the knowledge of which enables one to interpret any arbitrary sentence within the language. As a minimum, Davidson believes that such a theory must be capable of recursively characterizing the truth-conditions of every meaningful sentence within the language. A theory that does this is what Davidson calls a *theory of truth* for the natural language. In giving the truth-conditions of the sentences in the natural language, such a theory will necessarily attribute logical form to those sentences. In addition to this, the theory will have clauses that effectively interpret all logical and non-logical constants, where the logical constants consist of those terms that are governed by the rules of our logical theory, and the non-logical constants are the remaining singular terms and predicates.

Davidson claims that "above all else," an account of the logical form of a sentence should enable us to see the sentence in terms of a recursive theory of truth, which gives the truth-conditions for every sentence in the natural language. However, this is debatable. In the first place, assuming that it is possible to provide a recursive theory of truth for the natural language, there is at least some question concerning the need for, or the value of, such a theory. Davidson claims that "by giving such a theory we demonstrate in a persuasive way that language, though it consists of an indefinitely large number of sentences, can be comprehended by a creature with finite powers" (Davidson 1968,
But what does this amount to? Surely we do not need any demonstration of the fact that human beings have the capacity to understand an indefinitely large number of sentences. On the other hand, it would be advantageous to have some account of what this capacity actually consists in—what practical or theoretical knowledge a competent language user actually possesses. However, Davidson does not claim that a recursive theory of truth provides an answer to this question, and he even suggests that it does not (Davidson 1973e, p.125)

There is, then, at least some question concerning the need for a recursive theory of truth for the natural language, but there is also a question concerning the very possibility of providing such a theory. It is well known that Alfred Tarski, from whom the idea of a recursive theory of truth derives, doubted that a theory of this sort could be provided for the natural language; the theory that Tarski originally provided was designed specifically for formal languages, which at least appear to differ from the natural language in a number of important respects. It is crucial to the success of Davidson’s theory of meaning, therefore, that Tarski be proved wrong in doubting the application of his work to the natural language. In what follows, we will consider and evaluate some of Davidson’s reasons for thinking that Tarski is wrong, but it is beyond the scope of this thesis to attempt to settle the matter. For the present purposes, we will accept Davidson’s assumption that it is possible to provide a recursive theory of truth for the natural language, and that this therefore constitutes a legitimate
goal of any theory of logical form. However, we must guard against thinking that this is the only goal of such a theory.

According to Davidson, there are many valid inferences in the natural language that cannot be explained by recourse to a theory of logical form. Thus, he writes that "x > y" entails "y < x" but not as a matter of form, and "x is a grandfather" entails "x is a father" but not as a matter of form' (Davidson 1969a, p.125). Assuming that entailment is a relationship that holds between sentences, we may assume that the variables that Davidson employs in these examples are symbols for singular terms, rather than symbols for argument-places. Thus, to give a clearer example of the sort of claim that Davidson appears to be making, one might say that "0 < 1" entails "1 > 0," but not in virtue of logical form.

However, this remark is misleading, as is the suggestion that "0 < 1" by itself entails "1 > 0." Rather, the former sentence entails the latter only in virtue of an additional premise which has been left implicit—namely, "(\forall x)(\forall y)((x < y) \rightarrow (y > x))." This premise, which partially defines the relation "smaller than," is undoubtedly true, and it is clear that its truth is essential to the validity of the above inference, which is only elliptically stated. Only in conjunction with the suppressed premise is it the case that "0 < 1" entails "1 > 0," but once this is made explicit, it makes no sense to say that the argument is valid, but not in virtue of logical form. For the only way of demonstrating the validity of the argument is by attributing logical form to the premise and conclusion.
Similar remarks apply to the other argument that Davidson elliptically describes, whose validity depends upon a suppressed premise defining the term "grandfather." Evidently, the reason why Davidson neglects the suppressed premises in the examples he gives is that a recursive theory of truth may neglect them. A recursive theory of truth for the natural language need not define predicates like "x is a grandfather"; rather, it is sufficient that such a theory give the extension of these predicates. However, for the purpose of demonstrating validity, it is a mistake to neglect these definitional clauses. And so here we see one way in which one can be misled by treating the task of providing a recursive theory of truth for the natural language as the only goal of a theory of logical form.

Having considered what it means to attribute logical form to a sentence, and the reasons for doing so, we will be concerned in the rest of this chapter with Davidson's views on the logical form of what he calls action-sentences, sentences such as "Eve ate an apple," "Caesar crossed the Rubicon," and "Oedipus intentionally struck the old man at the crossroads." In addition to, and subsequent to this, we will consider the logical form of sentences that are in some sense constructed out of these more elementary action-sentences, such as "Hamlet killed the man behind the arras because he wanted to avenge his father's death."
Consider the following sentence:

(2a) Brutus stabbed Caesar.

We observed in the last section that in order to explain the inference involved in A2 we must treat "the man in the brown hat" in (1b) as a singular term. Similarly, the inference from (2a) to (2b), and from (2a) to (2c), demonstrate that "Brutus" and "Caesar" also function as singular terms.

(2b) Someone stabbed Caesar.

(2c) Brutus stabbed someone.

Since there are no other singular terms in (2a), it is tempting to suppose that the rest of the sentence consists of a two-place predicate "x stabbed y." Letting the term "Stabbed" stand for this predicate, let us tentatively propose the following account of the logical form of (2a):

(2a*) Stabbed(Brutus, Caesar).

According to Davidson, there is a very good reason for resisting this analysis, which can be brought out indirectly by asking for the logical form of (2d).

(2d) Brutus stabbed Caesar with his knife.

The same reasoning that leads us to recognize "Brutus" and "Caesar" as singular terms suggests that "his knife," understood as "Brutus's knife," is a singular term as well; for if Brutus stabbed Caesar with his knife, then Brutus certainly stabbed Caesar with something. Thus, it may be suggested that (2d) contains a three-
place predicate "x stabbed y with z," which relates Brutus, Caesar, and Brutus' knife. In accordance with the foregoing analysis of (2a), let us tentatively propose the following analysis of (2d):

(2d*) Stabbed-With(Brutus, Caesar, Brutus' knife).

Now the problem, according to Davidson, with the foregoing analysis of sentences such as (2a) and (2d) is that it fails to provide an explanation for the validity of the inference involved in A3.

A3  (2d) Brutus stabbed Caesar with his knife.

-----------------------------
(2a) Brutus stabbed Caesar.

The argument in A3 is clearly valid, but since (2a*) and (2d*) do not provide an explanation for the validity of the inference, Davidson rejects the idea that (2a*) and (2d*) give the logical form of (2a) and (2d). We will eventually return to question whether Davidson is justified in rejecting (2a*) and (2d*), but let us first consider the account that he proposes in place of this.

Davidson's basic idea is that adverbs and adverbial phrases, such as "with his knife" in (2d), should be treated on a logical par with adjectives and adjective-phrases. With a view towards explicating Davidson's proposal, let us review the logical form of sentences containing adjectives and adjective-phrases. Consider the following sentence:

(1d) The man in the brown hat, who is a clergyman, is a spy.

Now there are two different ways of representing the logical form of (1d). One is to treat it as a conjunction of two independent sentences, "The man in the brown hat is a clergyman," and "The man in the brown hat is a spy." Thus, letting "Clergyman," stand for
the one-place predicate "x is a clergyman," (1d) might be symbolized as follows:

(1d*) Clergyman(the man in the brown hat) & Spy(the man in the brown hat).

The only drawback to this sort of symbolization is that it fails to represent the pronominalization in (1d). This is only a superficial defect, but it is a defect that is nevertheless avoided by the following symbolization of (1d):

(1d**) (∃x)(Clergyman(x) & Spy(x) & x = the man in the brown hat).

From a logical point of view, (1d*) and (1d**) are equivalent, and either symbolization provides for a demonstration of the validity of the argument in A4.

A4 (1d) The man in the brown hat, who is a clergyman, is a spy.

(1b) The man in the brown hat is a spy.

If we represent (1d) as per (1d*), then the inference in A4 is just a matter of simplification; similarly, if we represent (1d) as per (1d**), then the inference in A4 is essentially a matter of simplification, although in this case additional rules are employed in the proof.

Consider, next, the argument in A5.

A5 (1e) Someone who is a clergyman is a spy.

(1c) Someone is a spy.

Clearly, the same sort of reasoning is involved here as is involved in A4. However, unlike the sentences of A4, the sentences of A5 contain no singular terms. The premise in A5 is an ordinary existential generalization, and can be represented as follows:
\[(1e^*) \ (\exists x) (\text{Clergyman}(x) \land \text{Spy}(x)).\]

Symbolized in this way, the argument in A5 receives a similar explanation to that given for the argument in A4 when its premise is symbolized as per \((1d**).\)

Now Davidson believes that the inference in A3 is analogous to the inference in A5, and that \((2a)\) and \((2d)\) have a logical form analogous to \((1c)\) and \((1e)\) respectively. The logical form of \((1c)\), we have already agreed, is given by \((1c**)\).

\[(1c**) \ (\exists x) (\text{Spy}(x)).\]

According to Davidson, \((2a)\) is also an existential generalization, although whereas the predicate in \((1c)\) is a one-place predicate, the predicate in \((2a)\) is three-place. The predicate that Davidson believes to be involved in \((2a)\) is a predicate that relates Brutus, Caesar, and something that is a stabbing. If we let "Stabbing" stand for a predicate of this sort, we can express Davidson’s proposal for the logical form of \((2a)\) as follows:

\[(2a**) \ (\exists x) (\text{Stabbing}(x, \text{Caesar}, \text{Brutus})).\]

On one interpretation, \((2a**)\) states that there is something that is a stabbing of Caesar by Brutus. Davidson claims that the thing or things that are related to Brutus and Caesar by this sentence—the things over which the sentence quantifies—are events. Davidson also insists that events are not to be confused with objects, or the sort of things over which we quantify in sentences like \((1c**)\). We will return to this question in the next chapter, but the present chapter is reserved for matters of logical form.
One of the more important features of (2a**) is that while it quantifies over events, it does not refer to any particular event. Already in 1947 Hans Reichenbach had claimed that sentences like "George IV was crowned at Westminster Abbey," should be construed as existential quantifications whose variables range over events. But at the same time, Reichenbach thought that this last sentence was equivalent to "The coronation of George IV took place at Westminster Abbey," which apparently does refer to a particular event (Reichenbach 1947, p.269). Davidson admonishes us not to conflate these sentences. According to Davidson, it is a mistake to think that "Leopold met Stephen on Bloomsday" is equivalent to the sentence "Leopold’s meeting with Stephen on Bloomsday occurred," or that "Caesar died," is equivalent to "Caesar’s death took place."

Thus, Davidson writes that:

"Caesar’s death," like "Leopold’s meeting with Stephen," is a singular term, and so "Caesar’s death took place," and "Leopold’s meeting with Stephen occurred," are true only if there was just one such meeting or death. But "Caesar died" is true even if Caesar died a thousand death’s, and Leopold and Stephen may meet as often as they please on Bloomsday without falsifying "Leopold met Stephen on Bloomsday" (Davidson 1969a, p.134)

Of course, it is reasonable to assume that a sentence like "Caesar died" asserts the existence of at most one event, but it is another thing entirely to think that the sentence refers to an event. Similarly, it may be thought that the sentence "There exists a prime between 20 and 28" refers to the number 23, since it is the only number that satisfies this condition. However, according to Davidson, there is a good reason against taking this line, which is that on this view someone could be uniquely referring without
Moreover, many of the sentences that are "about events" do not even suggest, let alone imply, that there is at most a single event under discussion. Thus, "Brutus stabbed Caesar," tells us nothing about how many times he did it, or in other words, how many stabbings of Caesar by Brutus there were. This last remark shows that the pronoun "it," in "he did it" can be misleading. Insofar as it makes sense to say something like "He did it sixteen times," it is clear that the "it" does not refer to any particular event; for a particular event cannot occur more than once. Neither "He did it," nor "Brutus stabbed Caesar," refers to a particular event, but if Davidson is right, both sentences are made true by the existence of certain events. In general, Davidson’s view is that a sentence like "The stabbing of Caesar by Brutus occurred," which contains a definite description of an action, entails but is not entailed by the corresponding sentence "Brutus stabbed Caesar"; according to Davidson, the former sentence is true only if the context picks out exactly one stabbing of Caesar by Brutus (Davidson 1971b, p.194).

We have now seen how Davidson accounts for the logical form of (2a), but it remains to be seen how he would explain the inference in A3, which is what motivates his account of (2a) in the first place. I remarked above that for Davidson, the inference in A3 is analogous to the inference in A5. To complete the analogy, Davidson construes the adverbial phrase "with his knife" in (2d) as consisting of a predicate which is bound by a variable to the same
quantifier as the predicate "Stabbing(x, Caesar, Brutus)." We have already observed that "Brutus' knife" functions as a singular term, and so "with" evidently functions as a predicate. However, if this predicate is to be bound to a quantifier, it must contain an argument-place it does not appear to have. Accordingly, Davidson proposes that the adverbial phrase "with his knife" consists of the two-place predicate "x with y," which relates Brutus' knife to the event or events which make (2d) true. In other words, Davidson's proposal for the logical form of (2d) is as follows:

\[(2d**) \ (\exists x) (\text{Stabbing}(x, \text{Caesar, Brutus}) \& \text{With}(x, \text{Brutus's knife})).\]

On this analysis, the validity of the argument in A3 can be demonstrated by the same sequence of steps needed to demonstrate the validity of the argument in A5. Moreover, the same explanation can be given for an indefinitely large number of valid arguments that are like this one, such as the argument in A6.

A6  \[(2e) \text{Brutus stabbed Caesar with his knife in the Forum.}\]

----------

(2d) \text{Brutus stabbed Caesar with his knife.}\]

The argument in A6 is obviously similar in structure to the argument in A3, and the sort of analysis that Davidson provides for (2a) and (2d) can be extended to account for the validity of the inference in A6 as well. In this case, the adverbial phrase "In the Forum" is given the same sort of treatment as the phrase "With his knife" is given in (2d**).

Before going on to consider some of the complications with, and applications of, Davidson's account of the logical form of action-sentences, let us pause to consider whether one ought to
accept it at all. Davidson seems to suggest that his account of the logical form of action-sentences is justified by the fact that it alone enables us to account for the inferences involving adverbial modification, inferences of the sort in A3 and A6 (Davidson 1969a, p.136). However, it is doubtful that this is true; indeed, as we will presently see, there is at least one, and possibly two, ways in which one might account for the inferences involving adverbial modification without endorsing Davidson's account of the logical form of action-sentences.

In the first place, let us note an alternative to Davidson's proposal that is not sufficient. At first glance, it may seem that another way of explaining the validity of the inference in A3 is by regarding (2a) as elliptical for "Brutus stabbed Caesar with something." On this construal of (2a), the inference in A3 is explained simply as a matter of existential generalization. Now one of the things that suggests that (2a) is elliptical for something like the foregoing is that it is inconceivable that one person could stab another and yet not do it with something. Similarly, it is inconceivable that one person could stab another without doing it somewhere (in the world), somewhere (on the victim's body), and also at some point in time. Thus, the inference in A6 can also be construed as a matter of existential generalization. And when we provide this sort of explanation for the inferences in A3 and A6, there is no need to accept Davidson's account of the logical form of these sentences.
Davidson at one point considers and rejects the foregoing idea (Davidson 1969a, p.136). He rejects it essentially because there is no limit to the number of adverbial modifications that a sentence like (2a) may undergo. Treating (2a) as elliptical for "Brutus stabbed Caesar with something" enables one to explain the inference in A3, but in order to provide the same sort of explanation for every inference like A3 in which (2a) is potentially involved, one would have to regard (2a) as elliptical in an indefinitely large number of ways. On the current supposition, (2a) must be elliptical, not just for "Brutus stabbed Caesar with something," but for "Brutus stabbed Caesar with something, somewhere, at some time, beside something, on something, near something, ...etc." The problem here is not exactly that it is implausible to suppose that a sentence like (2a) is so elliptical, but rather that there seems to be no way of determining just how elliptical it is. And unless it is possible to decide how many argument-places the predicate in (2a) has, there is no way of determining its logical form.

Moreover, even if it were possible to fix the number of adverbial modifications that a sentence like (2a) could undergo, it is clear that this procedure would not provide for an explanation of every inference involving adverbial modifications. In particular, it would not provide an explanation for those inferences involving the elimination of adverbs, rather than adverbial phrases. Thus, consider the argument in A7.

A7  (2f) Brutus fatally stabbed Caesar.

(2a) Brutus stabbed Caesar.
Evidently, it does not matter what we treat (2a) as elliptical for, there is no chance of explaining the inference in A7 as a matter of existential generalization. However, on Davidson's analysis, the validity of the inference in A7 is readily demonstrable. Thus, letting "Fatal" stand as a one-place predicate "x was fatal," (2f) would be rendered as follows:

\[(2f^*) \ (\exists x) (\text{Stabbing}(x, \text{Caesar}, \text{Brutus}) \ & \ \text{Fatal}(x)).\]

And on this analysis, the inference in A7 can receive the same explanation as the inferences involved in A3 and A6.

There is good reason, then, for rejecting the alternative proposal just considered for explaining the inferences involving adverbial modification. However, there is another idea worth considering, which avoids the flaw in the above proposal. Instead of regarding a sentence like (2a) as being elliptical in order to explain the validity of the inference in A3, one might simply claim that the inference in A3 involves a suppressed premise to the effect that "If Brutus stabbed Caesar with his knife, then Brutus stabbed Caesar." More generally, one could say that the inference depends upon a premise to the effect that "(\forall x)(\forall y) (\text{If } x \text{ stabs } y \text{ with } x's \text{ knife, then } x \text{ stabs } y)." This sentence is obviously true—indeed, necessarily true—and the inference in A3 clearly depends upon its truth. But by acknowledging this premise, we can demonstrate the validity of the inference in A3 without endorsing Davidson's account of the logical form of the contained sentences. On this account, (2a) and (2d) can be analyzed along the lines of (2a*) and (2d*), and the inference in A3 can be explained as a
matter of modus ponens. And since the same explanation can be given for the inferences in A6 and A7, it appears that Davidson’s account of the logical form of action-sentences is not justified by the need to explain the validity of inferences involving adverbial modification.

Davidson does not address this idea, but I think it is clear what his response would be—namely, that in explaining the validity of A3 by recourse to the suppressed premise described above, one is taking for granted that which one wants an explanation of in the first place. Thus, instead of asking for an explanation of why the inferences involving adverbial modification are valid, one could just as well ask for an explanation of why the conditionals that function as suppressed premises in these inferences are true. But if, for instance, we analyze (2a) and (2d) as per (2a*) and (2d*), then there is no structural explanation for why these suppressed premises are true. On the other hand, Davidson’s account of the logical form of (2a) and (2d) does provide such an explanation. We have already seen that, according to Davidson, it is not incumbent on a theory of logical form to explain the truth of every conditional that is necessarily true. However, since there seems to be nothing standing in the way of providing such an explanation in the case of the suppressed premises in inferences involving adverbial modification, it is plausible to suppose that a theory of logical form that does provide such an explanation is better than one that does not.
Accordingly, while the proposal most recently considered makes it doubtful that one must accept Davidson's account of the logical form of action-sentences in order to explain the validity of inferences involving adverbial modification, it may be held that Davidson's account of logical form is justified by the fact that it provides a more satisfying explanation than the proposal just considered. In particular, it provides a structural explanation for that which is taken for granted by the alternative proposal. However, if this is the reason for endorsing Davidson's account of the logical form of action-sentences, I believe it can be shown that there is an even better proposal than the one Davidson provides. In order to see this, consider the following three sentences:

(2g) Brutus did something.
(2h) Something was done to Caesar.
(2i) There was a stabbing.

Each of these three sentences is entailed by (2a), but Davidson's account of the logical form of action-sentences fails to provide a structural explanation for at least two of the three inferences. In the case of the inference from (2a) to (2i), Davidson could claim that the conclusion is elliptical for "There was a stabbing of someone by someone," and this non-elliptical version of (2i) is entailed by Davidson's analysis of (2a). Treating (2i) as elliptical for "There was a stabbing of someone by someone" is not implausible, since every stabbing must have an agent and a patient. However, there is clearly no paraphrase of (2g) or (2h) which is entailed by Davidson's analysis of (2a). How, then, are we to
explain the inferences from (2a) to (2g), and from (2a) to (2h), on Davidson’s analysis of action-sentences?

The only option here, it seems, is to recognize a suppressed premise in the relevant inferences. Thus, Davidson could claim that the inference from (2a) to (2g) depends upon a suppressed premise such as "(∀x)(∀y)(∀z)(Stabbing(x, y, z) → Done-By(x, z))." This premise, which tells us that whenever someone stabs another, the stabber does something, is necessarily true, and its truth is clearly relevant to the validity of the inference from (2a) to (2g). However, while Davidson’s account of the logical form of action-sentences forces us to take the truth of this sort of premise for granted, there is another way of accounting for the logical form of action-sentences that provides a structural explanation for its truth.

The solution, as Terrence Parsons has demonstrated, is to decompose Davidson’s three-place predicate in 2a**) into three distinct predicates (Parsons 1980, p.35). Using Parsons’ symbolization, (2a) would be rendered as follows:

(2a***) (∃x)(Stabbing(x) & Of(x, Caesar) & Agent(x, Brutus)).

In this case, "Of" stands for the two-place predicate "x was done of y," which relates Caesar to some event or events, and "Agent" stands for the two-place predicate "x was done by y," or "y was the agent of x," which relates Brutus to the same event or events. On this analysis of (2a), the inferences from (2a) to (2g), from (2a) to (2h), and from (2a) to (2i) receive the same sort of explanation as the inferences in A3, A6, and A7. Most importantly, this account
of the logical form of action-sentences provides a structural explanation for the truth of the premises which must be taken for granted, on Davidson's account, in order to explain the validity of inferences such as those from (2a) to (2g) and from (2a) to (2h). To this extent, it appears that the considerations that one might use to justify Davidson's account of the logical form of action-sentences provide better justification for Parsons' account.

Now Parsons' account of the logical form of action-sentences is slightly more complicated than what we have just observed, but (2a****) does give the essential idea, and illustrates the contrast between his account and Davidson's. Both of these accounts may be called quantificational analyses, or event-analyses, since they both treat action-sentences as existential generalizations whose variables range over events. For the reasons just considered, I think that if one is going to accept a quantificational analysis of action-sentences, one ought to accept Parsons' account over Davidson's. However, regardless of which version of the quantificational analysis one accepts, it is clear that this style of analysis will apply to only certain action-sentences. More precisely, only certain adverbs and adverbial phrases will submit to the sort of treatment provided by a quantificational analysis. Among the adverbs that cannot be treated as predicates of events are "allegedly," "probably," "quickly," "slowly," "accidentally," "deliberately," and "unintentionally." Thus, consider the following sentence:

(2j) Brutus allegedly stabbed Caesar.
If we were to treat the adverb in (2j) in the same way that an quantificational analysis treats the adverb in (2f), or the adverbial phrase in (2d), then (2j) would be analyzed in either of the following two ways:

\[(2j^*) \quad (\exists x)(\text{Stabbing}(x, \text{Caesar, Brutus}) \land \text{Allegedly}(x)).\]

\[(2j^{**}) \quad (\exists x)(\text{Stabbing}(x) \land \text{Of}(x, \text{Caesar}) \land \text{Agent}(x, \text{Brutus}) \land \text{Allegedly}(x)).\]

Now (2j*) entails (2a**), and (2j**) entails (2a***), but (2j) does not entail (2a). So it does not matter which version of the quantificational analysis we chose, in neither case can we treat "allegedly" as a predicate of the event(s) whose existence makes (2a) true. Davidson's response to this is to claim that a sentence like (2j) is actually elliptical for something like "It is alleged that Brutus stabbed Caesar." To put it in this form is not to give the logical form of (2j), but it is, according to Davidson, the first step in seeing exactly what its logical form is. Before we take a look at that next step, let us note another type of adverb that demands special treatment. Consider the following sentence:

\[(2k) \quad \text{Brutus intentionally stabbed Caesar.}\]

Once again, if we were to treat the adverb in (2k) in the same way that a quantificational analysis treats other adverbs and adverbial phrases, then (2k) would be analyzed in either of the following two ways:

\[(2k^*) \quad (\exists x)(\text{Stabbing}(x, \text{Caesar, Brutus}) \land \text{Intentionally}(x)).\]

\[(2k^{**}) \quad (\exists x)(\text{Stabbing}(x) \land \text{Of}(x, \text{Caesar}) \land \text{Agent}(x, \text{Brutus}) \land \text{Intentionally}(x)).\]

Both analyses of (2k) provide for an explanation of the valid inference from (2k) to (2a). However, they also both attribute
structure that allows for inferences that are actually invalid. Thus, consider the following argument:

A8  (2k) Brutus intentionally stabbed Caesar.
   (21) Caesar is identical to the inventor of the Julian calendar.

-----------------------------
(2m) Brutus intentionally stabbed the inventor of the Julian calendar.

Now there is at least one sense in which the foregoing argument is invalid; put differently, there is at least one sense of (2k) and (2m) according to which (2m) might be false, even though (2k) and (21) were true. However, there is a principle of logic, known as the principle of substitutivity, which provides that given a true statement of identity, one of its two terms may be substituted for the other in any true statement and the result will be true (Quine 1953, p.139). Admittedly, the validity of the principle is less obvious than other rules that a standard logical theory employs, but it is widely endorsed nonetheless. Davidson, for instance, claims that the principle of substitutivity is unexceptionable since it "no more than spells out what is involved in the idea of a (logically) singular term" (Davidson 1968, p.93). But given this principle, the argument in A8 should be valid if we analyze (2k) as per (2k*) or (2k**).

Davidson's response to this is analogous to his response in the case of adverbs such as "allegedly," and is contained in the following passage:

The solution that recommends itself is to take such words and phrases as "intentionally," "inadvertently," and "on purpose" as creating semantically opaque contexts, on a par with "possibly" and "probably." If this is right, then such adverbs are not to be treated as predicates of
events or actions at all. If they are predicates, then they are predicates of propositions, sentences, or utterances—whatever, that is, predicates like "Smith said x," or "Jones hopes that y," are predicates of" (Davidson 1985b, p.234).

According to Davidson, the first step in understanding the logical form of a sentence like (2k) is to treat it as being elliptical for something like "It was intentional of Brutus that he stabbed Caesar." The next step, Davidson believes, is to provide whatever sort of analysis for it that one would provide for ordinary propositional attitude-ascriptions, such "Philip is unaware that Tully denounced Cataline," or "Galileo said that the earth moves."

With a view towards evaluating Davidson’s account of the logical form of ascriptions of intentional action, let us now consider his account of the logical form of propositional attitude-ascriptions.

III

Verbs like "seeing," "believing," and "saying" are said to express propositional attitudes because they seem to describe a psychological relation between a person and a proposition. Correspondingly, these verbs appear to function from a semantic point of view as two-place predicates, one of whose argument-places is occupied by the name of an agent, and the other of which is occupied by a sentence that names or expresses the proposition to which the agent is related. In this respect, the verbs that introduce propositional attitudes seem to function as a sort of cross between an ordinary predicate and a sentential operator.
Consider the following sentence:

(3a) Ralph believes that the man in the brown hat is a spy.

At first glance, it appears as though the principle verb in (3a) relates "Ralph" to the sentence "The man in the brown hat is a spy." And since (1b**) gives the logical form of this last sentence, one might propose the following as an account of the logical form of (3a):

(3a*) Believes-That(Ralph, Spy(the man in the brown hat)).

Now Davidson, as we will presently see, would reject the idea that (3a*) gives the logical form of (3a). However, before we can properly appreciate the problem with (3a*), or Davidson's solution to it, we need to note a certain complication to sentences such as (3a). Consider the following argument:

A9  (3a) Ralph believes that the man in the brown hat is a spy.
    (3b) The man in the brown hat is the man at the beach.
    (3c) Ralph believes that the man at the beach is a spy.

It is widely agreed that there is a sense in which the argument in A9 is valid, and another sense in which it is not, and most philosophers explain this fact in terms of a distinction between two different interpretations of sentences like (3a) and (3c). Following a well-established practice, let us say that sentences like (3a) and (3c) have both a de dicto and a de re sense, which we will distinguish with subscripts. Thus, on the de dicto reading of (3a) and (3c), which we will mark as (3a₁) and (3c₁), the inference in A9 is invalid; on the de re reading, which we will mark as (3a₂) and (3c₂), the inference in A9 is valid.
It is also worth noting that "de dicto"/"de re" distinction is often marked in the literature with different terminology, such as "notional"/"relational," "opaque"/"transparent," or "intensional"/"extensional." And while it is not necessarily the case that the distinction between the de dicto and the de re reflects a difference in logical form, many philosophers, including Davidson, believe that it does. So as not to prejudge the issue, it is therefore best to consider the question of the logical form of propositional attitude-ascriptions as dividing into two distinct problems. Let us begin with the de dicto.

Now the basic reason why Davidson would reject the idea that (3a*) gives the logical form of (3a₁) relates to the fact that (3a₁) and (3b) do not, by definition, entail (3c₁). If (3a*) does give the logical form of (3a₁), and, hence, if (3c₁) has a similar structure, then (3a₁) and (3b) should entail (3c₁). This, at any rate, is what the principle of substitutivity leads us to expect. Given the truth of (3b), the principle tells us that "the man in the brown hat" in (3a*) can be replaced with "the man at the beach" without affecting the truth-value of the overall sentence. Thus, analyzing (3a₁) in terms of (3a*) seems to conflict with the validity of the principle of substitutivity.

The classic response to this sort of problem comes from Gottlob Frege. Frege held that the reference of a singular term cannot be identified independently of the sentence in which it occurs, and that when singular terms are used within propositional attitude-ascriptions, they refer to something other than what they
do elsewhere. According to Frege, when singular terms are used within attitude-ascriptions they undergo shift in reference away from their customary or direct reference, and refer instead to their indirect reference or what Frege calls their sense. Thus, on this view, "the man in the brown hat" refers, in (3a₁), to something other than what it refers in a sentence like (1b) or (3b). In the latter two sentences, we may agree that "the man in the brown hat" refers to the man in the brown hat, but according to Frege, in (3a) it refers rather to the sense of the term "the man in the brown hat."

Now if Frege's doctrine of referential shift is correct, then the inference from (3a₁) and (3b) to (3c₁) does not in fact conflict with the principle of substitutivity; rather, the appearance of a conflict is due to our mistaken belief that "the man in the brown hat" refers to the same thing in (3a₁) as it does in (3b). And so by embracing Frege's doctrine, one could conceivably maintain that (3a*) gives the logical form of (3a₁). However, this option is not available to Davidson, who rejects Frege's doctrine of referential shift. Let us first consider Davidson's argument against Frege's doctrine, and then consider his own proposal for the logical form of sentences like (3a₁).

Davidson's objection to the doctrine of referential shift is based upon the observation that attitude-ascriptions can occur as the content clauses in even larger attitude-ascriptions, and that this process of embedding one ascription within another can be repeated indefinitely. Thus, just as a sentence like (1b) can be
embedded within the fragment "Ralph believes that..." to form (3a₁), so too (3a₁) can be embedded within a fragment like "Quine believes that Ralph believes that..." to form (4a₁).

(4a₁) Quine believes that Ralph believes that the man in the brown hat is a spy.

Frege at one point suggested, in a letter to Bertrand Russell, that with each iteration of a propositional attitude, the singular term in the original content clause comes to refer to a new entity (McGuinness 1980, p.153). Thus, just as the "the man in the brown hat" refers to one thing in (1b), and another thing in (3a₁), Frege’s suggestion was that it refers to yet a third thing in (4a₁). If this is right, and we can iterate the attributions of belief indefinitely, then it follows that a singular term like "the man in the brown hat" can refer to any number of things. Davidson therefore claims that Frege’s doctrine of referential shift entails that every singular term is "infinitely many ways ambiguous" (Davidson 1968, p.99). He then rejects the doctrine of referential shift because it stands in the way, he believes, of providing a recursive theory of truth for the natural language. Such a theory, as we have noted, must contain clauses that give the extension of the non-logical constants, but there is clearly no chance of doing this if singular terms are infinitely many ways ambiguous.

There are a couple of responses to Davidson’s criticism that might be made on Frege’s behalf. In the first place, we should note that at least one philosopher, Alonzo Church, has tried to show how the ambiguity that affects singular terms, on Frege’s doctrine, can be systematized. Davidson, however, considers this point, and
responds by claiming that the systematization that Church produces nevertheless fails to satisfy the requirements of a recursive theory of truth. According to Davidson, Church proposed a language notationally superior to the natural language, in which one could keep track of the multiple referents assigned to any singular terms, but the expressions that Church introduced to fulfil this role are all semantic primitives. Thus, in the place of a finite number of singular terms that are infinitely many ways ambiguous, Church proposed a model of the natural language containing an infinite number of singular terms (Davidson 1968, p.99).

Another suggestion that has been made on Frege's behalf, and which Davidson has not explicitly considered, comes from Michael Dummett. According to Dummett, we can interpret Frege's doctrine of referential shift in such a way that it does not entail that singular terms are ambiguous in an infinite number of ways (Dummett 1973, p.268). In particular, instead of accepting the idea that a singular term comes to refer to a new entity with each iteration of a propositional attitude, Dummett insists that a singular term in an ordinary attitude-ascription, such as (3a₁), refers to the same thing as it refers to in an iterated attitude-ascription, such as (4a₁). Thus, on this interpretation, every singular term refers to either one of two things—its customary reference or its sense—and so it does not follow that singular terms are infinitely many ways ambiguous.

Assuming that Dummett's interpretation is viable, it appears as though Davidson's rejection of Frege's doctrine is based upon a
rather superficial defect in the theory. And if this is right, then it cannot be said that Davidson provides a convincing objection to the idea that (3a*) gives the logical form of (3a₁). Nevertheless, let us now go on to consider the analysis that Davidson himself provides for sentences like (3a₁).

Davidson's discussion of the logical form of propositional attitude-ascriptions is contained mostly in an article entitled "On Saying That," in which he focuses exclusively on the verb "saying," rather than "believing." However, he does indicate elsewhere that his account of the former is intended to serve as an account of the latter as well, and of propositional attitude-ascriptions generally (Davidson 1989, p.14). Applying his remarks in the case of saying to that of believing, Davidson would represent the logical form of (3a₁) as follows:

(3a₁*) Believes(Ralph, that) Spy(the man in the brown hat).

In the case of (3a*), "Believes-That" stands for the two-place predicate "x believes that y," in which the first argument-place is occupied by the singular term "Ralph," and the second is occupied by the sentence "The man in the brown hat is a spy." But in the case of (3a₁*), "Believes" stands for the two-place predicate "x believes y," in which the first argument-place is occupied by the term "Ralph", and the second is occupied by the demonstrative "that." The demonstrative, according to Davidson, refers to an utterance or inscription of the sentence that follows it—namely, "The man in the brown hat is a spy." And so the crucial difference between (3a*) and (3a₁*) relates to the role played by "that." On
Davidson's proposal, it is a demonstrative which has the effect of severing what appears to be a single sentence into two logically independent sentences. For this reason, Davidson's proposal is often called a *paratactic* analysis of attitude-ascriptions.

Davidson's paratactic analysis provides a structural explanation for why (3a₁) and (3b) do not entail (3c₁). If the analysis is correct, then it follows that when we attempt to substitute "the man at the beach" for "the man in the brown hat," we perform the substitution upon the second sentence in (3a₁*), but then judge the truth-value of the first. On this analysis, the idea that (3c₁) should follow from the truth of (3b) and (3a₁) is no less confused than the idea that (3c₁) should follow from the truth of (3b) and "The man in the brown hat is a spy." In fact, it is even worse; for it does not follow from the truth of (3a₁) that "The man in the brown hat is a spy" is true. Thus, Davidson concludes that the appearance of conflict with the principle of substitutivity is due to "our mistaking what are really two sentences for one" (Davidson 1968, p.107). And this, of course, contrasts sharply with Frege's explanation, according to which the appearance of conflict is explained in terms of our being mistaken about the semantic properties of the singular terms within the content-clauses of attitude-ascriptions.

Tyler Burge has objected to Davidson's paratactic analysis on the grounds that it leaves one unable to account for the validity of certain inferences involving propositional attitude-ascriptions. In particular, Burge claims that Davidson's analysis cannot explain
the validity of what he calls "arguments by repetition" (Burge 1986, p.204). As an example of such an argument, consider the inference from (3a₁) to (3a₁). It is clear that this inference is valid, but it is not clear how it can be explained if we analyze the premise and conclusion according to Davidson’s analysis. For on Davidson’s analysis, the demonstrative in the premise refers to one utterance or inscription, and the demonstrative in the conclusion refers to yet another. And there is no reason to expect, on the basis of logical form alone, that if the referent of the first demonstrative makes the premise true, the referent of the second demonstrative will make the conclusion true. Thus, if Davidson’s paratactic analysis is correct, then in order to explain the validity of the inference from (3a₁) to (3a₁), further premises are needed. This itself is not necessarily an objection to the paratactic analysis, but it is if the required premises cannot be supplied.

There is, moreover, another potential objection to Davidson’s analysis, which is closely related to Burge’s criticism. Consider the sentence "Ralph knows that the man in the brown hat is a spy." Like (3a), this sentence has both a de dicto and a de re reading, but on either reading the sentence entails (1b), "The man in the brown hat is a spy." However, if the de dicto reading of the original sentence receives a paratactic analysis, then how can the inference to (1b) be explained? Without the paratactic analysis we could explain the inference by recourse to a suppressed premise such as "If Ralph knows that the man in the brown hat is a spy,
then the man in the brown hat is a spy." However, this option is not available on the paratactic analysis; for what would function as the antecedent in the suppressed conditional is, on the paratactic analysis, two independent sentences. Thus, while it is clear that Davidson would have to appeal to suppressed premises in order to explain the inference just noted, it is not at all clear what they would be, and this certainly raises doubts about the paratactic analysis.

The foregoing criticism of the paratactic analysis may or may not be fatal, but assuming that Davidson can supply an adequate response, there is a further point worth mentioning, and which I think requires certain amendments to the paratactic analysis. In the first place, we must note that the verbs that express propositional attitudes are modified by adverbs and adverbial phrases just as the verbs of action are. Thus, we say things like "Clifford adamantly denies that he assaulted the woman," "The woman vividly remembers that her assailant had a mark on his face," and "The pilot calmly suggested that we prepare for a crash-landing."

Consider, then, the following sentence:

(3d₁) Ralph firmly believes that the man in the brown hat is a spy.

Evidently, (3d₁) stands to (3a₁) just as (2d) stands to (2a). And according to Davidson, the latter relation demonstrates that the predicate corresponding to the principle verb in (2a) contains an argument-place that it does not appear to have. By parity of reasoning, it seems that Davidson should treat the predicate corresponding to the principle verb in (3a₁) as containing an
argument-place that it does not appear to have. In other words, it appears as though Davidson is committed to something like the following as the logical form of (3d1):

\[(3d_1^*) (\exists x) (\text{Believing}(x, \text{Ralph, that}) \& \text{Firm}(x)) \]
\[\text{Spy(\text{the man in the brown hat})}.
\]

In this case, "Believing" stands for the three-place predicate "x is a believing by y, z," in which the first argument-place is occupied by a variable bound to a quantifier, the second is occupied by the term "Ralph," and the third is occupied by the demonstrative "that," which refers to an utterance or inscription of the second sentence in (3d1*). If our ontology includes states, in addition to events, then we may treat the variables in (3d1) as ranging over states, and perhaps speak of a belief of Ralph’s instead of a believing by Ralph. However, the important point is that there appears to be just as much reason for a quantificational analysis of sentences like (3a1) and (3d1) as there is for a quantificational analysis of sentences like (2a) and (2d). Accordingly, I think we say that Davidson is committed to something like the following as an analysis of (3a1):

\[(3a_1^{**}) (\exists x) (\text{Believing}(x, \text{Ralph, that})) \]
\[\text{Spy(\text{the man in the brown hat})}.
\]

Moreover, if we are right in thinking that the considerations used to justify Davidson’s quantificational analysis of action-sentences actually justify Parson’s quantificational analysis, then it seems that Davidson really ought to provide something more like the following as an account of the logical form of (3a1):

\[(3a_1^{***}) (\exists x) (\text{Believing}(x) \& \text{Agent}(x, \text{Ralph}) \& \text{Of}(x, \text{that})) \]
\[\text{Spy(\text{the man in the brown hat})}.
\]
We have so far considered Davidson’s approach to the logical form of attitude-ascriptions on their *de dicto* reading, but we have yet to say anything about the form of attitude-ascriptions on their *de re* reading. Unfortunately, Davidson has not explicitly addressed this question, and it is clear that the analysis he provides for the *de dicto* reading will not work as an analysis of the *de re*. Thus, it is clear that none of the versions of the paratactic analysis just considered for (3a₁) give the logical form of (3a₂); for these analyses are all designed to provide a structural explanation for why (3a₁) and (3b) do not entail (3c₁), but (3a₂) and (3b) do, by definition, entail (3c₂).

However, for the purposes of understanding the logical form of ascriptions of intentional action, to which we will now return, it is not necessary to have an account of the logical form of *de re* reading of attitude-ascriptions. For while ascriptions of intentional action have both a *de dicto* and a *de re* reading, the latter present no logical problem for the quantificational analysis of action-sentences. Thus, consider (2k) again.

(2k) Brutus intentionally stabbed Caesar.

We noted at the end of the last section that if adverbs of intention are to receive the same sort of treatment as a quantificational analysis provides for other adverbs, then (2k) would be analyzed in one of the following two ways:

\[(2k\ast) (\exists x)(\text{Stabbing}(x, \text{Caesar}, \text{Brutus}) \& \text{Intentionally}(x)).\]

\[(2k\ast\ast) (\exists x)(\text{Stabbing}(x) \& \text{Of}(x, \text{Caesar}) \& \text{Agent}(x, \text{Brutus}) \& \text{Intentionally}(x)).\]
For the sake of making a point, let us assume that we have settled on Parsons' version of the quantificational analysis and, hence, (2k**). Now there is no problem accepting this as an analysis of the de re sense of (2k), which we will designate (2k₂). What (2k**) does not provide an analysis of is the de dicto sense of (2k), which we will designate (2k₁). Thus, the question that was left hanging at the end of the last section, and to which we may now return, is the question of the logical form of (2k₁).

Davidson's response to this question, as we noted, divides into two steps, the first of which is to paraphrase (2k₁) into the form of a propositional attitude-ascription; hence, "It was intentional of Brutus that he stabbed Caesar." We are now in a position to appreciate the second step. Employing Davidson's version of the quantificational analysis, his proposal is that (2k₁) be analyzed as follows:

\[(2k₁*) \text{Intentional}(\text{Brutus}, \text{that}) \quad (\exists x)(\text{Stabbing}(x, \text{Caesar}, \text{him})).\]

On this analysis, "Intentional," stands for a two-place predicate "It was intentional of x, y," in which the first argument-place is occupied by the term "Brutus," and the second is occupied by the demonstrative "that." As usual, the demonstrative is taken to refer to an utterance or inscription of the sentence that follows—namely, "He stabbed Caesar." Since the adverb "intentionally" is not itself modified by further adverbs, the quantification analysis of (2k₁) need not insert an additional argument-place in the predicate corresponding to the adverb of intention. However, if Parsons version of the quantificational analysis is correct, the logical
form of \((2k_1)\) is slightly different. On Parsons' version of the analysis, \((2k_1)\) is rendered as follows:

\[
(2k_{1**}) \text{ Intentional(Brutus, that)} \cup(\exists x)(\text{Stabbing}(x) \& \text{Of}(x, \text{Caesar}) \& \text{Agent}(x, \text{he})).
\]

We recently observed that there may be an interpretation of Frege's doctrine of referential shift which escapes Davidson's objection to it, and it is worth noting at this point that if Frege's doctrine of referential shift were correct, there would be no need for a paratactic analysis of ascriptions of intentional action. Thus, we could simply say that adverbs of intention create a context in which singular terms refer to their sense as opposed to their customary reference. More precisely, the adverbs of intention create a context that affects only the singular terms that refer to the direct object of the action, such as "Caesar" in \((2k)\); the singular terms that refer to the agent of the action are clearly unaffected. But if Frege's doctrine of referential shift is correct, then there is no need to provide a different account of the logical form of \((2k_1)\) than is provided for \((2k_2)\). Thus, if we accept \((2k_{**})\) as an analysis of \((2k_2)\), and if the doctrine of referential shift is correct, then there is no reason not to accept \((2k_{**})\) as an analysis of \((2k_1)\) as well.

Moreover, one clear advantage to analyzing \((2k_1)\) along the lines of \((2k_{**})\), rather than \((2k_1*)\), is that the former analysis does, and the later analysis does not, provide an explanation for the inference from \((2k_1)\) to \((2a)\). And it is not at all clear how this inference can be explained if \((2k_1)\) is analyzed along the lines of \((2k_1*)\) or \((2k_{**})\). In other words, it is not clear what
sort of subsidiary premisses can be used to explain this inference. The point here is precisely the same as the point raised in connection with the paratactic analysis of a sentence like "Ralph knows that the man in the brown hat is a spy." Thus, I think we may conclude that Davidson's analysis of ascriptions of intentional action is no more, or less, objectionable than his analysis of propositional attitude-ascriptions generally.

IV

One sort of sentence that is of paramount importance in the philosophy of action is what philosophers call reason-explanations. These are sentences such as "James went to church because he wanted to please his mother," and "Hamlet killed the man behind the arras because he wanted to avenge his father's death." In Chapter 3, we will consider the meaning of these sentences, and how they do or do not differ from other kinds of explanation, but it will be of some assistance if we can first clarify the logical features of these sentences. Having considered the question of the logical form of action-sentences and attitude-ascriptions, let us now go on to consider the logical form of reason-explanations, which are constructed out of the former.

Notwithstanding a few brief remarks made in passing, the question of the logical form of reason-explanations is another one which Davidson has not dealt with in any direct way. Nevertheless, I believe that one can see the sort of account to which Davidson is
committed in virtue of his endorsement of a quantificational analysis of action-sentences and a paratactic analysis of propositional attitude-ascriptions. In the remainder of the chapter, I will try to sketch this account. However, for reasons that will become clear as we proceed, it will be best to approach the question of the logical form of reason-explanations indirectly, by first considering Davidson’s account of the logical form of causal statements.

In "Causal Relations," Davidson distinguishes between the following three sentences:

(5a) Jack’s fall caused the breaking of his crown.
(5b) Jack fell down, which caused it to be the case that Jack broke his crown.
(5c) The fact that Jack fell down explains the fact that Jack broke his crown.

According to Davidson, (5a) consists of the two-place predicate "x caused y," in which the first argument-place is occupied by the term "Jack’s fall," and the second is occupied by the term "the breaking of Jack’s crown." Using the logical symbol for definite descriptions, Davidson represents the descriptions given in (5a) as "(ix)Fall(x, Jack)," and "(ix)(Breaking(x, Jack’s crown))." And with these resources, Davidson gives the following account of the logical form of (5a):

(5a*) Caused(((ix)(Falling(x, Jack),
           (ix)(Breaking(x, Jack’s crown))).

On a fairly literal translation, (5a*) states that the fall of Jack caused the breaking of Jack’s crown.

Now if (5a*) correctly gives the logical form of (5a), it is clear that (5b) requires a slightly different analysis; for unlike
(5a), (5b) contains no definite descriptions of events. Thus, in the place of a two-place causal predicate, one might that claim (5b) contains the sentential connective "The fact p caused it to be the case that q," where "p" and "q" are replaced by the appropriate sentences. Davidson, however, rejects this idea, and he claims to have an argument which proves that a sentence like (5b) does not, from a logical point of view, contain the sentential connective just mentioned. Let us briefly consider this argument.

The argument begins with three observations or assumptions. The first is that the sentential connective "The fact that p caused it to be the case that q" is non-truth-functional. A truth-functional connective, such as "and," is one that combines with other sentences to form a more complex sentences whose truth-value is a function of the truth-value of its component sentences. Now a sentence like (5b) is true only if each each of its component sentences is true. However, since (5b) may change from true to false if one of the contained sentences is switched for another true sentence, the truth-value of (5b) is not a function of the truth-value of its component sentences. Thus, the alleged connective in (5b) is not truth-functional. The second observation that Davidson makes is that the singular terms in (5b) can be substituted salva veritate for other co-referential singular terms. Thus, Davidson writes that:

If the fact that there was a fire in Jones' house caused it to be the case that the pig was roasted, and Jones' house is the oldest house on Elm Street, then the fact that there was a fire in the oldest house on Elm Street caused it to be the case that the pig was roasted. (Davidson 1967b, p.152).
Now in order to establish the intended conclusion, Davidson makes one further assumption, which is that we cannot change the truth-value of a sentence like (5b) by substituting logically equivalent sentences for sentences within it (Davidson 1967b, p.153). Let us now see, how these assumptions are supposed to entail that (5b) does not contain the given sentential connective.

According to Davidson, the sentence "Jack broke his crown" is logically equivalent to "ś(s=s & Jack broke his crown) = ś(s=s)," and (5b) therefore retains its truth-value if we substitute the latter for the former. And since "ś(s=s & Jack broke his crown)" is co-extensive with "ś(s=s & Nero fiddled)," (5b) also retains its truth-value if we replace "Jack broke his crown" with "ś(s=s & Nero fiddled) = ś(s=s)." Finally, since this last sentence is logically equivalent to "Nero fiddled," (5b) retains its truth-value even if we substitute "Nero fiddled" for "Jack broke his crown." Thus, contrary to the original supposition, the alleged sentential connective in (5b) must be truth-functional. In order to escape the dilemma, Davidson rejects the idea that (5b) contains the alleged sentential connective.

The argument is clearly valid, and so if we accept the premises we ought to accept the conclusion, but there is at least some question as to whether we ought to accept the premises. In particular, there is at least some question concerning the crucial assumption that the substitution of logically equivalent sentences within (5b) does not affect its truth-value. Why should one accept this assumption? Certainly the assumption does not hold in the case
of other apparently non-truth-functional sentential operators. Thus, it surely does not follow from the fact that Jill believes that Jack broke his crown that Jill believes that \( \hat{s}(s=s & \text{Jack broke his crown}) = \hat{s}(s=s) \). And if the principle Davidson cites does not apply in this case, why must we accept it in the case of (5b)?

Davidson would likely respond to this by claiming that the crucial difference in the case of the alleged sentential operator "X believes that p" is that, unlike the alleged sentential operator in (5b), it does not allow for the substitution of co-referential singular terms. However, it seems that the foregoing objection applies even if we stipulate that the attitude-ascription is to be understood in its de re sense, which does permit the substitution of co-referential terms. In other words, even on its de re reading it seems that "Jill believes that Jack broke his crown" does not entail "Jill believes that \( \hat{s}(s=s & \text{Jack broke his crown}) = \hat{s}(s=s) \)."

There is, then, at least some doubt as to whether Davidson establishes that (5b) does not contain a sentential connective, but let us go on to consider the analysis that he provides in the place of this. According to Davidson, the only difference between (5a) and (5b) is that the later does not contain any singular terms that refer to events. However, the fact that (5b) does not contain terms that refer to any individual events does not entail that it does not contain a causal predicate that relates individual events. According to Davidson it does, and he therefore proposes the following as an account of the logical form of (5b):

\[(5b^*) (\exists x)(\exists y)(\text{Fall}(x, \text{Jack}) & \text{Breaking}(y, \text{Jack's crown}) & \text{Caused}(x, y))\]
On a fairly literal reading, (5b*) asserts that there is at least one fall of Jack and at least one breaking of Jack's crown, such that the former caused the latter. And this, Davidson believes, is logically equivalent to (5b).

Now there is, it seems, a problem with Davidson's analysis of (5b). As long as each breaking of Jack's crown is the result of a single fall of Jack, then (5b*) is unexceptionable. However, it may be that Jack's crown breaks as the result of some combination of falls by Jack, and if so, then (5b*) clearly gives the wrong sense. Thus, suppose that Jack fell down twice, broke his crown once, and that it was not the first or the second fall, but the first and the second fall, that caused Jack's crown to break. In the event, (5b) could be true, and yet (5b*) false. And if this is right, then we are in need of some other account of the logical form of (5b).

Davidson remarks, at one point in "Causal Relations," that there are a number of causal sentences, very much like (5b), that do raise the issue of apparently non-truth-functional sentential connectives" (Davidson 1967, p.161). One of the examples he gives is "The fact that the dam did not hold caused the flood." However, Davidson does not provide any account of these sentences; rather, he simply suggests that sometimes the word "caused" functions not to express a relation between events, but rather to express an explanatory relation amongst facts or propositions. Thus, he explicitly endorses an idea espoused by Zeno Vendler, among others, that there is a deep distinction in our use of "cause," "effect," and related words, between occurrences of verb-nominalizations that
are fact-like or propositional, and occurrences that are event-like (Davidson 1967b, p.162).

By a "deep distinction," Davidson means a distinction in logical form; the distinction he has in mind is a distinction between causal statements that do, and causal statements that do not, contain a sentential connective. The former, he suggests, are explanations, which typically relate statements rather than events (1967b, p.161). And to mark the difference, he claims that the causal idioms that do contain a sentential connective are best expressed with the words "causally explains" rather than "caused." But if this is right, then it may explain why (5b) does not seem to submit to the treatment Davidson provides for it. It may be that (5b) is a sort of causal explanation, rather than an ordinary causal statement, and, hence, that it should be analyzed along the lines of (5c). Let us therefore consider the logical form of (5c).

The first thing to notice about sentences like (5c) is that there is at least in one sense in which they do not allow for the substitution of co-referential singular terms. Thus, consider the sentence "The fact that the president was murdered explains the fact that there was a national crisis." Now the president may be the fattest man in Washington, but substituting the latter for the former may affect the truth-value of the given explanation. More precisely, I think we should say that there is a de dicto reading of the explanation, which does not allow for the substitution of the co-referential terms, and a de re reading, which does. As usual, let us mark the de dicto sense (5c₁) and the de re (5c₂).
It appears, then, that the question of the logical form of (5c) is directly analogous to the question of the logical form of propositional attitude-ascriptions, and should be handled in much the same way. And if this is right, then Davidson is no more justified in claiming that (5c) contains a sentential connective than he would be in claiming that a sentence like (3a) contains a sentential connective. Moreover, there is good reason why Davidson should not make this claim; for if (5c) does contain a sentential connective, then it is a non-truth-functional connective, and at least from his point of view, non-truth-functional connectives stand in the way of providing a recursive theory of truth for the natural language.

Now if we are right in thinking that (5b) should be treated as a kind of explanation, rather than an ordinary causal statement, then it is clear that it should be treated along the lines of (5c2) rather than (5c1); for (5b) allows for the substitution of coreferential singular terms. Davidson, as we have already observed, does not provide a semantic account of the de re idioms. However, since he does provide an account for the de dicto idioms, we can at least consider the sort of analysis that he ought to provide for a sentence like (5c1). Indeed, it seems that we can extend Davidson’s paratactic analysis of sentences like (3a1) and (2k1) to account for (5c1) as well.

Thus, let us suppose that each of the two occurrences of the term "that" in (5c1) are demonstratives that refer to an utterance or inscription of the sentence that follows them. We can then
acknowledge a two-place predicate "x explains y" whose argument-places are occupied by demonstratives that refer to the relevant utterances. Accordingly, we can provide, on Davidson’s behalf, the following account of the logical form of (5c₁):

\[(5c₁^*) \text{Explains}(\text{that}, \text{that})\]
\[\exists x (\text{Fall}(x, \text{Jack}) \land \exists x (\text{Breaking}(x, \text{Jack's crown})).\]

The foregoing is merely a more complicated form of the paratactic analysis that Davidson provides for attitude-ascriptions. In the case of (5c₁*), the first demonstrative is to be understood as referring to an utterance or inscription of the second sentence, and the second demonstrative is to be understood as referring to an utterance or inscription of the third sentence. There is, of course, the question of how it is that we know what the demonstratives refer to, but this is no less of a problem for the basic paratactic analysis. In general, it seems that (5c₁*) is no more problematic than Davidson’s analysis of (2k₁).

Finally, now that we have seen how Davidson would, or at least should, account for the logical form of (5c₁), we can also see the sort of account that he should provide for the logical form of reason-explanations. Consider the following sentence:

(6a) Brutus stabbed Caesar because he wanted to end the tyranny.

That there is a distinction to be drawn between two different senses of (6a) is clear from the following argument:

A10 (6a) Brutus stabbed Caesar because he wanted to end the tyranny.
(6b) Caesar is the inventor of the Julian calendar.

(6c) Brutus stabbed the inventor of the Julian calendar because he wanted to end the tyranny.
Given that there is a sense in which the inference in A10 is invalid, and another sense in which it is not, let us distinguish between a *de dicto* and a *de re* reading of the contained explanations. Once again, (6a₁) marks the *de dicto* sense of (6a), and (6a₂) marks the *de re*. Now if we are right in thinking that a paratactic analysis can be provided for (5c₁), then the same sort of treatment ought to be possible for (6a₁). The first step is to paraphrase (6a₁) as follows: "The fact that Brutus wanted to end the tyranny explains the fact that he stabbed Caesar." And the rest is obvious, with the sole complication being the need for an account of the logical form of the attitude-ascription in (6a), which we will label (6d).

(6d) Brutus wanted to end the tyranny.

Now it is reasonable to suppose that the content clause in (6d) is propositional in character, and that it contains a suppressed reference to Brutus. Let us make this explicit by paraphrasing (6d) as "Brutus wanted it to be the case that he ended the tyranny." Accordingly, the content-clause in (6d) is itself an action-sentence. Since we know how Davidson analyzes action-sentences, and also how he analyzes attitude-ascriptions on their *de dicto* sense, we can propose the following as a Davidsonian analysis of (6d₁):

(6d₁*) Wanted(Brutus, that) 
(∃x)(Ending(x, the tyranny, him)).

As usual, the demonstrative in the first sentence is to be understood as referring to an utterance or inscription of the second sentence. In fact, I think that Davidson is committed to something slightly more complicated than this, which we will
consider momentarily, but let us first use (6d,*) to bring out the sort of analysis which Davidson ought to provide for (6a,).

Roughly, it is as follows:

(6a,*) Explains(that, that)
    Wanted(Brutus, that) (\exists x)(Ending(x, the tyranny, Brutus))
    (\exists x)(Stabbing(x, Caesar, Brutus)).

On this analysis, (6a,*) consists of four independent sentences. The first sentence contains the two-place predicate "x explains y," both of whose argument-places are occupied by demonstratives. The first demonstrative is to be understood as referring to an utterance of the second sentence, and the second demonstrative is to be understood as referring to an utterance of the fourth sentence. The second sentence, in turn, contains a demonstrative that refers to an utterance of the third sentence. It goes without saying that the analysis is unwieldy, but it is not therefore incorrect.

However, given the results of the preceding sections, there are at least two ways in which (6a,*) ought to be adjusted. The first adjustment is to accommodate Parsons' version of the quantificational analysis of action-sentences, which appears to be better justified than Davidson's. The second adjustment is to accommodate the fact, noted above, that attitude-ascriptions ought to receive a quantificational analysis as well. Given these two points, the analysis of (6a,*) to which Davidson appears to be committed is actually something like the following:

(6a,**) Explains(that, that)
    (\exists x)(Wanting(x) & Of(x, that) & Agent(x, Brutus))
    (\exists x)(Ending(x) & Of(x, the tyranny) & Agent(x, Brutus))
    (\exists x)(Stabbing(x) & Of(x, Caesar) & Agent(x, Brutus)).
We observed at the end of the last section that there is at least some question concerning the viability of Davidson’s analysis of ascriptions of intentional action. The problem relates to the question of how, or if, one can explain the validity of inferences such as the one from \((2k_1)\) to \((2a)\) given Davidson’s analysis of \((2k_1)\). Now precisely the same point applies to the foregoing analysis of \((6a_1)\); for \((6a_1)\) entails both \((2a)\) and \((6d)\). Thus, if the foregoing analysis is correct, we are in need of some account of why these last two inferences are valid. It is not clear whether such an account can be supplied, but what we can say is that the problem here is no different than the problem that affects the paratactic analysis of sentences like "Ralph knows that the man in the brown hat is a spy." If a paratactic analysis can succeed in the case of this last sentence, then it should also succeed in the case of reason-explanations.

Having considered Davidson’s views on the logical form of action-sentences, let us now go on to examine his views concerning the individuation of action. The issue of individuation, like many other issues in the philosophy of action, depends in various ways upon an understanding of the logical features of action-sentences. And we will see as we proceed how Davidson’s views on logical form support, or fail to support, his views on individuation.
In "Actions, Reasons, and Causes," Davidson remarks that the word "action" does not very often occur in ordinary speech, and when it does it is usually reserved for fairly portentous occasions (Davidson 1963, p.5). But in "Agency," Davidson adopts "the positive assumption...that there is a fairly definite subclass of events that are actions," which he attempts in that paper to characterize (Davidson 1971, p.44). We will consider this assumption in detail below, but let us begin by noting what, according to Davidson, an event is.

There is a tradition in philosophy, tracing back as far as Aristotle, of treating language as a representation of reality. Correspondingly, there is the widespread practice of using language as an instrument of metaphysical investigation. Very few philosophers would deny that our familiar ways of speaking can be confused and misleading, but very many philosophers would assert that our language nevertheless has a discernible structure which, when made manifest, informs us about the structure of reality. And since it is the business of the logician to clarify our confused ways of speaking, and lay bare the logical structure of our language, logic has come to play a crucial role in these otherwise metaphysical questions.
We observed in the last chapter that, according to Davidson, there are two related reasons for attributing logical structure to the sentences in the natural language, the most important of which is to implement a recursive theory of truth. If logical structure reflects metaphysical structure, then in providing a theory of truth for the natural language, one necessarily paints a metaphysical picture. Davidson embraces this consequence, and consequently holds that a theory of truth not only tells us how the language works, but also tells us about how the world works, or at least what it contains. To use a theory of truth as a guide to ontology is to employ what Davidson calls the method of truth in metaphysics.

One of the more important metaphysical consequences that follows from a recursive theory of truth, according to Davidson, is the existence of events. Thus, Davidson believes that a theory of truth reveals that the natural language involves an ontology of events no less than, and in precisely the same sense as, it involves an ontology of objects. In an article appropriately entitled "The Method of Truth in Metaphysics," he writes that:

A metaphysician who is willing to suppose no sentences like "Vesuvius erupted in March 1944" or "Caesar crossed the Rubicon" are true will not be forced by a theory of truth to admit the existence of events or even, perhaps, of people or mountains. But if he accepts that many such sentences are true (whichever they may be), then it is obvious that he must accept the existence of people and volcanoes; and, if I am right, the existence of events like eruptions and crossings. (Davidson 1977a, p.214)

According to Davidson, a sentence like "Caesar crossed the Rubicon" should be analyzed as "(\exists x)(\text{Crossing}(x, \text{the Rubicon}, \text{Caesar})." This
is the logical form that Davidson believes we must attribute to the
given sentence if we are to account for the truth-conditions of
this sentence, and the many others that it entails and is entailed
by. Moreover, Davidson takes this as showing that the natural
language embraces an ontology of things like crossings in addition
to things like people and rivers; for Davidson interprets the
symbolized sentence just given as asserting the existence of at
least one thing that is a crossing, and which involves Caesar and
the Rubicon.

Now it should be noted that in interpreting the symbolized
sentence as such, Davidson employs an objectual interpretation of
quantificational sentences, and that on a substitutional
interpretation, the ontological implications that Davidson insists
upon do not obtain. This requires a qualification to the so-called
method of truth in metaphysics. Davidson’s metaphysical method is
based not only on the traditional idea of using logic as an
ontological tool, but also on the narrower idea that "to be is to
be the value of a bound variable." The well-known author of this
idea is W.V. Quine, who explains it as follows:

The artificial notation "\(\exists x\)" of existential
quantification is explained merely as a symbolic
rendering of the words "there is something x such that."
So, whatever more one may care to say about being or
existence, what there are taken to be are assuredly just
what are taken to qualify as values of "x" in
quantifications. The point is thus trivial and obvious.
(Quine 1990, p.27)

But the point is not at all trivial or obvious. Quine himself, in
the same text, speaks of the possibility of a substitutional
interpretation, and on a substitutional interpretation "\(\exists x\)" is not
a symbolic rendering of the words "there is something x such that." Rather, on a substitutional interpretation, a sentence of the form "(∃x)(Fx)" is interpreted as saying that there is a true substitution-instance of the open-sentence "(Fx)." In effect, it states that there is a term which, when substituted in the open-sentence, forms a true sentence. As such, it asserts the existence of nothing beyond a certain term. A substitutional interpretation does not in any way imply that nothing beyond the language exists, but it does imply that the question of what exists is not as straightforward as Quine suggests.

Now Davidson does indeed reject a substitutional interpretation of quantificational sentences, although he gives very little justification. What he does say is that a substitutional interpretation does not satisfy the demands of a recursive theory of truth (Davidson 1973a p.69). But it is not at all clear that this is true, and in a paper entitled "Is There a Problem About Substitutional Quantification?" Saul Kripke raises significant doubts about the objections that Davidson and others raise against a substitutional interpretation (Kripke 1976). Nevertheless, I will not pursue this point in the present context; for the purposes of this thesis, we will continue to interpret quantificational sentences as having the ontological implications that Davidson attributes to them.

Davidson believes that a theory of truth for the natural language requires an ontology of things like eruptions and crossings, and that these are events. This gives us Davidson’s
reason for believing that events exist, but we have yet to see what he thinks events are. All that we know so far is that they are the things over which we quantify in sentences like "Vesuvius erupted," and "Caesar crossed the Rubicon." What we need to know, in addition to this, is what the criteria are for individuating them.

Davidson, at one point, tentatively suggested that events that have the same causes and effects are identical, but he subsequently renounced this idea in response to an objection by Quine (Quine 1985, p.166). In place of the earlier suggestion, Davidson came to embrace Quine's proposal that events are identical if they have the same spatio-temporal co-ordinates (Davidson 1985c, p.175). However, even though Quine and Davidson both endorse the same criterion, they understand it slightly differently. According to Quine, spatio-temporal identity is also the criterion for individuating objects, and as such, he embraces the consequence that there is no distinction to be drawn between objects and events. Davidson, however, rejects this result. Thus, he writes that:

Occupying the same portion of space-time, event and object differ. One is an object that remains the same object through changes, the other a change in an object or objects. Spatiotemporal areas do not distinguish them, but our predicates, our basic grammar, our ways of sorting do. Given my interest in the metaphysics implicit in our language, this is a distinction I do not want to give up. (Davidson 1985c, p.176).

Elsewhere, Davidson suggests that events occur at a time in a place, while objects occupy places at times, and that this semantic distinction between occurring and occupying reflects a corresponding ontological distinction.
Now if events are things that occur at times in places, then it would seem that actions are also events; for they too occur at times in places. If Oswald killed Kennedy, then he did so at a certain time and in a certain place; if he did it, he did it on November 22, 1963 in Dallas, Texas. Similarly, spatio-temporal coordinates can, at least in principle, be assigned to anything that we would count as an action; for it is not possible for someone to do something and yet not do it somewhere and at some time. Of course, it is not always clear exactly where to draw the boundaries, but this sort of problem affects the individuation of other events and objects just as much.

Let us agree, then, that every action is an event. We may also agree that not every event is an action. The Big Bang, for instance, was an event, but aside from any supernatural possibilities, it was not an action. So this naturally raises the question of what distinguishes events that are actions from those that are not. Now it may well be that there is no interesting answer to be given to this question, but as we have already noted, Davidson assumes that there is a fairly definite subclass of events that are actions. Let us now consider this assumption.

In the case of the Big Bang, at least part of the reason why it does not count as an action is obvious—no one did it. For whatever else an action may be, it is clearly something that an agent does. If actions are events, then only those events that are done by an agent are actions. We have, then, at least a necessary condition for distinguishing actions from other events, but is it
also sufficient? Whenever an agent performs an action he does something, but does everything an agent do count as an action?

Many philosophers think the answer to this last question is no. Lawrence Davis, for instance, suggests that if a person is standing still, or hoping for the best, then in each case he is doing something, but not performing any action (Davis 1979, p.4). However, it is doubtful that Davis's examples establish the intended point. Thus, imagine an actor, whose job it is to stand still at a certain point in the performance of a play. It is clear that in these circumstances, the actor would be performing an action in standing still. Hoping for the best, on the other hand, is like praying for something, and this would seem to count as an action in at least some circumstances.

Evidently, what encourages Davis in the thought that standing still or hoping for the best are not actions is that they do not involve any bodily movement, but it is a mistake to think that an action must involve a bodily movement. Thus, if I calculate the square root of a number, recite a poem, or compose a song, then in each case I perform an action, but the action need not involve any bodily movement—at least no obvious bodily movement. Like Davis, Davidson also seems to identify actions with bodily movements, but at the same time, he agrees that in taking the square root of a number, or even in making a decision, an agent performs an action. Davidson reconciles this apparent conflict by interpreting the idea of a bodily movement extremely generously; indeed, he writes that "the generosity must be open-handed enough to encompass such
movements as standing fast, and mental acts like deciding and computing" (Davidson 1971a, 49). Suffice it to say that, on this conception of a bodily movement, anything an agent does involves a bodily movement.

The present question is whether an agent's actions are a subset of the things he does—his doings. If we consider all of the things that an agent does, can we distinguish between those that are actions and those that are not? Here we need to be clear about whether we are considering the types of the things that an agent does, or the individual things that an agent does. Davis evidently approaches the question at the level of types. Thus, he writes that "people hiccup, bleed, tremble, shudder, stumble, and fall. They give birth, grow taller, fall asleep, catch cold, recuperate, faint and die. Each of these is something people may be said to "do," yet none is an action." (Davis, 1979, p.4).

Each of the things that Davis mentions are types of things that people do. But it is clear that Davis cannot baldly say that none of them are actions. In fact, all of them are actions, in the sense that at least some of the individual instances of these types of things are actions. It is tempting to suppose no instance of bleeding counts as an action, but it is easy to see that this is not true of any of the other verbs that Davis mentions. Again, if we imagine an actor in a play, we can imagine circumstances in which his shuddering, trembling, or falling down would count as an action. Even in the case of bleeding, it is not difficult to imagine circumstances in which such an event would count as an
action. Thus, a woman might acquire the ability to control her own internal bleeding, perhaps by taking certain drugs, and bleed only when she wants to bleed.

It is an error, according to Davidson, "to think that verbs may be listed according to whether they do or do not impute agency to a subject or object" (Davidson 1071a, p.44). The foregoing considerations reinforce this point, and suggest that there is no plausible distinction to be drawn, amongst the types of things that people do, between those that are actions and those that are not. However, it does not follow from this that everything an agent does is an action. For it may be that there is a distinction to be drawn between our actions and our doings, but only at the level of the particular things we do.

Davidson, for one, believes that there is such a distinction, and that not everything that an agent does is an action. According to Davidson, an event is an action only if someone does it intentionally. However, for reasons having to do with the referential opacity associated with attributions of intentional action, Davidson expresses this criterion as follows: "a man is the agent of an event if and only if there is a description of what he did that makes true a sentence that says he did it intentionally" (Davidson 1971a, p.46).

The left-hand side of Davidson’s bi-conditional reads "a man is the agent of an event," which is Davidson's way of saying that a particular event, performed by some man, is an action. But men, of course, are not the only ones to perform actions; in addition to
other humans, and other animals, there may be things other than animals that also perform actions. We will consider some examples of this below, but in the meantime, we can avoid being sexist and anthropocentric by using the generic term "agent" for the sorts of things that are capable of performing actions—whatever they may be. Thus, let us restate the left-hand side of Davidson's bi-conditional to read as follows: An event e is an action of agent X.

Correspondingly, we might reword the right-hand side of Davidson's bi-conditional as follows: There is a description of e which makes true a sentence that says that X did it intentionally. However, it is misleading to suggest that a description of an event can make a certain sentence true. Evidently, what Davidson means is that there is a description of the event which, when substituted within a certain open-sentence, results in a true sentence. Thus, let us reformulate Davidson's criterion for distinguishing actions from other events as follows: An event e is an action of agent X if and only if there is a description of e which, when substituted for the variable "t" in the open-sentence "X did t intentionally," results in a true sentence. It is interesting to note, in passing, that one could express Davidson's criterion symbolically, given a substitutional interpretation of quantificational sentences.

Now the main reason why Davidson reverts to descriptions of events in distinguishing actions from other events has to do with the fact, as he puts it, "that one and the same action may be correctly described as intentional (when described in one way) and not intentional (when described in another)" (Davidson 1969a,
Thus, Oedipus intentionally struck the rude old man, but unintentionally struck his father, even though the rude old man was his father. Given this identity, and assuming that Oedipus struck the man once in this context, it follows that the striking of the old man by Oedipus was identical to the striking of his father by Oedipus. However, according to Davidson, the striking of the old man by Oedipus was intentional and the striking of his father was not. In order to avoid the conclusion that one and the same action is both intentional and not intentional, Davidson suggests that we understand sentences like the foregoing as containing an implicit reference to the descriptions under which the actions are given. Thus, Davidson writes that "the striking of the old man was intentional under one description but not under another" (Davidson 1971b, p. 195).

Davidson uses the technical notion of an action under a description in a number of different contexts, but in each case to make explicit the fact that a certain sort of action-sentence is referentially opaque (Davidson 1971b, p.194). In the present context, he uses it to draw attention to the referential opacity that affects attributions of intentional action. However, this use of the technical notion is certainly misleading. For referential opacity is a phenomenon concerning singular terms, or the positions occupied by singular terms, and while our ordinary attributions of intentional action are indeed referentially opaque, at least in one sense, they do not typically contain any singular terms that purport to refer to an action.
Thus, the sentence "Oedipus intentionally struck the old man" is referentially opaque, in one sense, since it does not entail "Oedipus intentionally struck his father," given the identity of the old man and Oedipus's father. And the sentence "Hamlet intentionally killed the man behind the arras" is referentially opaque, in one sense, because it does not entail "Hamlet intentionally killed Polonius," given the identity of Polonius and the man behind the arras. None of the four sentences just mentioned contain any terms that purport to refer to an action, and so it is misleading to draw attention to the sort of referential opacity that affects these sentences by speaking of the description under which the action is given. The referential opacity that affects these sentences relates, rather, to the descriptions under which the direct object of the action is given.

Of course, Davidson tends to employ slightly different sentences when explaining the idea that an action is intentional under one description and not under another. Thus, he uses sentences like "It was intentional of Oedipus that there was an event that was his striking of the old man at the crossroads" (Davidson 1971b, p.195). Unlike the related sentence in the previous paragraph, this last one does contain a definite description of an action. However, aside from justifying a use for the technical notion of an action under a description, it is not clear why Davidson reverts to these awkward sentences—and one must be careful about using them in this context. In particular, one must be careful not to conflate sentences like "Oedipus struck the
old man" and "The striking of the old man by Oedipus occurred." For
as we saw in the last chapter, Davidson holds that the latter
sentence entails but is not entailed by the former. And given this
semantic distinction between the two sentences, it is clear that we
cannot substitute one for the other within ascriptions of
intentional action. In other words, if Davidson's views on logical
form are correct, then it does not follow from the fact that
Oedipus intentionally struck the old man that it was intentional of
Oedipus that the striking of the old man by him occurred.

The danger involved in using the notion of an action under a
description to draw attention to the referential opacity that
affects ascriptions of intentional action is that it encourages the
mistaken idea that ascriptions of intentional action contain
definite descriptions, or other singular terms, that purport to
refer to actions. Most ascriptions of intentional action do not.
And given that there is no need for any technical notion to draw
attention to the referential opacity that affects these sentences,
one might reasonably dispense with the notion of an action under a
description altogether. However, there is at least one use for this
technical notion, which is not to clarify our ordinary talk of
intentional action, but rather to facilitate talk of Davidson's
criterion for distinguishing actions from other events.

According to Davidson, in order to determine whether or not
any particular event is an action we have to ask whether or not
there is a description of that event which results in a true
sentence when substituted for the variable "t" in an open-sentence
of the form "He didn't intentionally." In fact, if we were to be precise on matters of logical form, the question would be more complicated than this, but even as it stands this is an awkward question to have to repeat. So for the purposes of evaluating Davidson's criterion, it will be convenient to revert to the notion of an action under a description. Thus, instead of asking the question that Davidson's criterion literally requires, we can ask, of any given event, if there is a description under which it is intentional, or under which it was done intentionally. Given this elliptical locution, let us now go on to consider Davidson's criterion in more detail.

Davidson uses the following consideration to justify his proposed criterion. We must distinguish, he says, "between three situations in which it is correct to say that I spilled the coffee: in the first, I do it intentionally; in the second I do not do it intentionally but it is my action (I thought it was tea); in the third it is not my action at all (you jiggle my hand)." (Davidson 1971, p.45). If there is a difference between these last two cases of spilling the coffee, the difference is not a matter of whether the agent spills the coffee intentionally, since in both cases it is false to say that the agent intentionally spilled the coffee. According to Davidson, "the difference seems to lie in the fact that in one case, but not in the other, I am intentionally doing something" (Davidson 1971a, p.46). Having already decided that the second case is, and the third case is not, an expression of agency, Davidson concludes that an event is an action only if it is done
intentionally. But since an event is done intentionally, according to Davidson, only under a description, he expresses this point by saying that an event is an action only if there is a description under which it is done intentionally.

I think that Davidson's example fails to support his proposed criterion. Based on the minimal information that Davidson provides, there is really no way to tell whether or not the agent performs an action in the third case of spilling the coffee. It is tempting to suppose, as Davidson asserts, that the agent in this case does not perform an action, but it is just as tempting to suppose that the agent also does not spill the coffee. To see this, let us briefly consider what Davidson, following Joel Feinberg, calls the **accordion effect**. The accordion effect states that an agent is the agent of whatever events are caused by his actions. Davidson gives the following example:

A man moves his finger, let us say intentionally, thus flicking the switch, causing a light to come on, the room to be illuminated, and a prowler to be alerted. This statement has the following entailments: the man flicked the switch, turned on the light, illuminated the room, and alerted a prowler. Some of these things he did intentionally, some not; beyond the finger movement, intention is irrelevant to the inferences, and even there it is required only in the sense that it must be intentional under some description. (Davidson 1971a, p.53)

By stipulating that intention must be present in the first thing that the agent does, Davidson merely means to indicate that the first event in the causal chain must be an action; for the accordion effect applies only to the causal consequences of actions.
The accordion effect applies only to those events that are caused by an action; it does not necessarily apply when the action is a mere causal factor in the production of the event. "Burning down a forest," as Feinberg has observed, cannot be ascribed to a camper whose campfire is suddenly scattered by unprecedented hurricane winds, even though but for his action of making a fire, the forest would never have burned (Feinberg 1964, p.159). Nor does the accordion effect necessarily apply when the causal chain from action to event runs through the intentional actions of another agent. Thus, as Davidson points out, if Jones intentionally causes Smith intentionally to shoot Clifford to death, we certainly won't conclude that Jones shot Clifford, and we may or may not say that Jones killed Clifford (Davidson 1971a, p.53).

In this last case, it is probably be misleading to say, without qualification, that Jones’s action—whatever that was—caused either the shooting or the killing of Clifford. Rather, it is more likely that Jones did something that was a causal factor in the shooting or the killing. If so, then this would explain why the accordion effect does not apply in this sort of case. In any event, whether or not this is right, it is sufficient to note that the accordion effect does not necessarily apply when the causal chain runs through the intentional actions of another agent.

Given this brief sketch of the accordion effect, let us return to consider Davidson’s third case of spilling the coffee. Based on the information Davidson gives, we are led to imagine that one agent, whom we may call agent X, is holding a cup of coffee, and
that another agent, agent Y, causes the coffee to spill by jiggling the arm of X. At first glance, this looks like the sort of case in which the accordion effect does apply; for it appears that the coffee that agent X is holding spills as the direct result of what agent Y does. If so, then we can infer that agent Y spilled the coffee.

However, if this is right—if agent Y spills the coffee in this case—then the example provides no justification at all for Davidson's proposed criterion; for we do not need to endorse Davidson's criterion in order to explain why it is false to say that agent X performed an action in this case. A simpler explanation is available. In the scenario just sketched, the reason why it is false to say that agent X performed an action in spilling the coffee is not that there is no description under which he did it intentionally, but rather that there is no description under which he did it. Agent Y did it.

Now there is a difference between saying, on the one hand, that agent Y did something that caused the coffee that agent X was holding to spill, and, on the other hand, that agent Y did something that caused agent X to spill the coffee. The latter sentence does, and the former does not, entail that agent X spilled the coffee. And given the minimal information that Davidson provides in his description of the example, there is really no way to determine which locution is more appropriate in this case—the one that implicates agent X, or the one that does not. Davidson’s example relies on the claim that does implicate agent X, but he
gives us no reason for thinking that it is the right one. In order to determine whether or not it is, it is useful to consider a couple of different scenarios.

In the first case, imagine that instead of jiggling the hand of agent X, agent Y takes a two-by-four and smashes it across the arm of X. In this case we certainly would not say that agent X spilled the coffee. And for precisely the same reason it would be incorrect to say that agent Y did something that caused agent X to spill the coffee. In the second case, imagine that agent Y gently nudges the arm of X, that this causes the coffee to sway back and forth, and in an effort to stop the coffee from swaying, agent X overcompensates and spills the coffee. In this case, we can say that agent X spilled the coffee, and, arguably, that agent Y did something that caused agent X to spill the coffee. However, this sort of case is also of no assistance to Davidson in justifying his proposed criterion; for in this case there seems to be no reason to deny that agent X performed an action in spilling the coffee. Indeed, on Davidson’s criterion, agent X did perform an action in spilling the coffee, since he intentionally did something that caused the coffee to spill—namely, he intentionally moved his hands in such a way as to try to stop the coffee from spilling.

We may summarize the foregoing observations as follows. As we fill in the details in any case in which an agent might be said to have spilled a cup of coffee, the clearer it becomes that the spilling of the coffee was not an action of the agent’s, the clearer it becomes that the agent did not spill the coffee;
conversely, the clearer it becomes that the agent did spill the coffee, the clearer it becomes that he did perform an action in spilling the coffee. Due to the paucity of details in Davidson's example, there is no way to determine either whether the agent in question did spill the coffee or whether he did perform an action, but as we sketch in the details of different possible scenarios, it seems that these two things stand or fall together. Thus, I believe that the example Davidson provides fails to justify his proposed criterion for distinguishing actions from other events.

Davidson does not provide adequate justification for his criterion, but it remains to be seen whether or not the criterion is justified. Davidson's criterion for distinguishing actions from other events is, in effect, a definition of the term "action." But whatever else one may care to say about this definition, it is clear that it does not capture the ordinary usage of the term "action"; for we often use the term "action" to describe the behaviour of things that are not even capable of forming intentions. In chemistry, for instance, we say things like "if gold is immersed in hydrochloric acid, no chemical action takes place," or "with similar acid, the action on pure steel is hardly perceptible." Indeed, the term "action" is widely used throughout the physical sciences. Thus, scientists speak of "the action of ancient glaciers on the rocks and mountains," "the actions of the small particles of bodies on each other," and "the action, or attractive force, of the Sun and the Moon, on various objects" (Murray 1989, p.127).
Of course, any definition of the term "action," is bound to be at least partly stipulative, and so it may be thought that Davidson's proposed definition captures at least one sense of the term "action," which we might call "human action." However, I think that even this suggestion ought to be rejected. Thus, consider the following case. Last night, while I was sleeping and completely unconscious, I uttered the sentence "God is dead." Now is there any description of what I did under which it would be true to say that I did it intentionally? Clearly not; for I was not in the appropriate frame of mind to have done anything intentionally. It follows, then, on Davidson's definition, that my utterance was not an action. But is the situation really this clear? Surely there is at least some motivation for calling my utterance an action, albeit unintended and unintentional. And what if I got up and started walk about in my sleep—should we say that I did something, but that I did not perform any action? I see no reason at all for saying this, and if my sleep-walking or sleep-talking are actions, then they are surely human actions.

There is, moreover, a further sort of counterexample that might be raised against Davidson's criterion. Thus, imagine that Hamlet stabs the man behind the arras, and that he stabs him sixteen times. Now each of these stabbings is an event, and so it makes sense to ask of each of these events whether it is an action. Consider, for instance, the thirteenth stabbing. Is this an action? According to Davidson's criterion it is an action if there is a description of the event under which Hamlet did it intentionally.
But it may well be false to say that Hamlet did the thirteenth stabbing of the man behind the arras intentionally. Indeed, it surely would be false to say this if Hamlet had intended to stab the man behind the arras only twelve times, and had simply lost count after the tenth. So what is needed, if Davidson's criterion is correct, and if the thirteenth stabbing is an action, is some description that uniquely picks out the thirteenth stabbing, and at the same time results in a true ascription of intentional action. However, I see no reason for thinking there must be such a description. The problem raised by this sort of example stems from the fact that ascriptions of intentional action do not generally refer to particular events. And it is partly for this reason that it seems to be a mistake to use such ascriptions in order to distinguish amongst particular events.

Our most basic understanding of what it is to perform an action, as any dictionary will reveal, is just what it is to do something. An action is a doing, a deed, a thing that is done. And if, as we have already agreed, an action is an event, then we may say that an action is an event that is done or performed by someone or something. Insofar as we are concerned only with human actions, we may restrict our attention to those events that are performed by human beings. Obviously some of the events that can be attributed to someone count as actions more clearly, or in a more robust sense, than others, but given the variety of contexts in which the term "action" is used, it is highly doubtful that there is any interesting principle at work in distinguishing amongst the various
senses. Davidson, as we have observed, remarks that the term "action" does not very often occur in ordinary speech, but I think this is wrong. The term "action" is ubiquitous; it is rather the adverbs of intention that are reserved for portentous occasions.

In what follows, then, we will not make use of Davidson’s definition of action. However, this terminological decision will not affect any of the conclusions to be drawn below. Where there is any suggestion that the issue being discussed depends upon a particular definition of action, it will be made explicit. But it is noteworthy just how little in the philosophy of action depends upon any particular definition of action. The questions with which philosophers tend to be primarily concerned, questions such as "What is an intentional action?" can be formulated without recourse to the term "action" at all. Thus, to raise this last question, for instance, is just to ask for an account of what it means to do something intentionally. We will take up this very question in the next chapter, but prior to this there is another matter to consider.

II

According to Davidson, the natural language embraces an ontology of actions and events in addition to ordinary objects; collectively, they are the concrete particulars to which we refer and quantify over in ordinary and scientific discourse. That they are particulars, however, does not detract from the fact that they can
be composite. Thus, a particular object, such as my car, consists of a number of other objects, such as an engine, a steering wheel, some seats and tires. Similarly, individual actions and events can be composite. World War II, for instance, is a particular event, but also consists of a number of other actions and events, such as the invasion of Poland by the German troops, the landing of the Allied forces in Normandy, and the bombing of Hiroshima. Even simple actions, like my writing of the word "action," consist of other individual actions. First I write the letter "a," then I write the letter "c," and so on.

Individual actions, then, can consist of other actions and events, and where these constitutive actions and events stand in causal relations to one another, we can speak of the causal complexity of the action. Indirectly, we have already encountered the causal complexity of certain actions in discussing the accordion effect. We saw in the last section that if I flip a switch, and this causes the room to be illuminated, then it follows that I illuminate the room. Here we can say that I illuminated the room by flipping the switch, and the "by" in this last sentence is a mark of causal complexity. The sentence asserts the existence of at least one action that is causally complex in the sense that it consists of one action that was a flipping of the switch by me, and at least one event caused by that action.

Now the causal complexity of actions is an issue that has misled many philosophers, including, I think, Davidson. But without a proper understanding of what it means to perform one action by
means of another, we cannot begin to understand what it means to perform one action as a means of performing another. And since this latter issue is of paramount importance in the philosophy of action, and one with which we will be concerned in the following chapter, it is necessary that we first consider the question of what it means to perform one action by means of performing another. This is the issue with which we will be concerned in the remainder of the present chapter.

In some cases, we can tell that an action is causally complex merely in virtue of the way it is described; for the complexity is implicit in certain verbs of action. Davidson calls such verbs causal verbs, and cites "killing" and "thanking" as examples. Thus, he writes that:

\[
x \text{thanks } y \text{ if and only if there are two events, call them } e \text{ and } e', \text{ such that } x \text{ is the agent of } e, \text{ } e' \text{ is a being thanked by } y, \text{ and } e \text{ caused } e'. \text{ In other words, } x \text{ did something that caused } y \text{ to be thanked}"  \text{ (Davidson 1987, p.38).}
\]

Similarly, to kill someone is to do something that causes that person's death. Joel Feinberg adds "opening," "closing," and "rescuing" to the list, and I think we can say that most transitive verbs count as causal verbs. Actions that are described by causal verbs are actions that an agent performs by means of performing another action first, and which are therefore causally complex. However, as Arthur Danto pointed out long ago, not every action can be complex in this sense. According to Danto:

\[
\text{It may be argued upon the penalty of infinite regression that if, in order to perform an action, we must always perform some other action as part of performing it, that we could perform no actions whatever. So if we perform...}
\]
any actions at all, we must perform some which do not have further actions of ours amongst their component parts. These I designate basic actions. (Danto 1969, p.109)

In his original argument, Danto also argued that there are actions, from which it follows, given his first argument, that there are basic actions. He then argued that not every action is a basic action, from which it follows that there are non-basic actions.

Danto has given a number of different definitions of a basic action, but in each case it is clear that the distinction between basic and non-basic actions is a distinction that he draws at the level of types of actions. Thus, in his original article, Danto writes that "B is a basic action of a if and only if (i) B is an action and (ii) whenever a performs B, there is no other action A performed by a such that B is caused by A" (Danto 1963, p.435). In a later article he gives a slightly different definition: "if B is a basic action of a man m, then m performs B simply and directly, in the sense only that he need do nothing distinct from B, itself an action, as part of his doing B" (Danto 1969, p.109). And elsewhere he writes that "moving an arm is one of the standard basic actions" (Danto 1965, p.144). Joel Feinberg draws a similar distinction between "causally simple" and "causally complex" actions, but it is clear that Feinberg also draws the distinction amongst types of actions. According to Feinberg, "smiling and frowning are simple actions, and so are raising one’s arm and shutting one’s eye" (Feinberg p.145).

Danto and Feinberg both cite raising an arm as a basic or causally simple action. However, Davidson has observed that while
raising an arm is usually done without doing anything else, this is not always the case. Thus, he writes that "if I rig up a pulley and rope, I can raise my paralyzed left arm by pulling on the rope with my right arm, and in this case I do, of course, raise my left arm by doing something else" (Davidson 1987, p.37). Davidson concludes that "all we can say for sure is that not everything we do can be done by doing something else, or nothing would ever get done" (Davidson 1987, p.37). I think that what Davidson means to say is that not everything we do must be done by doing something else, but in any event, Davidson too is speaking of the types of things that one does, rather than particular doings.

Davidson’s observation shows that at least some cases of "moving an arm" are cases in which the agent moves his arm by doing something else. It follows from this that "moving an arm" cannot be a basic action according to the definition that Danto sets out. However, the same point that Davidson makes with respect to "moving an arm" can be made with respect to each of the other types of actions that Danto and Feinberg list as basic or causally simple actions. Thus, although it seems that one does not normally contract one’s facial muscles by doing something else first, it is easy to imagine circumstances, similar to those that Davidson describes, in which one could contract one’s facial muscles by doing something else. And even if one can calculate a sum directly, it is clear that one can doing it by doing something else, such as pushing buttons on a calculator.
Should we say, then, that there are no basic actions? Carl Ginet believes that there are basic actions, as Danto defines the term, but that the only actions that are truly basic are volitions. A volition, according to Ginet, is a willing or trying, on the part of the agent, to do something. Thus, Ginet writes that:

My turning on the light was not basic because it consisted in my moving the light switch in a certain way and that action causing the light to go on: I turned on the light by moving the switch. And my moving the light switch was not basic either, because I did that by exerting force with my hand and fingers in a certain way, thereby causing the switch to move. Nor was my exerting force with my hand and fingers basic, for I did that by means of a mental activity of volition, of willing to exert with my arm and hand in a certain way, which is a basic action. (Ginet 1990, p.117)

Now it is clearly not the case that an agent wills to do, or tries to do, everything that he does. Typically, an agent does not try to do that which he does unintentionally. Nor does an agent try to do everything that he does intentionally. However, even if we grant, in the case in which an agent turns on the light by moving his hand in a certain way, that he exerts force with his hand and that he wills or tries to exert force with his hand, it does not follow that the agent exerts force with his hand by trying to exert force with his hand. Nor is it clear that this is ever the case. Moreover, even if we grant that this is possible, it still does not follow that volitions are basic actions in the relevant sense. For if an agent can exert force with his hand by trying to exert force with his hand, then why can he not try to exert force with his hand by trying to try to exert force with his hand? Clearly, a further story needs to be told.
It seems to be the case, then, that no matter what type of action we consider, we can imagine circumstances in which an agent performs such an action by performing some other action. Should we conclude, then, that there are no basic actions? Well, if we define a basic action as Danto does, then I think that the answer is yes. For he defined a basic action B as one such that whenever an agent performs B, there is no other action A such that the agent performs B by means of A. However, we can preserve the idea behind the basic/non-basic distinction—the idea that some actions are causally simple and some are causally complex—by applying the distinction to particular actions instead of types of actions. Thus, we can say that an action is basic if the agent performs it directly, and not by means of performing some other action first. Now in this sense of the term, not only is it the case that there are basic actions, but there must be; for otherwise, there could be no actions at all. Thus, Danto was right in thinking that there must be basic actions; he was wrong only in giving the definition that he gave for them.

In order to avoid further confusion, let us jettison the terms "basic" and "non-basic," and instead speak of causally simple and causally complex actions. A particular action is causally simple if an agent performs it directly, and not by means of some other action; a particular action is causally complex if an agent performs it by means of performing some other action. We can still speak of certain types of actions as being "causally complex," in the sense that every token of that type is causally complex. Thus, verbs like "killing" are causal verbs because every action that is
a killing has a causally complex structure. On the other hand, we must be careful not to draw the distinction at the level of types. For if Davidson is right in claiming that "raising an arm is usually done without doing something else, but not always," then any particular arm-raising can be either causally simple or causally complex.

Now Davidson agrees that there must be causally simple actions, which he calls primitive actions. Thus, he writes that:

Not every event we attribute to an agent can be explained as caused by another event of which he is agent: some acts must be primitive in the sense that they cannot be analyzed in terms of their causal relations to acts of the same agent" (Davidson 1971a, p.49).

But while some actions are primitive, it is clear that not all actions are. Again, if I raise my left arm by pulling on the rope with my right arm, then my raising of my left arm evidently has a causally complex structure: it consists of the act of pulling on the rope with my right arm and the event or sequence of events that are caused by this one action.

Davidson, it seems, would agree. And he would also agree that any action that can be described by a causal verb, such as "killing" has a causally complex structure. However, in "Agency," Davidson makes the curious remark that every action is primitive. The following passage contains his explanation for this remark:

To say that all actions are primitive is merely to acknowledge, perhaps in a misleading way, the fact that the concept of being primitive, like the concept of being intentional, is intensional, and so cannot mark out a class of actions. If an event is an action, then under some description(s) it is primitive, and under some description(s) it is intentional. (Davidson 1971a, p.61)
The above passage contains two closely related ideas that need to be distinguished. If we are to use the technical locution of an action under a description at all, there are two questions that might be raised with respect to primitive actions; the first is whether one and the same action can be primitive under one description and not primitive under another; the second is whether every action has a description under which it is primitive. Davidson believes that the answer to both questions is yes. However, it seems to me that he is wrong on both counts. Let us now see why.

What is at stake here is a certain view concerning the individuation of action. According to Davidson, it is a mistake "to suppose that when I close the door by moving my hand, I perform two numerically distinct actions" (Davidson 1971a, p.56). To put it crudely, Davidson's own view is that if I close the door by moving my hand, then I perform one action of which two descriptions have been given. And while Davidson does say things to this effect, this can only be a rough formulation. For we have already seen that sentences like "I closed the door" and "I moved my hand" do not, according to Davidson, refer to particular actions at all; rather, they assert the existence of an action of a certain sort. And for this reason, we cannot infer from the truth of either sentence how many actions have been performed.

Similarly, we cannot infer how many actions have been performed from the truth of a sentence like "I closed the door by moving my hand." I may have closed the door any number of times,
moved my hand any number of times, and closed the door by moving my hand any number of times. All that follows from the sentence "I closed the door by moving my hand" is that there is at least one action that was a moving of my hand by me, and at least one action that was a closing of the door by me. We can now express Davidson's curious ontological view as follows: if I close the door by moving my hand, then there must be at least one action that is both a moving of my hand by me and a closing of the door by me.

The ontological view under discussion is of course more general than this. And while Davidson does not give a general statement of the view, it is relatively clear that he does endorse a thesis to the effect that if an agent performs at least one action of type A by performing at least one action of type B, then there must be at least one action that is of type A and B. This is what we will call Davidson's identity thesis concerning causally complex actions.

In order to see how this thesis leads to the claim that one and the same action can be primitive under one description and not primitive under another, consider again the case in which I close the door by moving my hand. And for the sake of discussion, let us assume that I move my hand just once, and that I close the door just once. Now if Davidson's identity thesis is correct, then there is one action that is both a moving of my hand by me and a closing of the door by me. But if we suppose that while I closed the door by moving my hand, I moved my hand directly, and not by means of doing anything else, then it follows that one and the same action
is primitive and non-primitive. Described as "the moving of my hand by me" it is primitive, and described as "the closing of the door by me" it is non-primitive.

Now in order to see what is wrong with Davidson’s identity thesis, we need to consider a couple of different cases. In "Actions, Reasons, and Causes," Davidson gives the following example: I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home. We are to imagine in this example that I turn on the light by flipping the switch, that I illuminate the room by flipping the switch, and that I alert the prowler by flipping the switch. Now Davidson claims that "I need not have done four things, but only one, of which four descriptions have been given." And in a footnote attached to the remark Davidson adds that "we might not call my unintentional alerting of the prowler an action, but it should not be inferred from this that alerting the prowler is something different from flipping the switch" (Davidson 1963, p.4). But Davidson must mean here that although we may not call the alerting of the prowler an action, really it is; for if it were not, and Davidson’s identity thesis were correct, one and the same event is and is not an action.

Davidson suggests, in the case just described, that there are not four actions, but one action of which four descriptions have been given. However, this again is a stronger conclusion than Davidson needs, or is entitled to give. To put his identity thesis in a defensible frame, Davidson must say that there is at least one
action of which four descriptions have been given. In other words, if Davidson's identity thesis is correct, there is at least one action that has the following properties: it is a flipping of the switch by me, a turning on of the light by me, an illuminating of the room by me, and an alerting of the prowler by me.

We saw, in the last chapter, that there were two different types of quantificational analysis that one might provide for action-sentences, and I argued that the analysis that derives from Terrence Parsons is better justified than the one that Davidson himself provides. According to Parsons, a sentence like (1) should be analyzed as (1*).

(1) I flipped the switch.
(1*) (\exists x)(Flipping(x) & Of(x, the switch) & Agent(x, I)).

Similarly, Parsons would analyze (2) as (2*).

(2) I alerted the prowler.
(2*) (\exists x)(Alerting(x) & Of(x, the prowler) & Agent(x, I)).

Now if Davidson's identity thesis is true, and if it is true that I alerted the prowler by flipping the switch, then the following sentence is also true:

(3) There is at least one action that was a flipping of the switch by me, and an alerting of the prowler by me.

And on Parsons' analysis, (3) would be rendered as follows:

(3*) (\exists x)(Flipping(x) & Of(x, the switch) & Agent(x, I) & Alerting(x) & Of(x, the prowler) & Agent(x, I)).

But we can infer from (3*) that I flipped the prowler, and also that I alerted the switch. Clearly, Parsons account of the logical form of actions sentences is inconsistent with Davidson's identity thesis. Parsons takes this as evidence against Davidson's identity
thesis, and Davidson, on the other hand, would probably take this
as evidence against Parsons’ account of the logical form of action-
sentences. According to Davidson, sentences (1), (2), and (3)
should be analyzed as follows:

$(1^{**}) \exists x(\text{Flipping Of}(x, \text{the switch, me}).$
$(2^{**}) \exists x(\text{Alerting Of}(x, \text{the prowler, me}).$
$(3^{**}) \exists x(\text{Flipping Of}(x, \text{the switch, me}) \&$
\hspace{1cm} \text{Alerting Of}(x, \text{the prowler, me}).$

And on this analysis of (3), the undesirable entailments mentioned
above do not follow. However, this does not show that Davidson’s
identity thesis is correct, for we can construct a slightly more
complicated case, which establishes the same point as the last one
even on Davidson’s version of the quantificational analysis.

The following example derives from John Wallace. Imagine that
in playing a game of pool I hit the cue-ball in such a way that it
causes the eight-ball to go into the corner pocket and the seven-
ball to go into the side pocket. If so, then we can say that I hit
the eight-ball into the corner pocket and the seven-ball into the
side pocket. These are the sorts of entailments that the so-called
accordion effect describes. And if Davidson’s identity thesis is
correct, then these are all descriptions of the same action. More
precisely, if Davidson’s identity thesis is correct, then there
must be at least one action that was a hitting of the eight-ball by
me into the corner pocket, and a hitting of the seven ball by me
into the side-pocket.

Now on Davidson’s version of the quantificational analysis, a
sentence like (4) should be analyzed as (4*).
(4) I hit the eight-ball into the corner pocket.
(4*) (∃x)(Hitting Of(x, the eight-ball, me) & Into(x, the corner pocket)).

Similarly, Davidson would analyze (5) as (5*).

(5) I hit the seven-ball into the side pocket.
(5*) (∃x)(Hitting Of(x, the seven-ball, me) & Into(x, the side pocket)).

In the case just described, if Davidson’s identity thesis is true, then it follows that (6) is also true.

(6) There was at least one action that was a hitting of the eight-ball by me into the corner pocket, and a hitting of the seven-ball by me into the side pocket.

But it is clear that undesirable entailments follow from (6) regardless of which version of the quantificational analysis we choose. On Davidson’s account, (6) is symbolized as follows:

(6*) (∃x)(Hitting Of(x, the eight-ball, me) & Into(x, the corner pocket) & Hitting Of(x, the seven-ball, me) & Into(x, the side pocket)).

And from this we can infer that I hit the eight-ball into the side pocket, and also that I hit the seven-ball into the corner pocket, both of which are false.

There are two responses that Davidson could conceivably make to this in order to protect his identity thesis. One would be to change his account of the logical form of action-sentences so that sentences like (4) and (5) each contained a four-place predicate, and so that (6) contained two such predicates. This sort of analysis would preclude the undesirable entailments, but it would also preclude the desirable ones. Thus, analyzing (4) in terms of a four-place predicate leaves one without a structural explanation of the inference from (4) to "I hit the eight-ball."
It is no surprise, then, that Davidson makes a less radical response. His response is contained in the following passage:

There are, I think, a number of events involved here, but only one action. "Hit" may be irreducibly transitive (I’m not sure about this), but the hitter need not be an agent—one ball may hit another. So "I hit the eight-ball into the corner" should be analyzed "There were two events such that I was the agent of one, another was a hitting of the eight ball into the corner pocket, and the first event caused the second." Even if "e was an into-the-corner-pocket event" is an independent predicate of events, it is not a predicate of the action, but of an event that was caused by the action. (Davidson 1985b, p.240)

Davidson attempts to block the undesirable entailments that follow from (6) by ruling that (6) is false, and that (6) is false because (4) and (5) are false. According to Davidson, (4) is false, strictly speaking, because although there was a hitting of the eight-ball into the corner pocket, it was not done by me—I was not the agent. For precisely the same reason, Davidson would claim that (5) is false. And if (4) and (5) are false, then it follows that (6) is also false.

I think it is clear that Davidson’s response to this example is inadequate. Davidson denies that it is correct to say, in this case, that I hit the eight-ball into the corner pocket; instead, he maintains that we must say that I hit the cue-ball, which in turn caused the eight-ball to be hit into the corner pocket. But as long as the accordion effect applies here—and this would seem to be a paradigm case—the sentence that Davidson tries to paraphrase away is entailed by the very paraphrase he gives. If, as Davidson maintains, I hit the cue-ball, which in turn caused the eight-ball to be hit into the corner pocket, then it follows that I hit the
eight-ball into the corner pocket. This is precisely what the accordion effect tells us.

Davidson’s only recourse, it seems, would be to produce some explanation for why the accordion effect does not apply in this case. But if it does not apply in this case, it is doubtful that it ever applies—and this is highly implausible. For there is an abundance of linguistic evidence in support of the accordion effect. Indeed, if it is false to say, in the case at hand, that I hit the eight-ball, then one should ask why it is true to say that I hit the cue-ball; for this is something I did by moving my cue in a certain way, which caused the cue-ball to be hit.

Part of the confusion in the case under discussion may relate to the suspicion that I do not hit the eight-ball in the same sense in which the cue-ball hits the eight-ball. Thus, it may be thought that (4) is false, in a strict sense, because the verb "to hit" correctly applies only to what the cue-ball does to the eight-ball. This, I think, is a mistake, but for the sake of making a point, it may help to choose a different verb. Let us therefore reserve the verb "to hit" for what the cue-ball does to the eight-ball, and use the verb "to sink" to describe what I do to the eight-ball. For surely it makes sense to say, in this case, that I sunk the eight-ball in the side-pocket. Similarly, we can say that I sunk the seven-ball in the side pocket. However, I did both of these things by hitting the cue-ball in a certain way. Thus, if Davidson’s identity thesis is true, it follows that there is at least one action that was a sinking of the eight-ball by me into the corner-
pocket and a sinking of the seven-ball by me into the side pocket. And from this we can infer, on either version of the quantificational analysis, that I sunk the eight-ball into the side pocket and the seven-ball into the corner pocket. Once again, I think the correct conclusion to draw is that Davidson’s identity thesis is false.

Let us return to the two questions raised above. The second question is whether every action has a description under which it is primitive. If the foregoing is correct, then the answer to this question must be no. This can be demonstrated by means of the foregoing example. For the sake of discussion, let us again suppose that I hit the cue-ball just once, and that this caused both the eight ball to go into the corner pocket and the seven-ball to go into the side pocket. Now Davidson claims that there is at most one action in this case, but we have just seen that this is wrong. My sinking of the eight-ball into the corner pocket cannot be identical with my sinking of the seven-ball into the side pocket; for it is from this alleged identity that the undesirable entailments follow. But given that these two actions are distinct, neither can be identical with my hitting of the cue-ball; for if either was, both would be, in which case they would be identical with each other. Thus, there must be at least three actions here, rather than one.

Now let us suppose that while I hit the cue-ball by moving my hand in a certain way, I moved my hand directly, and not by means of doing anything else. Thus, we can say that by moving my hand in
a certain way, I sunk the eight-ball in the corner pocket, and the seven-ball in the side pocket. But the argument that was just used to show that neither my sinking of the eight-ball nor my sinking of the seven-ball is identical to my hitting of the cue-ball can also be used to show that neither of the first two actions is identical to my action of moving of my hand. Indeed, the conclusion will hold regardless of which action we consider as primitive in this sequence. For whatever action we choose, if it is identical with my sinking of the eight-ball, it must also be identical with my sinking of the seven-ball. Thus, neither my sinking of the eight-ball, nor my sinking of the seven-ball, is identical with whatever action is primitive in this sequence. Here, then, we have an example of a non-primitive action that is not identical to the primitive action by means of which it was performed.

This last result shows that not every action has a description under which it is primitive, and this in turn naturally leads one to wonder whether any non-primitive action can be identified with a primitive action. The very language that we use to characterize non-primitive actions suggests that such identifications are mistaken; for we say that a non-primitive action is one that an agent performed by means of performing another. Of course, this is not a decisive consideration against identifying any non-primitive action with a primitive action, but it does put the burden of proof on those who endorse these identities. And why exactly would one think that a non-primitive action is identical with the primitive action by means of which it was performed?
Davidson gives nothing like a general argument for this claim. What he does is argue, in particular cases, against the opposing view: the view that if an agent performs one action $a_1$ by means of performing another action $a_2$, then $a_1$ and $a_2$ must be numerically distinct. Still, Davidson draws general conclusions. Thus, he writes:

There are, I think, insuperable difficulties that stand in the way of considering these various actions, the primitive actions like moving a hand, and the actions in describing which we refer to the consequences, as numerically distinct. (Davidson 1971a, p.57)

But what are these difficulties? The only thing resembling a difficulty comes in Davidson’s discussion of an example in which the queen kills the king by pouring poison in his ear. We are to suppose here that the queen pours poison in the king’s ear by moving her hand in a certain way. Now Davidson claims that "this movement [the queen moving her hand in a certain way] was by itself enough to cause the death of the king—there was no point to any further action on the part of the queen" (Davidson 1971a, p.57). Since there was no need for any further action by the queen, Davidson concludes that there was no further action by the queen. But this is fallacious. It does not follow from the fact that the queen did not need to do anything else in order to kill the king that she did not do something else in killing the king. Evidently, Davidson misses this distinction. For he writes, "Is it not absurd to suppose that, after the queen has moved her hand in such a way as to cause the king’s death, any deed remains for her to do or complete?" (Davidson 1971a, p.57).
If there is any explanation for why Davidson misses this point, it seems that it would relate to his tendency to identify actions with bodily movements. Since the queen does not, in killing the king, do anything *with her body* besides moving her hand in a certain way, Davidson believes that she does not perform any other action than moving her hand in a certain way. But on the contrary, I think we should take this as further evidence against identifying actions with bodily movements; not only does it require us to stretch the idea of a bodily movement beyond recognition in order to accommodate mental actions, it also encourages certain ontological errors. Actions are not identical with bodily movements; rather, certain actions involve certain bodily movements.

Danto was one of the first philosophers to stress the importance of the distinction between causally simple and causally complex actions. At the same time, he made two errors that have added as much confusion as clarity to the subject; the first was to draw the distinction between causally simple and causally complex actions at the level of types of actions; the second was to characterize causally complex actions as actions that are caused by other actions of the same agent. It was only in subsequent discussion of the issue that Danto realized this second mistake, and re-characterized causally complex actions as actions that have other actions as causal *components*. It is crucial to appreciate the difference here. If the queen kills the king by moving her hand in a certain way, then the act of moving her hand caused the king’s
death, but it did not cause the killing of the king. Rather, the killing of the king by the queen consists of, among other things, the queen moving her hand in a certain way. It also consists of an event that this last action caused—namely, the king's death. Thus, the one action of killing the king, assuming she killed him once, consists of one action which was a moving of her hand in a certain way and one of its causal consequences—the king's death.

Once we are clear on the structure of a causally complex action, or in Davidson's terminology, a non-primitive action, and distinguish its causal components from its causes, the idea that any causally complex action is identical with one of its causal components should strike one as plainly false. There is no more reason to identify the queen's killing of the king with the movement of her hand than there is to identify my car with its steering wheel, or World War II with the bombing of Pearl Harbour. Such identifications simply ride roughshod over obvious part/whole distinctions. Thus, I think we must conclude that Davidson's second claim is also mistaken—that no action is both primitive, under one description, and non-primitive, under another.

Let us sum up the results of the foregoing discussion in diagrammatic fashion. The following diagram, D1, illustrates the sort of structure involved in a causally complex action:

\[
\begin{align*}
  a_1 & \rightarrow e_1 \rightarrow e_2 \\
  a_2 & \\
  a_3 & \\
\end{align*}
\]
In this case, $a_1$, $a_2$, and $a_3$, are individual actions, while $e_1$ and $e_2$ are individual events that are not actions; $a_1$ is a causally simple action, while $a_2$ and $a_3$ are complex actions, with $a_3$ being more complex than $a_2$. The arrows represent the direction of causation, where $a_1$ causes $e_1$, which in turn causes $e_2$. As an example of the sort of action whose structure D1 illustrates, consider my turning on a light by flipping the switch. Here, we can identify $a_1$ with the act of moving my hand in a certain way, $a_2$ with the act of flipping the switch, and $a_3$ with the act of turning on the light. Each of the last two actions are performed by means of $a_1$, and consist of $a_1$ in addition to the relevant events that are caused by $a_1$.

When philosophers speak of causally complex actions, they often focus on rather elementary cases, such as turning on a light by flipping a switch, or shooting someone by pulling a trigger. In conjunction with the assumption that every action is either causally simple or causally complex, this focus on the elementary cases has an unfortunate distorting effect. For many of our actions are very different than the act of pulling a trigger or shooting someone. Consider, for instance, what is involved in building a house, writing a novel, or climbing a mountain. One of the more important ways in which climbing a mountain differs from turning on a light is that there is no possibility of analyzing an action of this sort merely in terms of a causally simple action and some subsequent event, or series of events, that are caused by that action.
If I climb Mt. Everest, then my climbing of Mt. Everest would consist, not of one causally simple action and a series of ensuing events, but rather of a very large series of causally unrelated simple and complex actions. Thus, my action of driving in the first spike is neither a cause, nor an effect, nor a causal component, of my action of driving in the second spike. Rather, both actions are causally unrelated constituents of my climbing of the mountain. To distinguish the actions like climbing a mountain, from those like turning on a light, let us call the former aggregate actions. Thus, an aggregate action is one which consists of two or more causally unrelated actions. And so on this definition, while an aggregate action can be extremely complex, like climbing Mt. Everest, it can also be rather elementary, like writing the word "action."

Let us illustrate the sort of structure involved in an aggregate action by means of the following diagram:

\[ D2: \]

\[ \begin{align*}
    &a_5 \\
    &\{ a_2 \} \\
    &\{ a_4 \} \\
    &a_1 \rightarrow e_1 + a_3 \rightarrow e_2
\end{align*} \]

Once again, the a’s designate actions, and the e’s designate events that are not actions; \( a_1 \) and \( a_2 \) bear the same relation as their counterparts in \( D1 \), and \( a_3 \) and \( a_4 \) stand in the same relation as well. The only difference between \( D1 \) and \( D2 \) relates to \( a_5 \), which consists of at least two causally unrelated actions. As an example of the sort of action whose structure \( D2 \) illustrates, consider my
act of typing the number "11." Here, we can identify $a_1$ with my pressing the appropriate key on the keyboard, which causes a certain event to occur in the computer, and in virtue of which I perform $a_2$, or type the number "1." Similarly, $a_3$ and $a_4$ can be associated with my typing of the number "1" for a second time. And then $a_5$ is just the action that I perform by performing $a_2$ and $a_4$.

Now that we have considered what it means to perform an action indirectly, or by means of performing another, we may go on to consider what it means to perform an action as a means of performing another. These expressions are not equivalent, nor the converse of each other. It does not follow from the fact that an agent B-ed by A-ing, that he A-ed as a means of B-ing. To say that an agent A-ed as a means of B-ing is to say, among other things, that the agent intended to B by A-ing, but an agent can B by A-ing without intending to B at all. If Jones swings a bat that hits a ball that breaks a window, then Jones breaks the window by swinging the bat whether or not he intended to break the window. Thus, "by means of" is not to be confused with "as a means of." The latter is the mark of actions that are performed for a reason, which is the topic of the next chapter.
In her 1957 monograph entitled *Intention*, Elizabeth Anscombe raises the question "What distinguishes actions which are intentional from those which are not?" The answer she there provides is that intentional actions "are the actions to which a certain sense of the question "Why?" is given application," where the relevant sense of "Why," she goes on to explain, "is that in which the answer, if positive, gives a reason for acting" (Anscombe 1957, p.9). In other words, Anscombe's claim is that an intentional action is one that is done for a reason. Six years after the publication of *Intention*, Davidson published "Actions, Reasons, and Causes," in which he endorses Anscombe's claim and suggests that it gives the definition of intentional action (Davidson 1963, p.6). By calling it a definition, Davidson apparently accepts two claims—that every intentional action is done for a reason, and that every action that is done for a reason is intentional.

Now we have already seen that there are problems with the very notion of an intentional action. According to Davidson, if we take literally the idea that an action can be intentional, then we are forced to accept the result that one and the same action can be intentional and not intentional. Since this is a result that ought not be accepted, Davidson concludes that "intentional actions are
not a class of actions" (Davidson 1967a, p. 121). But if this is right, then it would seem that the question that Anscombe raises, as well as the response that she provides, are misconceived. Nevertheless, Davidson endorses Anscombe's response, and he does so by endorsing another idea of hers—that an action is intentional or unintentional only relative to a certain description of the action. According to Davidson, actions themselves are not intentional, but under certain descriptions they are. And if this is right, then one can say that Anscombe's question and response are misleading, but not necessarily misconceived.

However, it is doubtful that the idea of an action being intentional under a certain description can be taken any more literally than the idea of an action being intentional. For if Davidson's semantic analysis of ascriptions of intentional action is correct, then terms like "intentional" and "unintentional" function logically as predicates of propositions or utterances of sentences—whatever, that is, a predicate like "Smith believes that x" is true of. And so if Davidson's views on logical form are correct, then to say that an action is intentional under a certain description is, at best, an abbreviated way of saying something much more complicated.

However, for the purposes of evaluating Davidson's definition of intentional action, we need not enter into any of these complications. For the so-called definition of intentional action is really a definition of the term "intentionally," or an analysis of sentences of the form "X A-ed intentionally." Thus, the claim
that Davidson endorses can be expressed in terms of the following bi-conditional: X A's intentionally if and only if X A's for a reason. Let us call this bi-conditional Davidson's definition of intentional action. Now one of the two conditionals in Davidson's definition is relatively uncontroversial—namely, that if X A's for a reason, then X A's intentionally. However, I think that the other conditional is false, and, hence, that the definition is also false. Let us now see why.

We briefly noted at the end of the last chapter that an agent can perform one action as means of performing another action, or, in other words, that an agent can A as a means of B-ing. Thus, if I flip a switch in order to turn on the light, then we can say that I flipped the switch as a means of turning on the light. On the other hand, an agent can perform a certain action, not as a means of performing some further action, but as an end in itself. Thus, a man might step on a snail, not as a means of doing anything else, but simply for the sake of stepping on a snail. Here, the agent A's, not as a means of B-ing, but for the sake of A-ing. Now this last sort of case puts at least some pressure on the idea that if an agent A's intentionally, then he A's for a reason; for it would not be inappropriate to say of a man, who intentionally steps on a snail simply for the sake of stepping on a snail, that he did it "for no reason."

Davidson responds to this point by claiming that when we say "for no reason" in this sort of case we do not really mean that there is no reason, but rather that "there is no further reason, no
reason that cannot be inferred from the fact that the action was
done intentionally; no reason, in other words, besides wanting to
do it. (Davidson 1963, p.6). In fact, it is false to suggest, as
Davidson does, that we can infer from the fact that an action was
done intentionally, that the agent wanted to do it. However, this
is irrelevant to the present point. In the sort of case we are
presently considering, we may agree that the agent intentionally
A’s because he wants to A; the question is whether it is correct to
say, in such a case, that the agent A’s for a reason. Davidson
maintains that it is, and he does so by drawing a distinction, at
least implicitly, between two types of reasons for which an agent
might perform an action.

Let us employ the term extrinsic to designate the sort of
reason that Davidson describes as a "further" reason, and contrast
this with what we may call an intrinsic reason. An extrinsic
reason, on this conception, is the sort of reason that is given in
an explanation like "I flipped the switch because I wanted to turn
on the light." More generally, it is the sort of reason given by a
sentence of the form "X A-ed because he wanted to B." An intrinsic
reason, on the other hand, is revealed by a sentence like "He
stepped on a snail because he wanted to step on a snail." If we are
prepared to accept the idea that a sentence of the form "X A-ed
because he wanted to A" is explanatory—that it does give a reason
why the agent A-ed—then we can accept Davidson’s claim that even
those actions that are performed as ends in themselves are
performed for a reason.
The foregoing response protects Davidson’s definition of intentional action from one sort of counterexample, but other counterexamples remain. Whenever an agent A’s as a means of B-ing, the agent A’s for a reason. And if Davidson is right, it is correct to say that an agent A’s for a reason even when the agent A’s simply for the sake of A-ing. However, it seems that an agent can A intentionally without A-ing as a means of B-ing, or A-ing for the sake of A-ing. Alvin Melden gives the following example:

I pass the salt to my dinner companion not in order to please him or with any other motive or purpose in mind, but because I am polite. I act out of politeness, rather than for the sake of politeness. (Melden 1956, p.527)

Now there are any number of reasons why an agent might, on some occasion, pass the salt to his dinner partner. He may want to appear to be polite, he may simply want to remove the salt from his section of the table, or he may believe that their is poison in the salt and want to eliminate his companion. However, if Melden is right—and I believe he is—it may be the case that an agent intentionally passes the salt to his dinner partner for no reason at all. It may be that the agent has simply been trained to do this sort of thing, and that he does it without any deliberation or intention to do it.

It is important to note, in connection with this last point, that it does not follow from the fact that an agent passed some salt to his dinner partner without actually intending to do so, that he did not do it intentionally. There are plethora of examples in the literature to illustrate this sort of point; the following comes from Gilbert Harman:
Someone who foresees that his attempts to extricate himself from a tight parking spot will dent your fender may reluctantly go ahead, intentionally denting your fender in the process, without having aimed at or intended this in any way. (Harman 1986, p.136)

Assuming that the examples such as this one are sound, it does not follow from the fact that an agent A-ed intentionally that he intended to A; in other words, intending to A is not a necessary condition of A-ing intentionally. Accordingly, we cannot conclude from the fact that an agent passes the salt to his dinner partner without intending to do so that he does not do it intentionally. Nevertheless, there remains the question of whether an agent could intentionally pass the salt to his partner and yet do it for no reason.

There are relatively few circumstances in which one would be inclined to say of someone who passed a container of salt to his dinner partner that he did not do it intentionally. Perhaps we would say this in the event that the agent was unconscious while he did it, or in the event that he passed the salt thinking it was pepper. But let us assume that none of these sorts of abnormalities obtain, and that we are considering the range of possible cases—which ever they may be—in which it would be correct to say that the agent intentionally passed the salt to his dinner companion. The question is whether it could also be correct to say, in any of these cases, that the agent did it for no reason.

To ask whether it could be the case that the agent did it for no reason is not to ask whether it could be the case that there is no reason why he did it—in the sense of no explanation for why he
did it; for we may agree that there is sure to be some explanation for anything that an agent does. The question is rather whether it must be the case that there is a reason-explanation for why he did it—an explanation such as "He passed the salt because he wanted to appear to be polite," or even "He passed the salt because he wanted to pass the salt."

Melden suggests, in the passage recently quoted, that the correct explanation in some case might be something like "He passed the salt because he was polite." Of course, this remark is more like a gesture in the direction of an explanation, rather than a complete explanation, but the more complete explanation that it points towards is not a reason-explanation; rather, it is an explanation concerning facts associated with the agent’s training in etiquette. That there is an explanation of this sort does not rule out the possibility that there might also be a reason-explanation, but I see no reason for insisting that there must be a reason-explanation in this sort of case. Accordingly, I think that Melden presents a legitimate counterexample to the claim that if an agent A’s intentionally, then he A’s for a reason.

Much of our intentional behaviour, it seems, consists of actions that we have been trained or conditioned to perform, and which we do not necessarily perform in order to satisfy or promote any personal goals. And in addition to the actions that we are trained to perform are the actions that we have in some sense trained ourselves to perform. Here too we find counterexamples to Davidson’s definition of intentional action.
Thus, a man might be in the habit of going to bed at midnight, and it may be that he has gone to bed at midnight every night for the past twenty years. On some particular night, he may look at his watch, recognize that midnight is fast approaching, and straightaway go to bed. Now it would certainly be correct to say, in this sort of case, that the man intentionally went to bed at midnight, but there seems to be no inconsistency here with denying that there is any reason-explanation for why he did it. The explanation for why he did go to bed at midnight on the night in question might simply be that he has a habit or routine of going to bed at that time every night. And while it could be the case that he went to bed at midnight on this particular occasion in order to maintain his routine, or to promote whatever goal the routine might serve, I see no reason for insisting that this must be the case.

Consider yet another example. As I reflect upon these abstruse philosophical matters, I get up out of my chair and start pacing around the room. However, the situation is not one in which I feel like I need a stretch; nor is it the case that I believe that by pacing about the room I will clarify my thoughts. Rather, I simply get up at some point and starting walking about. Now it seems to me that in this sort of case, it would be misleading, and a distortion of the situation, to suggest that I am pacing about the room for a reason. And it would be just as misleading to suggest that what I am doing is in any way accidental or unintentional.

It may not follow from the fact that I am doing it, and not doing it unintentionally, that I am doing it intentionally. But we
can at least imagine circumstances in which it would be correct to say that I am pacing about the room intentionally. Imagine, for instance, that for a brief moment I become conscious of the fact that I am pacing about the room, but then quickly return to my philosophical reflection. From that moment on, we could certainly say that I am knowingly pacing about the room, and given that I am free to do otherwise, I think we could also say that I am intentionally pacing about the room. But there appears to be no contradiction in asserting this, while at the same time denying that I am pacing about the room for a reason.

I claimed above that one of the two conditionals implicit in Davidson's definition of intentional action is relatively uncontroversial. Indeed, I think we can accept the claim that whenever an agent A's for a reason, the agent A's intentionally. However, if the foregoing is correct, then the other conditional implicit in Davidson's definition is false, which is enough to show that the definition itself is flawed. This flaw is significant, and I think it is responsible for many of the misconceptions surrounding the question of what it means to act for a reason. We will return to this point below, but in the first place, let us consider Davidson's account of what it means to act for a reason.

II

The question of what it means to act for a reason is the subject of Davidson's most influential paper in the philosophy of
action—"Actions, Reasons and Causes." The thesis that Davidson seeks to establish in that paper is that reason-explanations are a species of causal explanation, and he does so by arguing for the claim that when an agent acts for a reason, the agent’s reason for acting is the cause of his action. The argumentative strategy here may be explained as follows. To say that an agent acted for a reason is to say that his action is explicable in terms of a reason-explanation, where a reason-explanation is the sort of explanation that states the agent’s reason for acting. So if the agent’s reason for acting is the cause of his action, when he acts for a reason, then it follows that a reason-explanation necessarily states the cause of the action in question. Hence, if the agent’s reason for acting is the cause of his action, when he acts for a reason, it follows that reason-explanations are, at least in some sense, causal explanations.

What we have so far called "reason-explanations" other philosophers have called by various other names. William Dray uses the title "rational explanations," as does Carl Hempel in his response to Dray (Dray 1957; Hempel 1962). Both Alvin Goldman and Lawrence Davis call them "reasons-explanations," Jon Elster calls them "rational-choice explanations," and Kathleen Lennon refers to them as "reason-giving explanations" (Goldman 1970; Davis 1979; Elster 1985; Lennon 1990). However, in "Actions, Reasons, and Causes," Davidson unfortunately designates these explanations rationalizations. The designation is infelicitous because a rationalization, as the notion is commonly understood, is not an
explanation at all. For to say of some putative explanation that it is a rationalization is to imply that the reason it states is not the reason for which the agent acted, even though it may well have been a reason he had for acting. Like Aesop’s fox who tried unsuccessfully to get the grapes, we rationalize our behaviour when we want to conceal, rather than reveal, either the real reason for which we acted, or the fact that we acted for no reason at all.

Accordingly, it would be a mistake to discuss Davidson’s thesis in "Actions, Reasons, and Causes" in precisely the terms in which he presents it; instead we will continue to employ the term "reason-explanations" to designate the sort of explanations that we provide when an agent acts for a reason. Davidson himself, it should be noted, has made the same terminological adjustment in the subsequent papers he has published on the topic, such as "Hempel on Explaining Action," and "Problems in the Explanation of Action."

"Actions, Reasons, and Causes" is regarded by many philosophers as the locus classicus for the so-called causal theory of action. However, it is doubtful that there is any single causal theory of action, or even any single claim that every causal theorist endorses. Indeed, there is great variation to be found in the literature, and even if we restrict our attention to what philosophers sometimes call Davidson’s causal theory, there are a number of different claims that fall under this rubric. One idea that we can surely count as central to Davidson’s version of the causal theory is embodied in the claim that when an agent acts for a reason the agent’s reason is the cause of his action. And since
this claim, if true, entails that reason-explanations are in some sense causal explanations, it is possible to identify Davidson’s causal theory with this last claim as well. In other words, we can treat Davidson’s causal theory of action either as stating a necessary condition of what it means to act for a reason, or as stating a necessary condition of the truth of any reason-explanation.

However, Davidson’s causal theory is also often construed as stating a necessary condition of what it means to act intentionally. Thus, consider the following passage from Fred Stoutland:

By a causal theory of action I mean one that makes two distinctive claims: 1) that behaviour is not intentional unless it is caused in the right way—unless it has certain specific kinds of causes; 2) that all acceptable explanations of intentional behaviour are causal, that when we explain an agent’s act by giving the reasons for his acting as he did, we specify the causes of his behaviour, so that reasons are causes of a certain kind. A causal theory sees these two claims as necessarily connected: an intentional act is just an act that has a certain kind of explanation, namely, one that is done for a reason, and an act is done for a reason only if it caused in a certain way. (Stoutland 1980, p.351)

Stoutland goes on in the passage from which this is extracted to identify Davidson as "the most influential advocate of a causal theory of action," and it is clear that in defining a causal theory as he does, Stoutland is thinking specifically of Davidson’s views. For the two claims that Stoutland holds to be essential to a causal theory of action, and which he says are necessarily connected, are necessarily connected only in virtue of Davidson’s definition of intentional action. However, a causal theorist need not accept
Davidson’s definition of intentional action, and if the results of the previous section are correct, no philosopher should.

Nevertheless, the observations made in the previous section do not show that reason-explanations are not causal explanations, or that an agent’s reason for acting is not the cause of his action when he acts for a reason. Thus, I think it is a mistake to conflate these various claims under the rubric of "Davidson’s causal theory of action," even if Davidson himself believes them to be synonymous. In what follows, then, we will distinguish, at least nominally, amongst the various claims associated with Davidson’s causal theory of action. In this, and the next, section we will be primarily concerned with what we might call Davidson’s causal analysis of acting for a reason. Of course, if Davidson’s definition of intentional action were correct, then his causal analysis of acting for a reason would at the same time serve as a causal analysis of what it means to act intentionally. But as his definition is not correct, the question of what it means to act intentionally must receive a different response. In the fourth section we will pursue this last point.

The central argument that Davidson provides in "Actions, Reasons, and Causes" proceeds by stating and defending two related theses, C1 and C2, concerning the sort of reason for which an agent acts, when he acts for a reason. Collectively, C1 and C2, are intended to provide at least a partial account of what it means to act for a reason. The second of these two theses is the more controversial of the two, but before we can properly evaluate it,
we need to consider the first of Davidson's two theses, which he introduces by way of the following passage:

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 (knowing, perceiving, noticing, remembering) that his action is of that kind. Under a) are to be included desires, wanting, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions, and public and private goals and values in so far as these can be interpreted as attitudes of an agent directed towards actions of a certain kind. (Davidson 1963, p.3)

According to Davidson, explaining why an agent did something is often a matter of citing a pair of attitudes that satisfy conditions (a) and (b), and Davidson calls this pair the primary reason why the agent performed the action. It is unfortunate, I think, that Davidson employs this technical term; for there is certainly no need for it, and by using it freely in two slightly different ways, it confuses more than it clarifies.

In particular, there is a distinction to be noted between reasons that merely justify a certain type of action, and reasons that not only justify a certain type of action, but which also explain why an agent performed such an action. In the former case, we might speak of "a reason for a certain sort of action," and in the latter case "a reason why an agent performed such an action." Now Davidson seems to want to use the technical notion of a primary reason to speak of the latter sort of reasons—those that are explanatory. However, Davidson not only speaks of "a primary reason why an agent performed an action," but also of "a primary reason for an action."
In the latter case, it is not always clear whether Davidson intends to designate any pair of attitudes that satisfy conditions (a) and (b)—in other words, any pair of attitudes that justify a certain type of action—or only those attitudes that not only satisfy conditions (a) and (b), but which also explain why an agent performed such an action. As there is no need for the technical term "primary reason," one might just dispense with it altogether; indeed, the best way to interpret many of the sentences in which the term "primary reason" occurs is simply to read it without the term "primary." However, for our purposes here, it is not possible to dispense with the technical term altogether; for Davidson states both of his two theses in "Actions, Reasons, and Causes" in terms of it. The best we can do is to clarify in context.

Before giving a more precise statement of his first thesis, Davidson notes an important feature, at least in his view, of the relation between reasons and actions. According to Davidson, a reason may explain an action when the action is described in one way, but not in another. The example he gives to make this point is of a case in which an agent flips a switch in order to turn on a light, which causes a light to go on, and a prowler to be alerted. Given his identity thesis concerning causally complex actions, Davidson assumes that there is one action here under a number of different descriptions. Thus, Davidson assumes that there is one action that was both a flipping of the switch and an alerting of the prowler. Now the reason why the agent flipped the switch was that he wanted to turn on the light and believed that by flipping
the switch he would turn on the light. This, however, is certainly not the reason why the agent alerted the prowler. Davidson therefore concludes that a reason explains an action only under a certain description.

We have already observed that Davidson’s identity thesis concerning causally complex actions is false, and so the example that he gives to establish that reasons explain actions only under descriptions is flawed. However, it seems easy enough to produce a better example on his behalf. Thus, suppose that the reason why Hamlet killed the man behind the arras is that he wanted to avenge his father’s death and believed that by killing the man behind the arras he would avenge his father’s death. Since the man behind the arras is identical with Polonius, we may say that there is one action here, which is both a killing of the man behind the arras and a killing of Polonius. But there is at least one sense in which the reason that explains why Hamlet killed the man behind the arras does not explain why Hamlet killed Polonius. Accordingly, one might say that the reason explains the action only under a description.

One might say this, but there are good reasons against doing so. In the first place, in using the notion of an action under a description to draw attention to the referential opacity that affects a sentence like "Hamlet killed the man behind the arras because he wanted to avenge his father death" one is naturally lead to the mistaken idea that the sentence contains a description that refers, or at least purports to refer, to a particular action. Indeed, Davidson makes precisely this mistake in "Actions, Reasons,
and Causes," where he remarks that reason-explanations are quasi-intensional. In a footnote attached to this remark, Davidson provides the following explanation:

[They are] "quasi-intensional" because, besides their intensional aspect, the description of the action must also refer in rationalizations; otherwise it could be true that an action was done for a certain reason and yet the action not have been performed. (Davidson 1963, p.5)

According to Davidson, the description of the action contained in a reason-explanation must refer to a particular action, and this makes reason-explanations only quasi-intensional. In contrast, one might say that a sentences like "Ralph believes that the man in the brown hat is a spy" is fully intensional because the description "the man in the brown hat" does not refer. This, I believe, is a mistake at least in the respect that it neglects the de re sense of the given ascription. However, regardless of whether or not the singular term refers on the de dicto reading of the ascription, Davidson is clearly mistaken in claiming that the description of the action in reason-explanations must refer; for reason-explanations do not typically contain any definite description or other singular term that even purports to refer to an action—neither on their de dicto nor their de re reading.

Thus, a sentence like "Hamlet killed the man behind the arras because he wanted to avenge his father’s death" contains no definite description of an action; the referential opacity that affects this sentence relates rather to the descriptions of the direct object of the action—the man behind the arras. We have already encountered this same point in connection with ascriptions.
of intentional action, which is the other main context in which Davidson uses the technical notion of an action under a description. Reason-explanations are intensional in precisely the same sense as ascriptions of intentional action. Both sorts of sentences have both an intensional, or \textit{de dicto}, and an extensional, or \textit{de re}, reading. However, neither sort of sentence typically contains a singular term that refers, or purports to refer, to an action, which is why it is misleading to use the notion of "an action under a description" to draw attention to the fact that these sentences have an intensional reading.

Part of the problem here is that when Davidson wrote "Actions, Reasons, and Causes," he had not yet dealt with the question of the logical form of action-sentences. This he freely admits in the introduction to \textit{Essays on Actions and Events}, where he notes that:

When I wrote Essay 1 ["Actions, Reasons, and Causes], it had not yet occurred to me that a sentence like "Eve ate the apple," should not be taken to contain a singular reference to an event; it is distinct in logical form from "Eve’s eating of the apple occurred," though the latter does imply the former. (Davidson 1980, p.xiv).

"Eve ate an apple" does not contain a singular term that purports to refer to an event, nor does "Hamlet killed the man behind the arras," and nor does "Hamlet killed the man behind the arras because he wanted to avenge his father’s death." And the fact that this last sentence entails the one that precedes it does not in anyway show that it is "quasi-intensional," or that it contains a description that refers to an action.

The second reason for resisting the idea that a reason explains an action only under a description is that it promotes
confusion concerning the sort of thing that a reason does explain. If Oedipus struck the old man at the crossroads because he wanted to clear a path back to Thebes, then what is explained by reference to Oedipus’s desire is not the striking of the old man by Oedipus—under this or some other description—but rather the fact that Oedipus struck the old man. What a reason explains is not an action, or an action under a description, but rather the fact that the agent in question performed a certain sort of action.

Of course, if Davidson’s paratactic analysis of intensional contexts is correct, then it may be possible to reinterpret facts and propositions in terms of utterances or inscriptions of sentences. This, however, is irrelevant to the present point. The important point here is that what is explained by a reason is propositional in character. It is the same sort of thing of which a predicate like "Jones perceives that x" or "Smith knows that y" is true of; whatever it is, it is not an action, event, or any other ordinary sort of object. For the sake of drawing attention to this point, without getting embroiled in semantic theory, we can harmlessly call the entities of which these predicates are true propositions, which is what Davidson sometimes does when he is not attempting to make detailed semantic points (Davidson 1989, p.14). And since the proposition that is explained in an explanation must be true, we can say that what a reason explains is a fact.

For both of the above two reasons, it is best to avoid altogether the idea that a reason explains an action under a description. And since Davidson employs this terminology in stating
his first thesis concerning primary reasons, certain terminological adjustments are required. Thus, Davidson states his first thesis as follows:

Cl: R is a primary reason why an agent performed the action A under description d only if R consists of a pro-attitude of the agent towards actions with a certain property and a belief of the agent that A, under description d, has that property. (Davidson 1963, p.5)

The fact that Davidson attempts to characterize primary reasons as reasons directed towards particular actions—as in "R is a primary reason why an agent performed the action A under description d"—stems from the idea, which we have just shown to be mistaken, that reason-explanations refer to particular actions. The sort of sentences to which the analysis of primary reasons is responsible are sentences like "Brutus stabbed Caesar because he wanted to end the tyranny," or "Achilles returned to the battle because he wanted to avenge the death of Patroclus." And since these sentences do not refer, or purport to refer, to particular actions, primary reasons should not be construed as reasons for particular actions. Thus, let us reword the antecedent in Cl to read as follows: "R is a primary reason why an agent performed an action of type A," or just "R is a primary reason why an agent A-ed."

Correspondingly, the characterization of the belief involved in an agent's reason for acting should be adjusted as well. Roughly speaking, what an agent believes when he performs an action of type A for a reason R is not that his action A is an action of the type he desires, but rather that by performing an action of type A he will perform an action of the type he desires. Consider, for
instance, a case in which I flip a switch because I want to turn on the light. How should we characterize my reason for acting in this sort of case? According to C1, it consists of a desire to turn on the light and the belief that my flipping of the switch was a turning on of the light. Now this belief-ascription makes sense only if the context picks out a single flipping of the switch by me, but I may have flipped the switch more than once. This is one problem with Davidson's characterization of the belief, but there is a further problem.

Let us suppose that I do flip the switch just once, and that I do turn on the light by flipping the switch. Since I turned on the light by flipping the switch, it follows that my flipping of the switch is not identical with my turning on of the light; rather, it is a causal component in my turning on of the light. As such, my flipping of the switch is not a turning on of the light. But according to C1, if I flipped the switch because I wanted to turn on the light, then I must have believed that my flipping of the switch was a turning on of the light. And so if C1 is correct, we can infer that I hold a false belief from the fact that I flipped the switch because I wanted to turn on the light, and turned on the light by flipping the switch. Clearly, this is an undesirable consequence.

Both of the problems just mentioned can be avoided by giving a slightly different characterization to the belief that an agent has when he acts for a reason. As a first approximation, to be qualified further below, let us restate C1 as follows:
R is a reason why an agent performed an action of type A only if R consists of a pro-attitude towards actions of type B, and a belief held by the agent that by performing an action of type A he would perform an action of type B.

For obvious reasons, we may call the sort of belief described here an *instrumental* belief. However, before we cite such a belief as an essential component in any reason for acting, there are a couple of qualifications to be observed.

In the first place, the characterization of the instrumental belief needs to be weakened to include cases in which the agent is less than certain about the consequences of his action. Suppose that I want to produce ten copies of a certain document, and that in response to this you tell me to transcribe a single copy using your special carbon-paper. I am incredulous towards the suggestion, but I give it a try anyway. To my surprise, the carbon-paper does the trick. In such a case, it could be true that I wrote on the carbon paper because I wanted to produce ten copies of the document, but false that I believed that by writing on the carbon-paper I would produce the ten copies. Nevertheless, while I need not have believed that by writing on the carbon-paper I *would* produce the desired result, it seems that I must have believed that I might. For if, on the contrary, I did not believe that I might produce the ten copies by writing on the carbon paper, then my desire to produce the ten copies would not explain why I wrote on the carbon paper; the question of why I did write on the carbon paper would remain.

In the second place, there is need for a qualification even once we weaken the characterization of the instrumental belief so
as to accommodate the foregoing sort of case. Thus, suppose that I draw the curtains because I want to wash the windows. Here, my desire to wash the windows is part of my reason for drawing the curtains, and yet it is not quite right to say that I believed that by drawing the curtains I would, or might, wash the windows. In this case, the relevant belief seems to be a belief that by drawing the curtains I would enable myself to wash the windows. However, instead of attempting to make further adjustments to the characterization of the instrumental belief involved in an agent's reason for acting, I think the appropriate response in this case is to accept the description of the instrumental belief just given, and to regard the corresponding reason-explanation as elliptical. For example, I think we should regard a sentence like "I drew the curtains because I wanted to wash the windows" as elliptical for something like "I drew the curtains because I wanted to enable myself to wash the windows." If this is right, then this last sort of case does not present any counterexample to the idea that when an agent acts for a reason, the agent believes that by performing an action of the type he performs, he will, or might, perform an action of the type he wants to perform.

There is, however, one clear counterexample to this suggestion; for there certainly is no instrumental belief involved in an agent's reason for acting when the agent acts for an intrinsic reason. We observed above that an intrinsic reason is the sort of reason an agent has when he performs an action of one sort simply for the sake of performing such an action; it is the sort of
reason given in an explanation like "He stepped on a snail because he wanted to step on a snail." Now one of the things that Cl fails to make explicit is the difference in content between an intrinsic and an extrinsic reason for acting. But there is indeed a difference, which ought to be made explicit; the difference is precisely that an extrinsic reason does, and an intrinsic reason does not, contain an instrumental belief.

Thus, pending a further qualification to be made below, let us give the following characterization to the two types of reason for which an agent can perform an action:

i) R is an intrinsic reason why an agent performed an action of type A only if R consists of a pro-attitude towards actions of type A.

ii) R is an extrinsic reason why an agent performed an action of type A only if R consists of a pro-attitude towards actions of type B, and a belief that by performing an action of type A he would, or might, perform an action of type B.

This, I believe, is an improvement upon the characterization of an agent's reason for acting that Davidson provides in Cl, but we can simplify, and improve, matters even further.

In particular, I think that we can do away with the technical notion of a pro-attitude in place of the more familiar notion of a want or desire. Davidson remarks in "Actions, Reasons, and Causes" that the thesis he defends in that paper is "ancient," which is an allusion to Aristotle, who expounded a theory of action that Davidson believes to be sufficiently like his own. But there is one respect in which Davidson admits they differ. Aristotle held that an agent's reason for performing an action always consists of a
desire. In coining the omnibus term "pro-attitude," Davidson explicitly rejected this aspect of Aristotle's theory. However, it seems to me that this is a mistake.

Davidson claims that a primary reason consists of a belief and a pro-attitude, but that it is generally otiose to mention both. He writes that "If you tell me that you are easing the jib because you think that will stop the main from backing, I don't need to be told that you want to stop the main from backing" (Davidson 1963, p.6). Evidently, the ascription of a desire in this case does not need to be made explicit because it is clearly implicit in the original statement. But how is it implicit if the agent's reason for acting could have consisted of a pro-attitude other than a desire?

Davidson claims that certain beliefs count as pro-attitudes, and, hence, that a reason for acting can consist of a pair of beliefs, without any desire. As an example of the sort of thing Davidson has in mind, consider the sentence "James went to church because he believed it was his duty to please his mother." Davidson seems to think that the reason, in this case, consists of James' belief that it is his duty to please his mother, and his belief that by going to church he would please his mother. But is this right? What if James' had no desire to do his duty, or did have a desire to shirk his duty? In the event, the two beliefs just cited would in no way constitute a reason for him to go to church, nor explain why he went if he did go.

What this shows, I think, is that there is an implicit ascription of desire in the above explanation just as much as there
is in Davidson’s example of an agent easing the jib because he thinks that will stop the main from backing. We assume, in other words, that James did want to do his duty, and this desire is an essential part of his reason for acting. The notion of a pro-attitude is that of an attitude concerning a type of action towards which the agent in question is positively disposed. But beliefs by themselves do not indicate any positive disposition at all, and provide reasons for acting only in the context of desires. As such, there is reason to believe that Davidson’s conception of a pro-attitude is too broad.

Let us sum up the results of the foregoing discussion by giving the following characterization of the sort of reason for which an agent can be said to perform an action.

C1* i) R is an intrinsic reason why an agent A-ed only if R consists of a desire to A.

ii) R is an extrinsic reason why an agent A-ed only if R consists of a desire to B, and a belief that by A-ing he would, or might, B.

We will see as we proceed that when an agent acts for a reason, the agent necessarily has further attitudes beyond those listed in C1*. However, while these further attitudes are essential to the understanding of what it means to act for a reason, they are not part of the content of the agent’s reason for acting. To act for a reason involves more than having a reason for acting, and C1* only serves to characterize the sort of reason an agent has when he acts for a reason. As it stands, I think that C1* is complete, but there is one final point to be noted in connection with it, before moving on to consider Davidson’s second thesis concerning primary reasons.
There is a certain temporal qualification to the attitudes that an agent has when he acts for a reason, and which is left implicit in both Cl as well as Cl*. Roughly speaking, the attitudes that an agent has when he acts for a reason are attitudes that the agent must have at the time of acting. For if I form the desire to stretch only after I raise my arm, then my desire to stretch is irrelevant to the explanation of why I raised my arm; in other words, this desire cannot be part of my reason for acting. Moreover, it seems that an agent must not only have the attitudes that constitute his reason for acting at the time of acting, but that he must also have them throughout the performance of his action. Thus, if I set out to build a squirrel house because I want to win the affection of a certain person, and decide half-way through the project that I no longer want that person’s affection, then my desire to win the affection of that person cannot be used to explain why I built the squirrel house. It may explain why I started the project, and got half-way through it, but the question of why I completed it will remain.

However, it is not quite right to say that an agent must in all cases have the attitudes that constitute his reason for acting throughout the performance of the action. This is made manifest in the case of causally complex actions where the intended event is brought about long after the action which caused it. Thus, suppose the queen kills the king because she wants to gain the throne, and that she kills him by putting poison in the grapefruit he is to have for breakfast in the morning. Now the queen might lose her
desire to gain the throne at any point after placing the poison in the grapefruit, but if the king eats it and dies, then it would be true to say that the queen killed the king because she wanted to gain the throne, regardless of whether she wants to gain the throne at the time the king dies.

Thus, in the case of causally complex and aggregate actions, the agent need not have the attitudes that constitute his reason for acting throughout the entire performance of his action; all that is required is that the agent have them throughout the performance of whatever causally simple actions are involved in the performance of his action. In most cases, however, such as flipping a switch, building a squirrel house, or climbing a mountain, the temporal difference between the completion of the component actions and completion of the action itself is insignificant. Thus, it will not be too imprecise to express the relevant point concisely by saying that when an agent A's for a reason R, the agent must have the attitudes that constitute his reason R while he A's.

III

At one point in "Actions, Reasons, and Causes" Davidson writes that "a person can have a reason for performing an action, and perform the action, and yet this not be the reason why he did it" (Davidson 1963, p.9). Accordingly, one might ask what else is involved in performing an action for a reason than performing the action and having a reason for performing it. It is in response to this sort
of question that Davidson advances his second thesis concerning primary reasons, which he states as follows:

C2 A primary reason for an action is its cause.

Now before we can evaluate C2 in any detail, there are a couple of respects in which it ought to be clarified.

In the first place, we need to address Davidson's use of the phrase "a primary reason for an action." In stating his first thesis concerning primary reasons, Davidson speaks of "a primary reason why an agent performed an action," which clearly indicates that what is being characterized is the sort of reason that not only justifies a certain sort of action, but which also explains why an agent performed such an action. Davidson's second thesis is certainly intended as a correlative thesis to the first and concerns these explanatory reasons as well, but the terminology he uses in stating the thesis fails to make this clear. This, however, is easy enough to adjust. As a first step towards clarification, let us restate C2 as follows: R is a primary reason why an agent performed the action A only if R caused A.

As a second step, we need to accommodate the changes that were required in the case of C1. We observed above that it is a mistake to attempt to characterize an agent's reason for acting as a reason for performing a particular action. For reason-explanations do not typically refer to any action, but rather assert the existence of an action of a certain sort. Thus, if C2 is a correlative thesis to C1, and if C1 ought to be reformulated as per C1*, then C2 ought to be adjusted as well. We can do this by restating C2 as follows:
C2* R is a primary reason why an agent A-ed only if R caused the agent to A.

In addition to its compatibility with C1*, this last formulation corresponds most closely to the examples Davidson gives of the causal statements that are allegedly true in any case in which an agent performs an action for a reason. For instance, Davidson writes that "not only did Oedipus want to kill his father, and actually kill him, but his desire caused him to kill his father" (Davidson 1973d, p.232).

Now C2 is a core claim of Davidson's causal theory of action, and in conjunction with C1, constitutes what we might call Davidson's causal analysis of acting for a reason. But for the reasons we have just considered, it will be best to identify the causal analysis rather with C1* and C2*. This, however, will not negatively affect Davidson's position in any way; for the reformulated theses are intended merely as a rational reconstruction of Davidson's causal analysis. And we will see as we proceed that whatever argument or defence Davidson provides for C2 can easily be formulated in terms of C2*.

The greater part of "Actions, Reasons, and Causes," is concerned with defending the causal analysis of acting for a reason against various possible objections. However, to defend a theory against objections is one thing, and to provide reasons for believing that the theory is true is yet another, and "Actions, Reasons, and Causes" contains very little in the form of positive justification for the causal analysis. What Davidson does provide is contained in the following, rather terse passage:
In order to turn the first "and" to "because" in "He exercised and wanted to reduce and thought that exercise would do it," we must, as the basic move, augment condition C1 with C2. (Davidson 1963, p.12)

The claim that Davidson makes in this passage relates to the difference in meaning between the following two sentences:

(1) He exercised and wanted to reduce and thought that exercise would do it;
(2) He exercised because he wanted to reduce and thought that exercise would do it.

We are to suppose here that there is a semantic distinction between these two sentences—that the latter entails the former, but not conversely. Assuming this is right, one might ask what must be added to (1) in order to give the meaning of (2)? And the answer that Davidson apparently wants to provide, or at least part of the answer, is (3).

(3) His desire to reduce and belief that by exercising he would reduce caused him to exercise.

To say that (3) must be added to (1) in order to give the meaning of (2) is to say that (3) is a necessary condition of (2). And the idea that (3) is a necessary condition of (2), is precisely what C2* entails in this case. So if (3) really is the correct response to the question raised above, then this suggests that C2* is true.

Now the semantic distinction between (1) and (2) is not entirely clear, and so one may also be unclear about the need to suppose that something like (3) is a necessary condition of (2). However, it is easy to produce a slightly different example in which the relevant distinction is more conspicuous. Thus, for the sake of understanding what exactly motivates the idea that C2* is true, let us consider another example.
Imagine that Jack and Jill are friends, and that Jill is on a hiring committee for a job that Jack has applied for. Now Jill wants to help Jack out and knows that by voting for Jack she would help him out. At the same time, Jill wants to vote for whoever she believes to be the best candidate. In this case, however, it turns out that Jack is the best candidate in Jill’s honest opinion. And so when it comes time to vote, Jill has two different reasons to vote for Jack. On the one hand, she wants to help Jack out and believes that by voting for Jack she will help him out; let us call this reason \( R_1 \). On the other hand, she wants to vote for the best candidate and believes that by voting for Jack she will be voting for the best candidate; let us call this reason \( R_2 \).

Now even though Jill has these two reasons to vote for Jack, it seems fairly clear that she could vote for Jack for only one of the two reasons, for both reasons, or indeed for neither reason. As a plausible example of the first possibility, we might suppose that Jill takes her job very seriously, and though she likes Jack and wants to help him out, she would never vote for anything other than purely professional reasons. Let us suppose, then, that even though Jill has both reasons to vote for Jack, she votes for Jack only for reason \( R_2 \), and not for reason \( R_1 \). Accordingly, it makes sense to ask the following question: in virtue of what is it the case that Jill voted for Jack for \( R_2 \), but not for \( R_1 \)?

This last question is essentially the same as the question raised above concerning the semantic distinction between (1) and (2), and Davidson would surely provide a similar sort of response.
In the case at hand, Davidson’s response would be something to the effect that if Jill voted for Jack for R₂, and not for R₁, then this is because R₂ did, and R₁ did not, cause Jill to vote for Jack. So assuming that there is nothing abnormal in this case, if the foregoing response is correct, then there is reason to believe that an agent A’s for reason R only if R causes the agent to A. In other words, insofar as the foregoing response is correct, there is reason to believe that C₂* is true.

The argument just provided is at least implicit in "Actions, Reasons, and Causes," and is more or less explicitly endorsed by a number of other philosophers who have attempted to reconstruct Davidson’s argument (Goldman 1970, p.77; Davis 1979, p.89; Antony 1989, p.159; Bishop 1989, p.101). There is, however, a problem with this sort of argument, which is by no means unknown, but which is generally neglected. Thus, the crucial question in the argument just given is the following: "In virtue of what is it the case that Jill voted for Jack for R₂, but not for R₁, even though she had both reasons to vote for Jack?" And the response that the causal theorist wants to provide—at least roughly speaking—is that it is in virtue of the fact that only R₂ caused Jill to vote for Jack.

Now the reason why it is necessary to add some such qualification to the causalist’s response is that without it the response is plainly false. For it does not follow from the fact that an agent has a reason R for A-ing, which causes him to A, that the agent A’s for reason R. We will see some examples of this momentarily, but the present point is to note how this fact demands
that an important caveat be added to the sort of argument we are considering. For if it does not follow from the fact that R₂ caused Jill to vote for Jack, that Jill voted for Jack for R₂, then one cannot claim that the reason why Jill voted for Jack for R₂, and not for R₁, is that only R₂ caused Jill to vote for Jack. The most that one can claim here is that the supposition that R₂ caused Jill to vote for Jack provides a partial response to the given question.

This, however, is a very important qualification, and one which tends to be insufficiently stressed. For insofar as we know that the supposition that R₂ caused Jill to vote for Jack provides at most only a partial response to the given question, then it is difficult to see what exactly is so appealing about that supposition. Indeed, it looks as if we are in need of a further story to justify the idea that the partial response provided by the causalist is indeed part of the correct response, rather than merely an incorrect response.

Judging by the remarks that are made by causal theorists, the retort that one is likely to get at this point is that if we do not suppose, in the case at hand, that R₂ caused Jill to vote for Jack, then how are we to account for the fact that R₂ explains why Jill voted for Jack? But this, of course, is a bad question insofar as it suggests that we do account for the fact that R₂ explains why Jill voted for Jack by supposing that R₂ caused Jill to vote for Jack. The better question is simply "How are we to account for the fact that R₂ explains why Jill voted for Jack?" Now the answer to this question—whatever it is—is not that R₂ caused Jill to vote for
Jack, and until we are relatively confident that we know what the correct answer is, it is difficult to see any justification for insisting that this must be a part of the correct response. At the very least, it cannot be said that the argument we have just considered provides such justification.

Before proceeding any further, let us confirm the important point mentioned above—that it does not follow from the fact that an agent has a reason R for A-ing, which causes him to A, that he A's for reason R. Davidson was well aware of this point, even at the time of writing "Actions, Reasons, and Causes," and it is for this reason that he explicitly cancels the suggestion that C1* and C2* constitute sufficient conditions of what it means to act for a reason (Davidson 1963, p.12). In "Freedom to Act," Davidson provides two examples to demonstrate the relevant point; the first of these, which he credits to Daniel Bennett, is as follows:

A man may try to kill someone by shooting at him. Suppose the killer misses his victim by a mile, but the shot stampedes a herd of wild pigs that trample the intended victim to death. Do we want to say that the man killed his victim intentionally? The point is that not just any causal connection between rationalizing attitudes and wanted effect suffices to guarantee that producing the wanted effect was intentional. The causal chain must follow the right sort of route. (Davidson 1973b, p.78)

There is no mention of what the rationalizing attitudes are supposed to be in this case, but for the sake of making a point we can provide some. Let us suppose that the rifleman was a hired assassin and that the reason why he was trying to kill the other man was that he wanted to earn the bounty that he was promised. Now Davidson's point is that even if we suppose that this reason caused
the assassin to fire his gun, which ultimately resulted in the
death of the other man, it would not be correct to say that the
assassin killed the other man for that reason. According to
Davidson, the assassin did not kill the other man for any reason at
all, since he killed him unintentionally.

However, I think that this is the wrong conclusion to draw,
and that Bennett’s example fails to establish the intended point.
In particular, it does not seem correct to say that the assassin in
this case killed his victim unintentionally. Nor does it seem
correct to say that he killed him intentionally. Rather, I think
the correct conclusion to draw in this case is that the assassin
did not kill the other man at all. We noted in chapter 2 that the
accordion effect, which transfers agency from an agent’s action to
the events that are caused by his action, only applies when the
given action is indeed the cause of the event in question; it does
not necessarily apply when the action is a mere causal factor in
the production of the relevant event. But the situation that
Bennett describes seems to be a case in which the action that the
assassin performs in firing the gun is at most a causal factor in,
rather than the cause of, the death of the other man. And in this
case it does seem to be incorrect to attribute the man’s death to
the assassin, even though the man might not have died had the
assassin not fired his gun.

As it stands, Bennett’s example does not show that C1* and C2*
do not provide sufficient conditions of what it means to act for a
reason, but it may be possible to modify the example so that it
does. However, for the sake of demonstrating the relevant point, we may turn to consider Davidson’s second example, since it avoids the defect in Bennett’s example. The second example Davidson gives is the famous rock-climber example, which he describes as follows:

A climber might want to rid himself of the weight and danger of holding another man on a rope, and he might know that by loosening his hold on the rope he could rid himself of the weight and danger. This belief and want might so unnerve him as to cause him to loosen his hold, and yet it might be the case that he never chose to loosen his hold, nor did he do it intentionally. (Davidson 1973b, p.79)

Here we have a case in which an agent has a reason for loosening his hold on the rope, which causes him to loosen his hold on the rope, and yet he does not loosen his hold on the rope for that reason. Since an agent acts intentionally whenever he acts for a reason, Davidson infers from the fact that the agent did not perform the action in question intentionally, that he did not perform it for the given reason, or for any reason at all. Accordingly, Davidson’s example presents us with a case in which an agent has a reason R for A-ing, which causes him to A, and yet it is not the case that he A’s for reason R.

Another example that philosophers often use to establish the same point, and which is worth mentioning, comes from Roderick Chisholm (Chisholm 1966, p.30). Suppose a man wants to inherit his uncle’s fortune, and believes that by killing his uncle he will inherit the fortune. This belief and desire agitate the nephew so severely that he drives excessively fast, which causes him to run over and kill a pedestrian who, unbeknownst to the nephew, is none other than the uncle. Here we can say that the nephew killed his
uncle and, arguably, that his reason for killing his uncle caused him to kill his uncle. However, we would not say that the nephew killed his uncle for that reason, or that he killed him intentionally; rather, it was something the agent did accidentally.

Both of the last two examples show that $C_1*$ and $C_2*$ do not constitute sufficient conditions of what it means to act for a reason, and the fact that they do not casts doubt upon the argument that is typically given in support of $C_2*$. For $C_2*$ is ultimately proposed in response to the question of what else is involved in acting for a reason beyond having a reason for performing a certain sort of action and performing such an action. And given that $C_1*$ and $C_2*$ do not constitute sufficient conditions of what it means to act for a reason, the most that can be said on behalf of $C_2*$ is that it provides a partial response to the question. But once again, until we are relatively confident that we know what the correct response is, it is difficult to see any justification for believing either that $C_2*$ does, or that it does not, provide part of the correct response. Certainly the argument that is given in support of $C_2*$ cannot be said to provide such justification.

What is needed, clearly, is some sort of convincing account of what it means to act for a reason. In the following section we will consider such an account, and at that point we will be able to sensibly discuss whether or not $C_2*$ is an essential part of the account. However, prior to this, it is necessary to dispense with a suggestion, which Davidson has advanced, that it not possible to provide a complete account of what it means to act for a reason.
The original publication of "Actions, Reasons, and Causes," contains a footnote in which Davidson expresses his belief that C2* could be strengthened so as to provide a complete account of what it means to act for a reason. However, this remark is deleted in the subsequent republication of the article. And in more recent articles, Davidson has simply denied that it is possible to complete the analysis of what it means to act for a reason. In "Psychology as Philosophy," he gives the following explanation:

What prevents us from giving necessary and sufficient conditions for acting on a reason also prevents us from giving serious laws connecting reasons and actions. To see this, suppose we had sufficient conditions. Then we could say: whenever a man has such-and-such beliefs and desires, and such-and-such further conditions are satisfied, he will act in such-and-such a way. There are no serious laws of this kind. (Davidson 1973d, p.233)

Davidson believes that there are no serious laws relating an agent's reasons and his actions not just as a matter of fact, but as a matter of principle. His reasons for believing this are complicated to say the least, but we need not consider them here, for Davidson is clearly mistaken in thinking that the fact that there are no such laws entails that it is not possible to complete the account of what it means to act for a reason.

In particular, it is a mistake to suppose that what we are seeking, in asking for an account of what it means to act for a reason, are conditions that would enable us to predict whether or not an agent will perform an action of type A, given that he has a certain reason R for A-ing. What we are seeking are rather conditions, including the fact that the agent A-ed and had a certain reason R for A-ing, which collectively entail that the
agent A-ed for reason R. This is an entirely different sort of investigation from the one in which Davidson suggests we are engaged; the investigation in which we are engaged is not about predicting behaviour, but rather analyzing a certain concept. Thus, even if Davidson were right in thinking that there are no laws relating reasons and actions, this would in no way suggest that it is not possible to complete the account of what it means to act for a reason. Let us now go on to complete the account.

IV

Whenever an agent acts for reason, the agent has a reason for performing a certain sort of action, and performs such an action; if an agent A's for reason R, then the agent has the reason R for A-ing and he A's. This much is uncontroversial. In addition to this, causal theorists maintain that an agent A's for reason R only if R causes the agent to A. However, it does not follow from the fact that an agent has reason R for A-ing and R causes the agent to A, that the agent A's for reason R. This too is uncontroversial, and there are numerous examples in the literature to demonstrate this fact. In the last section we considered two of these.

Now among the various things that one might learn from these examples, there is one very obvious point that we have yet to address. Both of the examples considered above demonstrate that even if we suppose that the agent's reason for A-ing caused him to A, we would still not say he A-ed for that reason, since it is
clear in these cases that the agent did not A intentionally. But why, then, do we not just take these examples as proof that acting intentionally is a necessary condition of acting for a reason? In other words, if what we are seeking is an analysis of what it means to say that an agent A-ed for reason R, then why not just include a clause that states that the agent A-ed intentionally?

The reason, of course, at least in the case of Davidson, concerns his definition of intentional action. Since Davidson defines an intentional action as one done for a reason, it is clear that he will not include the concept of intentional action in the analysis of what it means to act for a reason. However, we have already seen that Davidson’s definition of intentional action is flawed; in particular, we found that it does not follow from the fact that an agent A-ed intentionally that he A-ed for a reason. But if this is right, and if A-ing intentionally is a necessary condition of A-ing for a reason, then there is no good reason not to include the concept of intentional action in the analysis of acting for a reason. Of course, one may still want an analysis of what it means to act intentionally, but that is another matter, which ought to be kept apart from the analysis of acting for a reason. Similarly, one may want an analysis of what it means to believe something, but that should not prevent one from using the concept of belief in analysis of what it means to know something.

We have already observed that in order to account for what it means to act for a reason, it is necessary to distinguish between the two types of reasons for which an agent can act. Given this, we
may express the foregoing point by means of means of the following, partial analysis of what it means to act for a reason:

Case 1: Acting for an intrinsic reason

X A’s because he wants to A only if:
i) X A’s intentionally,
ii) X wants (while A-ing) to A.

Case 2: Acting for an extrinsic reason

X A’s because he wants to B only if:
i) X A’s intentionally,
ii) X wants (while A-ing) to B,
iii) X believes (while A-ing) that by A-ing he would, or might, B.

It is important to note that neither of these analyses is complete; the conditions stated in both cases provide necessary but not sufficient conditions of the sentence they analyze. Since the two different cases are perfectly parallel, it will be sufficient to consider only one of them. To see how the analysis in Case 2 is insufficient, consider again the example of Jack and Jill.

In this example, Jill has two different reasons to vote for Jack. On the one hand, she wants to help Jack out and believes that by voting for Jack she will help him out; this is reason $R_1$. On the other hand, she wants to vote for the best candidate and believes that by voting for Jack she will be voting for the best candidate; this is reason $R_2$. Now even though Jill has these two different reasons to vote for Jack, it is at least possible that she could vote for Jack for only one of the two reasons, for both reasons, or indeed for neither reason. But notice that none of this changes if we add the fact that Jill intentionally votes for Jack. For it does not follow from the fact that Jill intentionally votes for Jack and has reason $R_1$ for voting for Jack, that she votes for Jack for $R_1$. 
Nor does it follow that she votes for Jack for $R_3$, or for the combination of both reasons, or indeed for any reason.

In accordance with the argument given in the last section, let assume once again that Jill voted for Jack only for $R_3$, and not for $R_1$. And let us add the fact that Jill intentionally voted for Jack. As such, we are faced with a similar question to the one considered above; namely, "In virtue of what is it the case that Jill intentionally voted for Jack for $R_3$, but not for $R_1$, even though she had both reasons to vote for Jack?" And here, the causal theorist might supply a similar response—that if Jill intentionally voted for Jack for $R_3$, and not for $R_1$, then this is because $R_3$ did, and $R_1$ did not, cause Jill to vote for Jack. So if the argument given in the previous section supports the idea that $C_2^*$ is true, then it appears that this last argument provides just as much support.

However, it is important to realize that this last argument, like the previous one, provides very little support indeed; for once again, the causalist’s response is at most only a partial answer to the given question. To see this, imagine that although Jill intentionally voted for Jack and that $R_3$ caused Jill to vote for Jack, Jill did not intend to vote for the best candidate. Thus, imagine that although Jill wanted to vote for the best candidate and believed that by voting for Jack she would be voting for the best candidate, this belief and desire caused Jill to vote for Jack without her yet having formed or acquired the intention to vote for the best candidate. Now if it were true that Jill did not intend, while voting for Jack, to vote for the best candidate, then we
would surely deny that she voted for Jack because she wanted to vote for the because candidate. In other words, if Jill lacked this intention while voting for Jack, then we would deny that she voted for Jack for \( R_2 \), regardless of whether \( R_2 \) caused Jill to vote for Jack, and regardless of whether Jill intentionally voted for Jack.

The foregoing considerations reveal an obvious lacuna in the analysis of acting for a reason outlined above; for that analysis makes no mention of intentions. But when an agent acts for a reason, the agent necessarily has a corresponding intention—an intention whose content matches the content of the constitutive desire in the agent's reason. To be more precise, there is a distinction to be noted between the content of an agent's intention when he acts for an intrinsic reason and the content of an agent's intention when he acts for an extrinsic reason.

If an agent A's because he wants to A, then the agent must intend to A. But if an agent A's because he wants to B, then the content of the agent's intention will be slightly more complex—the agent must intend to B by A-ing. Thus, if I flip the switch because I want to turn on a light, then it must be the case that I intend not only to turn on the light, but to turn on the light by flipping the switch. Corresponding to the sort of belief involved in an extrinsic reason, we may call the intention that an agent has when he acts for an extrinsic reason an *instrumental* intention.

Given the foregoing, let us fill the lacuna in the analysis outlined above, by providing the following, more complete account of what it means to act for a reason:
Case 1: Acting for an intrinsic reason

X A's because he wants to A only if:
   i) X A's intentionally,
   ii) X wants (while A-ing) to A.
   iii) X intends (while A-ing) to A.

Case 2: Acting for an extrinsic reason

X A's because he wants to B only if:
   i) X A's intentionally,
   ii) X wants (while A-ing) to B,
   iii) X intends (while A-ing) to B by A-ing,
   iv) X believes (while A-ing) that by A-ing he would, or might, B.

Now each of the conditions listed in these two analyses is certainly necessary, but what is interesting is that they appear to be sufficient as well. In other words, the reason-explanation stated in each case, appears not only to entail, but also to be entailed by, the conjunction of the sentences below it. And what is interesting about this is that if it is right, then there is no evident need for a causal statement relating reason and action in the analysis of acting for a reason.

As we have seen, the reason for thinking that a causal statement of this sort is needed in the analysis of acting for a reason is that we are in need of some account of the criterion that distinguishes the reason that does explain why an agent did what he did from the other reasons he may have had for doing it. However, it appears that what distinguishes the explanatory reason from the rest is not what caused the agent to do what he did, but rather what the agent intended to do by doing what he did.

Thus, in the case of Jack and Jill, if Jill intentionally voted for Jack for R₂, but not R₁, even though she had both reasons to vote for Jack, then this is because she intended to vote for the
best candidate by voting for Jack, but did not intend to help Jack out by voting for Jack. On the other hand, if she intentionally voted for Jack for both \( R_1 \) and \( R_2 \), then this is because she intended to accomplish both things by voting for Jack. And finally, if she intentionally voted for Jack for neither reason, than this is because she did not intend to accomplish either of these things by voting for Jack.

Assuming the foregoing is right, the demand for an account of what remains in the analysis of acting for a reason beyond having a reason for performing a certain sort of action and performing such an action provides no justification at all for the idea that \( C_2^* \) is true. Thus, one cannot claim, as Davidson does, that in order to turn the first "and" to "because" in "He exercised and wanted to reduce and thought that exercise would do it" we must as the basic move augment \( C_1^* \) with \( C_2^* \). Rather, what we must do is acknowledge two simple facts: i) that if he exercised because he wanted to reduce and thought that exercise would do it, then he must have intended to reduce by exercising; ii) that if he exercised because he wanted to reduce and thought that exercise would do it, then he must have exercised intentionally.

That Davidson has failed to fully appreciate these facts is readily understandable, or at least predictable, given other aspects of his philosophy of action. Indeed, neglect for the first fact can be traced back to an ontological prejudice against intentions that Davidson held at least at the time of writing "Actions, Reasons and Causes." And the neglect for the second fact,
on the other hand, is an obvious consequence of Davidson’s definition of intentional action. I will comment on each of these points in turn.

In "Actions, Reasons, and Causes" Davidson explicitly sought to provide an account of what it means to act for a reason that made no use of intentions. That this was his intention is a rather remarkable fact; for Davidson was well aware that whenever an agent acts for a reason, he has a corresponding intention. Thus, Davidson writes that:

To know the primary reason why someone acted as he did is to know an intention with which the action was done. If I turn left at the fork because I want to get to Katmandu, my intention in turning left is to get to Katmandu. (Davidson 1963, p.7)

But if it follows logically from the fact that an agent acted for a certain reason that he had a corresponding intention, it would be astounding if this fact were not relevant to the question of what it means to act for a reason. Why, then, did Davidson fail to acknowledge the role that intentions play in acting for a reason? Part of the answer to this is contained in the following passage, which immediately succeeds the one just given:

The expression "the intention with which James went to church" has the outward form of a description, but in fact is syncategorematic and cannot be taken to refer to an entity, state, disposition, or event. Its function in context is to generate new descriptions of actions in terms of their reasons; thus "James went to church with the intention of pleasing his mother" yields a new, and fuller, description of the action described in "James went to church." (Davidson 1963, p.8)

This last passage has puzzled many philosophers, who have rightly wondered why James’ intention to please his mother should
correspond to a state of mind any less than James’ desire to please his mother? In fact, there is no good reason why Davidson took this position in "Actions, Reasons, and Causes"; he simply believed at the time that there were no independent mental states of intention. Years later, in an article entitled "Intending," Davidson sought to give an account of what it means to intend to do something, and it was at this point that he recognized his mistake in refusing to acknowledge the existence of intentions. In a more recent article, entitled "Problems in the Explanation of Action," Davidson explicitly acknowledges this error.

At one time (about 25 years ago when I wrote "Actions, Reasons, and Causes) I thought there were no such states as intending; there were just intentional actions. This was, I now believe, an error. This is clear in the case where an intention is formed long before the intended action is performed, and even clearer in the case where the intended action is never performed. Intentions are also required to explain how complex actions are monitored and controlled. (Davidson, 1987 p.39)

However, having renounced his earlier prejudice against intentions, Davidson has by no means given up the causal analysis of acting for a reason; rather, he has simply produced a more complicated causal analysis.

Davidson now believes that whenever an agent acts for a reason, the agent has a corresponding intention, and that sometimes the agent forms this intention prior to acting, and sometimes the intention is formed only when the agent begins to act. In the former case, Davidson maintains that the agent’s reason causes him to form the intention, which in turn causes him to act; in the latter case, Davidson believes that the formation of the intention
and the performance of the action are both directly caused by the agent’s reason (Davidson 1985a, p.221). In the one case, the agent’s reason is the direct cause of his action; in the other case, it is the indirect cause. But since the reason is the cause of the action in both cases, Davidson maintains his commitment to the core claim of the causal analysis—that when an agent acts for a reason, the agent’s reason is the cause of his action.

Davidson himself remarks that this account of the relations between actions, reasons, and intentions is "rather absurdly mechanical," but one ought for that very reason to reconsider it, and to reflect upon what sustains it. What this mechanical model of the mind reveals, I believe, is a failure to appreciate the role that intentions play in acting for a reason. Having renounced his earlier ontological prejudice against intentions, Davidson has nevertheless failed to recognize that intentions play the very role that he originally set aside for causal relations between reasons and actions. So instead of giving up the idea that an agent’s reason must be the cause of his action when he acts for a reason, Davidson has simply produced a more complicated causal story.

Now if there is any explanation for why, having acknowledged the existence of intentions, Davidson has failed to appreciate the role they play in acting for a reason I believe that it relates to his definition of intentional action. For given Davidson’s definition, an analysis of what it means to act for a reason is at the same time supposed to yield an analysis of what it means to act intentionally. And this is something that we surely do not achieve
simply by acknowledging the fact that an agent has a corresponding intention when he acts for a reason. Consequently, it is easy to miss the role that intentions do play in acting for a reason. But the mistake here, of course, is the idea that an analysis of acting for a reason should yield an analysis of what it means to act intentionally. And this mistake is the obvious consequence of Davidson’s definition of intentional action.

I argued in the first section of this chapter that A-ing intentionally is a necessary but not sufficient condition of A-ing for a reason, and that Davidson’s definition of intentional action is therefore false. Given this result, there appears to be no good reason not to use the concept of intentional action in the analysis of what it means to act for a reason. And now the apparent success of the non-causal analysis of acting for a reason outlined above serves to reinforce this approach as well as the original criticism of Davidson’s definition of intentional action. Nevertheless, one may still seek an analysis of what it means to act intentionally, and if it were possible to provide one, then we could further reduce the non-causal analysis outlined above. However, I do not believe that there is any interesting analysis of intentional action forthcoming, and so for the sake of defending the non-causal analysis as it stands, let us pursue this point in a little more detail. To begin with, let us briefly consider the attempts that other causal theorists have made at analyzing the concept of intentional action.
According to Davidson, whether or not an agent acts intentionally depends not only on whether the action is caused by a reason the agent has for performing it, but also on the way in which the reason causes the action. Davidson tends to put this point by saying something to the effect that an agent performs an intentional action only if the agent’s reason causes his action in the right way. This qualification is not intended as a means of completing the analysis, but only as a means of indicating the respect in which the analysis is incomplete.

Now the problem of specifying the way in which an agent’s reason must cause his action in order for it to be the case that the agent acts intentionally has come to be known as the problem of causal deviance, and while Davidson is convinced that the problem is insurmountable, a number of philosophers have spent a great deal of effort at trying to solve the problem. Following a suggestion of Davidson’s, most causal theorists regard the problem of causal deviance as dividing into two separate problems. Thus, in "Freedom to Act," Davidson suggests that we ought to distinguish between deviant internal causal chains and deviant external causal chains (Davidson 1973b, p.79). And this same distinction is drawn by other philosophers in slightly different terms; Myles Brand distinguishes the problem of antecedential waywardness from the problem of consequential waywardness, and John Bishop distinguishes basic from non-basic deviance (Brand 1984, p.18; Bishop 1989, p.132).

Employing Davidson’s terminology, the problem of external causal deviance is a problem that only affects causally complex
actions, and it concerns the causal relations between the causally simple action that the agent performs and the subsequent events that are caused by this action. The problem of internal deviance, on the other hand, is a problem concerning the causal relations that are alleged to exist between the agent’s reason and his causally simple actions. Davidson offers the rock-climber example as an example of internal deviance, and Bennett’s example of the lucky marksman as an example of external deviance. We have already observed that Bennett’s example is flawed, but for the sake of understanding the distinction that is being drawn here, Chisholm’s example of the murderous nephew will suffice as an illustration of the problem of external deviance.

Now if we are right in rejecting the idea that every intentional action is done for a reason, then the problem of internal deviance is surely an illusion; for if an agent need not have any reason for performing his action when he acts intentionally, then the question of whether or not an agent acts intentionally clearly has nothing to do with facts about the causal chain between the agent’s reason for acting and his action. On the other hand, since the problem of external deviance concerns the legitimate causal chains that run between an agent’s causally simple actions and the events that are caused by those actions, the problem of external deviance is not necessarily misconceived. And it is perhaps tempting to suppose that the reason why the nephew, in Chisholm’s example, does not intentionally kill his uncle has to do with the fact that while the nephew did do something that caused
his uncle's death—he drove over him—the uncle’s death was not caused in the right way. But what, then, is meant by the qualification "the right way?"

Various causal theorists have attempted to answer this question, and while there are differences in the details of the various responses, almost all attempt to answer the question by recourse to the notion of an action-plan (Goldman 1970, p. 59; Armstrong 1981, p. 79; Brand 1984, p. 25; Bishop 1989, p. 132). Essentially, the idea is that in the case of a causally complex action, an agent performs his action intentionally only if there is a match or correspondence between the agent's action-plan for bringing about the event in question and the means by which the given event is actually brought about. For instance, to take Chisholm's example, the suggestion is that the nephew did not kill his uncle intentionally because the way in which the uncle's death was brought about does not correspond to the nephew's plan of action for bringing about his uncle's death.

Now it is probably stretching the notion of a plan to suggest that an agent has a "plan of action" whenever he does something intentionally. However, let us ignore this point and focus on the cases where the agent clearly does have a plan of action. In these cases, it is certainly reasonable to assume that there must be some sort of correspondence between the agent's action-plan for bringing about the relevant event and the means by which it is in fact brought about. However, it is clear that there need not be a perfect correspondence. Thus, if Oswald had planned on shooting
Kennedy by putting a bullet in his heart, and had instead got him in the head, there would be absolutely no inclination to deny that Oswald intentionally shot Kennedy. So the most that can be said is that there must be some rough correspondence between the agent’s action and his action-plan.

This is one problem with the proposal to analyze the concept of intentional action in terms of a match between the agent’s action and his action-plan, but there is yet another. For in the case of causally complex actions of any significant duration, or aggregate actions, the agent’s action-plan may change and evolve throughout the performance of his action. Thus, if a man sets out to sail around the world, his plan for achieving that effect may undergo any number of modifications along the way. As it does, the agent will acquire new beliefs, and dispense with old ones, and the means by which he actually sails around the world may in no way resemble the plan with which he set out. So in addition to having to tolerate the idea of a rough correspondence between the agent’s action and his action-plan, one also has to tolerate the idea of an evolving and variable action-plan. Clearly, this is not the sort of thing that provides for any interesting analysis of sentences of the form "X A-ed intentionally."

Instead of speaking of action-plans, we might instead speak of expectations, and say that an agent A’s intentionally if and only if he A’s roughly in the way he expected to A. This definition is not as uninformative as it may at first appear. It tells us that an agent expects to A whenever he A’s intentionally, but it avoids the
idea, which we have already observed to be false, that an agent intends to A whenever he A’s intentionally. Secondly, it tells us that whenever an agent A’s intentionally, the agent has certain beliefs about the manner or the means by which he will A, which is also right. Thirdly, it captures the sort of idea that we have just been considering—that when an agent A’s intentionally, the agent’s beliefs about the manner or the means by which he will A must be mostly true. However, the problem with the foregoing definition is that there is no hope of identifying exactly which beliefs, or how many of these beliefs, must be true. For this reason, there is no hope of any precise analysis of what it means to act intentionally. Accordingly, I submit that the non-causal analysis of acting for a reason outlined above cannot be reduced in any interesting way.

V

If the argument of the preceding section is sound, then we need not suppose that C2* is true in order to account for what else is involved in acting for a reason beyond having a reason for performing a certain sort of action and performing such an action. This not only undermines the justification that is typically given in support of C2*, but also strongly suggests that C2* is false. Certainly it shifts the burden of proof onto those who insist that it is true. Having dispensed with the argument that is typically given in support of C2*, let us now see if any support for the thesis can be found in the concept of causality.
To begin with, it is perhaps worth mentioning that there is obviously some interpretation on which C2* is true; for the term "cause" is sometimes treated as being synonymous with "reason." Thus, in the Physics, Aristotle distinguished amongst the various sorts of questions that one might ask, and the responses that they demand. And the result of this famous investigation is what we now routinely call Aristotle's doctrine of the four causes. Most would agree, I think, that it would be better to call it the doctrine of the four reasons or types of explanation, but the fact remains that the term "cause" is sometimes used in this loose sense. Accordingly, one might suggest that C2* must be true because to say that a reason R caused an agent to A is just to say that R is the reason why the agent A-ed. However, it is perfectly clear that C2* is not to be interpreted in this trivial sense; for it is agreed by all that C2* states at most necessary but not sufficient condition of what it means to act for a reason. In other words, it is agreed by all that it does not follow from the fact that a certain reason R caused an agent to A that R is the reason why the agent A-ed.

In what sense of the term, then, has it been thought that C2* is true? Let us approach this question by considering what Davidson has said about the concept of causality. Both in "Actions, Reasons, and Causes" and in "Causal Relations," Davidson seems to endorse a roughly Humean conception of causality. In both articles, he explicitly quotes a definition of cause that Hume gives in the Inquiry, which is as follows:
We may define a cause to be an object followed by another, and where all objects, similar to the first, are followed by objects similar to the second. Or, in other words, where, if the first object had not been, the second never had existed. (Hume 1955, p. 87)

Though Hume, evidently, regarded these two definitions as being equivalent, it is unlikely that they are, and in both cases in which Davidson quotes Hume, he quotes only the first of the two definitions Hume gives. Nevertheless, it worth noting that Hume’s second definition hardly supports the idea that C2* is true. Thus, consider the following sentence:

(4) Brutus stabbed Caesar because he wanted to end the tyranny.

According to C2*, (4) entails (5).

(5) Brutus’s desire to end the tyranny and his belief that by stabbing Caesar he would end the tyranny caused him to stab Caesar.

On a natural interpretation of Hume’s second definition, (5) states that had Brutus not wanted to end the tyranny and believed that by stabbing Caesar he would do so, then he would not have stabbed Caesar. But this is certainly not entailed by (4), which is consistent with the possibility that Brutus might have stabbed Caesar for some other reason, or no reason at all. Thus, if there is any truth to Hume’s second definition, I think it suggests that C2* is false. However, it may be for other reasons that Davidson ignores this definition, so let us now consider the first.

Hume speaks of objects as causes and effects, but since they are objects that follow one another, it is natural to interpret him as speaking of events. Thus, one might restate Hume’s first definition as follows: \( e_1 \) caused \( e_2 \) if and only if all events
similar to \(e_1\) are followed by events similar to \(e_2\). An important ambiguity in Hume's first definition concerns the relation of similarity; for similarity is an inherently imprecise notion. One event is similar to another only in certain respects, and it is easy to find some respects in which any two events are similar.

Consider, for instance, a case in which a rock strikes a window and causes the window to break. Hume's definition tells us that the striking of the window by the rock caused the window to break if and only if the rock struck the window, the window then broke, and every event similar to this striking of the window by the rock is followed by an event similar to this breaking of the window. But similar in what respects? Certainly not in being "a striking of the window by a rock" and "a breaking of a window"; for not every window that is struck by a rock breaks.

Now this substantial ambiguity in Hume's definition would seem to be a problem, but in "Actions, Reasons, and Causes," Davidson apparently embraces precisely this aspect of Hume's definition. Thus, he writes that "\(e_1\) caused \(e_2\)" is true if and only if there are descriptions of \(e_1\) and \(e_2\) such that the sentence obtained by putting these descriptions for "\(e_1\)" and "\(e_2\)" in "\(e_1\) caused \(e_2\)" follows from a true causal law (Davidson 1963, p. 16). Unlike Hume, who defines causal relations in terms of generalizations, Davidson speaks of causal laws, which he says are distinguished from generalizations in that they are confirmed by their instances and they support counterfactuals. However, in one important respect the two definitions are alike; Hume's definition fails to specify the
respects in which the alleged cause and effect must be similar to the events described in the generalization, and Davidson’s definition imposes no restrictions on the permissible descriptions of the alleged cause and effect which instantiate the causal law.

The absence of any restrictions on permissible descriptions makes it doubtful that Davidson’s definition states a sufficient condition of causation. For if there are no restrictions placed on the permissible descriptions of the alleged cause and effect, it seems possible that any two events can be described in such a way that they instantiate a causal law. In "Causal Relations," Davidson himself seems to agree with this point, and consequently denies that he has given a definition of causation (Davidson 1967b, p.160). What he does provide is at most a necessary condition, which he expresses elsewhere as follows: e₁ causes e₂ only if there are descriptions of e₁ and e₂ which show that they fall under a causal law (Davidson 1976, p.262).

Davidson does not provide a full definition of cause in terms of which one might evaluate C₂*, and the necessary condition that he does provide is so weak that it cannot be said to constitute a meaningful criterion either. Indeed, consider any false causal statement. How can one possibly show that there are no descriptions of the alleged cause and effect under which they instantiate a causal law? Thus, the best that we can do here is to ask if there is any good reason to believe that the necessary condition that Davidson does cite is indeed a necessary condition of the truth of any reason-explanation? By a good reason I mean a reason other than
one related to the theoretical possibility that any two events have descriptions under which they instantiate a causal law.

Consider (4) again. Is there any good reason to believe that there exists a causal law governing the alleged cause and effect in this case? Notice that the alleged cause in the case of a reason-explanation is not an event, but rather a complex mental state. Davidson considers this point in "Actions, Reasons, and Causes" and responds by noting that states are often named as the causes of events, as in "The bridge collapsed because of a structural defect" (Davidson 1963, p.12). In the case of (4), then, the alleged cause is Brutus’s reason for stabbing Caesar, and the alleged effect is an event that is a stabbing of Caesar by Brutus. The question is whether there is a causal law governing this reason and action.

Now it is clear that there is no law to the effect that whenever Brutus wants to end the tyranny and believes that by stabbing Caesar he would do so, he stabs Caesar. Brutus may have wanted to stab Caesar for years, and only brought himself to do it on this one occasion. But this, of course, does not show that there is no law governing the reason and action; for the relevant law may be expressed in terms of very different concepts. Indeed, according to Davidson, they must be. Thus, he writes that:

The laws whose existence is required if reasons are causes of actions do not, we may be sure, deal in the concepts in which rationalizations must deal. If the causes of a class of events (actions) fall in a certain class (reasons) and there is a law to back each singular causal statement, it does not follow that there is any law connecting events classified as reasons with events classified as actions—the classifications may even be neurological, chemical, or physical. (Davidson 1963, p.17)
By suggesting that the law, whose existence is entailed by a reason-explanation, must be expressed in terms of concepts that are radically different from those employed in the reason-explanation itself, Davidson sketches a way of defending the supposition that reason-explanations entail the existence of causal laws. But this defence, of course, does not provide any grounds for believing that reason-explanations do entail the existence of causal laws.

Davidson suggests that rough generalizations are usually what counts as evidence for the existence of a causal law. Thus, he says that a generalization like "Windows are fragile, and fragile things tend to break when struck hard enough, other conditions being right," provides evidence for the existence of a causal law governing the events at hand in a case in which a rock strikes a window and causes it to break. (Davidson 1963, p.16). However, I think that this seriously underestimates the evidence that is relevant in this sort of case. In addition to whatever rough generalization we might possess concerning rocks and broken windows, we know what the salient variables are, and have at least some idea of the relationships amongst them. Thus, we know that by increasing the momentum of the rock, or by decreasing the thickness of the window, we increase the chances that the window will break. This is clearly what we would call causal knowledge, and it is this sort of knowledge, if anything, that serves as evidence for the existence of a causal law governing the events at hand.

However, there is nothing analogous to this in the case of reason-explanations. Thus, it is perfectly consistent with the
truth of (4) that we possess no generalization—however much we relax the standards—concerning what Brutus does when he wants to end the tyranny and believes that by stabbing Caesar he will end the tyranny. Nor will it do to provide a generalization as vacuous as "People tend to do that which they have a reason for doing." For as Davidson himself admits, if we were to guess at the frequency with which people perform actions for which they have reasons, it would be vanishingly small (Davidson 1976, p.264).

What is even more important in the case of reason-explanations is that we often have no knowledge at all of the salient variables. If Brutus wants to end the tyranny and believes that by stabbing Caesar he would do so, then he might stab Caesar, but maybe not if he really likes Caesar, or has a long-standing principle to harm no one, or if he lacks the courage to do it, or if he simply loses interest. Clearly, there is no limit to the kind of provisos one would be forced to acknowledge in any attempt to make a prediction on whether Brutus would stab Caesar given the fact that he has reason for stabbing Caesar. In short, what we lack here is what we can call causal knowledge. More precisely, it is consistent with the truth of (4), and our being justified in believing that (4) is true, that we lack such knowledge. Our evidence for the truth of (4) may simply be that this is the explanation that Brutus himself gave for why he stabbed Caesar. Typically, this is just the sort of evidence we use in explaining an agent’s behaviour, and this sort of evidence certainly does not provide evidence for the existence of any causal law.
It is clear that Davidson's belief that reason-explanations entail the existence of casual laws has little to do with the kind of evidence we consider in providing these explanations. It is rather by a theoretical route that Davidson arrives at this belief. Thus, in order to answer the question of what is involved in acting for a reason beyond having a reason for performing a certain sort of action and performing such an action, Davidson believes that we must suppose that C2* is true. And given what he cites as a necessary condition of causation, Davidson is thereby committed to the idea that reason-explanations entail the existence of causal laws governing the agent's reason and action. But of course the flaw in this pattern of reasoning is the idea that we must suppose that C2* is true in order to answer the question of what remains in the analysis of acting for a reason.

There is, I submit, no justification at all for the idea that reason-explanations entail the existence of a causal law governing the agent's reason and action. And if the existence of a causal law is a necessary condition of the truth of any causal statement, then the idea that reason-explanations entail a causal statement relating the agent's reason and action is also unjustified. Accordingly, the concept of cause with which Davidson has sought to defend C2* fails to provide any reason at all for believing that C2* is true. It remains to be shown, then, that there is any interesting sense in which C2* is true, and until this can be shown, the presumption clearly lies in favour of the idea that it is not. For C2* purports to state a necessary condition of what it
means to act for a reason, but we have already seen that we can account for what it means to act for a reason without supposing that C2* is true.

Most of the interest in the debate concerning C2* relates to the question of whether or not reason-explanations are a species of causal explanation, and I think that the foregoing results provide us with a clear answer to this question. Reason-explanations differ from causal explanations both in terms of what they explain, and what does the explaining. What is explained in the case of a causal explanation is the fact that an event, or an event of a certain sort, occurred. But what is explained by a reason-explanation is not merely the fact that an event occurred, or that an agent performed a certain sort of action, but rather the fact that an agent intentionally performed a certain sort of action.

Secondly, a causal explanation explains the fact that a certain event occurred by pointing to the fact that some other event occurred. Thus, we might explain the fact that a window broke by reference to the fact that it was struck by a rock. Here, one can ask "In virtue of what does this fact, rather than some other, explain the fact in question?" And the answer, evidently, is that the explanatory fact correctly identifies the cause of the relevant event. Thus, the fact that the window was struck by a rock explains the fact that the window broke only because the striking of the window by the rock caused the window to break. On the other hand, a reason-explanation explains the fact that an agent intentionally performed a certain sort of action by reference to the fact that he
had a certain reason for performing such an action. Thus, we might explain why Brutus intentionally stabbed Caesar by reference to the fact that he wanted to end the tyranny. Once again, one can ask "In virtue of what does this fact, rather than some other, explain the fact in question?" But in this case, there is an entirely different answer available, which is that the explanatory fact is the one that correctly identifies what the agent intended to do by doing what he did. Thus, the fact that Brutus wanted to end the tyranny explains the fact that he intentionally stabbed Caesar only because Brutus intended to end the tyranny by stabbing Caesar. There is no analogue to this in the case of causal explanations, and it is this, more than anything, that distinguishes the two types of explanation.
The foregoing examination of Donald Davidson’s philosophy of action reveals a number of respects in which he has erred. Any attempt to summarize these errors as species of a single error would certainly distort the relevant facts. Nevertheless, there does appear to be a common theme that runs through many of the issues we have discussed, and I believe that something of a general nature can be said about why Davidson tends to embrace the views that he does. In particular, I would say that much of Donald Davidson’s philosophy of action can be seen in terms of a commitment to a physicalist world-view.

It is well-known that Davidson embraces a physicalist theory of the mind, according to which every mental state—every belief, desire, and so on—is identical to a physical state. However, in speaking of a physicalist world-view, I mean to indicate something much more general than the position that Davidson espouses in the philosophy of the mind. A world-view is not necessarily a theory, but more like an attitude or disposition that serves to explain why one holds the various theories that one does.

The physicalist world-view to which I am alluding is associated with empiricism in general, and with the logical positivists in particular, who regarded physics as a paradigm of intellectual inquiry. Of course, it would be anachronistic, if nothing else, to accuse Davidson of being a positivist.
Nevertheless, it is clear that Davidson’s philosophy, like that of his most famous teacher, W.V. Quine, is steeped in the tradition of positivism. Aside from the various differences amongst these various philosophers, one of the things that unites them is a definite ontological bias towards that which is physical, and a corresponding suspicion towards that which is not.

In the philosophy of mind, Davidson betrays this ontological bias by treating "the mental" and "the physical" as different ways of describing events, and then insisting that every event has a description under which it is physical, while suggesting that not all events can be described as mental (Davidson 1970b, p.214). In the philosophy of language, on the other hand, Quine and Davidson both reveal their physicalist disposition in terms of a reluctance to countenance intensional entities, which Quine pejoratively calls "creatures of darkness" (Quine 1956, p.104). One place, of course, where intensional entities seem useful, if not indispensable, is in the semantic analysis of propositional attitude-ascriptions. However, Quine and Davidson have both attempted to analyze these sentences without recourse to such entities. Thus, according to Davidson’s paratactic analysis, attitude-ascriptions contain a two-place predicate that relates a person to an utterance, which is conceived as a concrete event.

Now we observed in the first chapter that there are significant doubts concerning Davidson’s paratactic analysis. In particular, we noted that the paratactic analysis of a sentence like "George IV is unaware that Scott is the author Waverly" leaves
one without any clear explanation for the inference from this last sentence to another token of the same type of sentence; it also leaves one without any clear explanation for the inference from the original sentence to "Scott is the author of Waverly." The problem here is not just that subsidiary premises are needed to explain these inferences, but that on the paratactic analysis it is not clear how these inferences can be explained even with subsidiary premises. Precisely the same point applies with respect to Davidson's analysis of ascriptions of intentional action, and to the paratactic analysis that we considered for reason-explanations.

The question of the logical form of attitude-ascriptions is a complicated one to which I have proposed no general solution. However, in light of the problems that confront the paratactic analysis, I think that one ought to reconsider Davidson's rejection of the analyses that make use of intensional entities. Davidson, as we have seen, claims to reject Frege's doctrine of indirect reference on logical rather than ontological grounds. However, we have also observed that Davidson's objection seems to be based upon a rather superficial defect in Frege's theory.

In the second chapter we considered Davidson's views on what an action is, and the relation between causally simple and causally complex actions. Here again, I think we encountered a form of physicalism. According to Davidson, if the queen kills the king by pouring poison in his ear, and pours poison in his ear by moving her hand in a certain way, then the queen has performed but one action of which three descriptions have been given. Evidently, what
leads Davidson to this view is the idea that since the queen does only one thing with her body—she moves her hand in a certain way—she therefore does only one thing. Accordingly, the other things that she apparently does by doing this one thing—such as pour poison in the king’s ear and kill the king—must just be different ways of describing this same thing. Davidson sums up this view by saying "our primitive actions,...mere movements of the body—these are all the actions there are" (Davidson 1971a, p.59).

We have already seen that this view rides roughshod over obvious part/whole distinctions and leads to contradiction. If the queen kills the king by moving her hand in a certain way, then the moving of her hand is not identical with her killing of the king, but is rather a causal component in her killing of the king. However, in order to appreciate this point one must first give up the mistaken idea that actions are mere movements of the body. Many of our actions involve no bodily movements at all, and many involve much more than mere movements of the body.

In the third chapter we considered and rejected Davidson’s most influential idea in the philosophy of action—namely, that reason-explanations are a species of causal explanation. In this case, I think that physicalist inclinations were made manifest in a couple of respects. In the first place, it is a striking feature of Davidson’s argument in "Actions, Reasons, and Causes" that causal explanations are taken for granted as a model of clarity in comparison to the obscurity associated with reason-explanations. Thus, Davidson writes that:
One way we can explain an event is by placing it in the context of its cause; cause and effect form the sort of pattern that explains the effect in a sense of "explain" that we understand as well as any. If reason and action illustrate a different pattern of explanation, that pattern must be identified. (Davidson 1963, p.10)

If the results of the last chapter are correct, then we have already identified the pattern of explanation involved in the case of reason-explanations, but the present point is to note Davidson's assumption that causal explanations represent a standard of clarity in terms of which other explanations are to be understood. This assumption, I believe, reflects a more general tendency to treat the physical sciences, where causal explanation is indeed the standard, as a sort of paradigm. But one of the results of this assumption, it seems, is a failure to appreciate the respects in which reason-explanations do differ from causal explanations.

Among the respects in which reason-explanations differ from causal explanations, the most important difference relates to the fact that reason-explanations depend for their truth or falsity upon whether they correctly identify, not what caused the agent to do what he did, but rather what the agent intended to do by doing what he did. Davidson's failure to appreciate this difference can be traced not only to his tendency to treat causal explanations as a sort of model of understanding, but also to an ontological prejudice against intentions, which he clearly demonstrates in "Actions, Reasons, and Causes."

There are, then, a number of respects in which physicalist inclinations tend to mislead in the philosophy of action. However, it would be a mistake to recoil from this into some sort of
immaterialism. Once again, the physicalist world-view to which I am attributing some of the flaws in Davidson’s philosophy of action is not a well-defined theory in opposition to other theories, but is rather a disposition to theorize in a certain sort of way.
Bibliography


Brandl, Johannes and Gombocz, Wolfgang (eds.) *The Mind of Donald Davidson*. Amsterdam: Rodopi.


(1973d) "Psychology as Philosophy," reprinted in Davidson 1980, pp.229-44.


---- (1972) "Quantification and Ontology," *Nous* v.6, pp. 240-250.


---- (1961) "Reasons for Action and Matters of Fact," A.P.A. Presidential Address, delivered at the 35th annual meeting at the University of British Columbia.


Appendix to Chapter 1

Arguments:

A1  (1a) If the man in the brown hat is a spy, someone is a spy. 
     (1b) The man in the brown hat is a spy.  
          -----------------------------------------  
     (1c) Someone is a spy.

A2  (1b) The man in the brown hat is a spy.  
          -----------------------------------------  
     (1c) Someone is a spy.

A3  (2d) Brutus stabbed Caesar with his knife.  
          -----------------------------------------  
     (2a) Brutus stabbed Caesar.

A4  (1d) The man in the brown hat, who is a clergyman, is a spy.  
          -----------------------------------------  
     (1b) The man in the brown hat is a spy.

A5  (1e) Someone who is a clergyman is a spy.  
          -----------------------------------------  
     (1c) Someone is a spy.

A6  (2e) Brutus stabbed Caesar with his knife in the Forum.  
          -----------------------------------------  
     (2d) Brutus stabbed Caesar with his knife.

A7  (2f) Brutus fatally stabbed Caesar.  
          -----------------------------------------  
     (2a) Brutus stabbed Caesar.

A8  (2k) Brutus intentionally stabbed Caesar.  
     (2l) Caesar is identical to the inventor of the Julian calendar.  
          -----------------------------------------  
     (2m) Brutus intentionally stabbed the inventor of the Julian calendar.

A9  3a) Ralph believes that the man in the brown hat is a spy.  
     3b) The man in the brown hat is the man at the beach.  
          -----------------------------------------  
     3c) Ralph believes that the man at the beach is a spy.

A10 (6a) Brutus stabbed Caesar because he wanted to end the tyranny.  
     (6b) Caesar is the inventor of the Julian calendar.  
          -----------------------------------------  
     (6c) Brutus stabbed the inventor of the Julian calendar because he wanted to end the tyranny.
Sentences:

(1a) If the man in the brown hat is a spy, someone is a spy.
(1b) The man in the brown hat is a spy.
(1b**) Spy(\text{the man in the brown hat}).
(1c) Someone is a spy.
(1c**) \((\exists x) \text{Spy}(x)\).
(1d) The man in the brown hat, who is a clergyman, is a spy.
(1d*) Clergyman(\text{the man in the brown hat}) \&
       \text{Spy}(\text{the man in the brown hat}).
(1d**) \((\exists x) \text{Clergyman}(x) \& \text{Spy}(x) \&
       x = \text{the man in the brown hat})\).
(1e) Someone who is a clergyman is a spy.
(1e*) \((\exists x) \text{Clergyman}(x) \& \text{Spy}(x))\).
(2a) Brutus stabbed Caesar.
(2a*) Stabbed(\text{Brutus}, \text{Caesar}).
(2a**) \((\exists x) \text{Stabbing}(x, \text{Caesar}, \text{Brutus}))\).
(2a***) \((\exists x) \text{Stabbing}(x) \& \text{Of}(x, \text{Caesar}) \& \text{Agent}(x, \text{Brutus}))\).
(2b) Someone stabbed Caesar.
(2c) Brutus stabbed someone.
(2d) Brutus stabbed Caesar with his knife.
(2d*) Stabbed-\text{With}(\text{Brutus}, \text{Caesar}, \text{Brutus' knife}).
(2d**) \((\exists x) \text{Stabbing}(x, \text{Caesar}, \text{Brutus}) \&
       \text{With}(x, \text{Brutus's knife}))\).
(2e) Brutus stabbed Caesar with his knife in the Forum.
(2f) Brutus fatally stabbed Caesar.
(2f*) \((\exists x) \text{Stabbing}(x, \text{Caesar}, \text{Brutus}) \& \text{Fatal}(x))\).
(2g) Brutus did something.
(2h) Something was done to Caesar.
(2i) There was a stabbing.

(2j) Brutus allegedly stabbed Caesar.

(2j*) (\exists x)(\text{Stabbing}(x, \text{Caesar}, \text{Brutus}) \& \text{Allegedly}(x)).

(2j**) (\exists x)(\text{Stabbing}(x) \& \text{Of}(x, \text{Caesar}) \& \text{Agent}(x, \text{Brutus}) \& \text{Allegedly}(x)).

(2k) Brutus intentionally stabbed Caesar.

(2k*) (\exists x)(\text{Stabbing}(x, \text{Caesar}, \text{Brutus}) \& \text{Intentionally}(x)).

(2k**) (\exists x)(\text{Stabbing}(x) \& \text{Of}(x, \text{Caesar}) \& \text{Agent}(x, \text{Brutus}) \& \text{Intentionally}(x)).

(2k,*) \text{Intentional(} \text{Brutus, that)}
\quad (\exists x)(\text{Stabbing}(x, \text{Caesar, him})).

(2k,**) \text{Intentional(} \text{Brutus, that)}
\quad (\exists x)(\text{Stabbing}(x) \& \text{Of}(x, \text{Caesar}) \& \text{Agent}(x, \text{he})).

(2l) Caesar is identical to the inventor of the Julian calendar.

(2m) Brutus intentionally stabbed the inventor of the Julian calendar.

(3a) Ralph believes that the man in the brown hat is a spy.

(3a*) \text{Believes-That(} \text{Ralph, Spy(} \text{the man in the brown hat})\).

(3a,*) \text{Believes(} \text{Ralph, that) Spy(} \text{the man in the brown hat})\).

(3a,**) (\exists x)(\text{Believing}(x, \text{Ralph, that}))
\quad \text{Spy(} \text{the man in the brown hat})\).

(3a,***)(\exists x)(\text{Believing}(x) \& \text{Agent}(x, \text{Ralph}) \& \text{Of}(x, \text{that}))
\quad \text{Spy(} \text{the man in the brown hat})\).

(3b) The man in the brown hat is the man at the beach.

(3c) Ralph believes that the man at the beach is a spy.

(3d,1) Ralph firmly believes that the man in the brown hat is a spy.

(3d,*) (\exists x)(\text{Believing}(x, \text{Ralph, that}) \& \text{Firm}(x))
\quad \text{Spy(} \text{the man in the brown hat})\).

(4a,1) Quine believes that Ralph believes that the man in the brown hat is a spy.
(5a) Jack’s fall caused the breaking of his crown.

(5a*) Caused((ix)(Falling(x, Jack), (ix)(Breaking(x, Jack’s crown))).

(5b) Jack fell down, which caused it to be the case that Jack broke his crown.

(5b*) (∃x)(∃y)(Fall(x, Jack) & Breaking(y, Jack’s crown) & Caused(x,y)).

(5c) The fact that Jack fell down explains the fact that Jack broke his crown.

(5c,*) Explains(that, that)
(∃x)(Fall(x, Jack) (∃x)(Breaking(x, Jack’s crown))).

(6a) Brutus stabbed Caesar because he wanted to end the tyranny.

(6a,*) Explains(that, that)
Wanted(Brutus, that) (∃x)(Ending(x, the tyranny, Brutus)) (∃x)(Stabbing(x, Caesar, Brutus)).

(6a,**) Explains(that, that)
(∃x)(Wanting(x) & Of(x, that) & Agent(x, Brutus))
(∃x)(Ending(x) & Of(x, the tyranny) & Agent(x, Brutus))
(∃x)(Stabbing(x) & Of(x, Caesar) & Agent(x, Brutus)).

(6b) Caesar is the inventor of the Julian calendar.

(6c) Brutus stabbed the inventor of the Julian calendar because he wanted to end the tyranny.

(6d) Brutus wanted to end the tyranny.

(6d,*) Wanted(Brutus, that)
(∃x)(Ending(x, the tyranny, him)).