ACTIONS AND EVENTS:

A STUDY IN ONTOLOGY AND ETHICS

by

Tracy Lynn Isaacs

B.A. Scarborough Campus, University of Toronto (1987)

M.A. University of Toronto (1988)

Submitted to the Department of
Linguistics and Philosophy
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY in Philosophy

at the

Massachusetts Institute of Technology May, 1992

© Tracy Lynn Isaacs

The author hereby grants to MIT permission to reproduce and to distribute copies of this thesis document in whole or in part.

Signature of author	
	Department of Linguistics and Philosophy, May 1992
Certified by	r vvina aure v v v valentina
	Judith Jarvis Thomson
	Professor, Philosophy
	Thesis Supervisor
Accepted by	
	George Boolos
	Chair Committee on Graduate Studies

COmmittee on Graduate Studies
ARCHIVES Philosophy

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JUN 1 1 1992

For Norma and Ray, my wonderful parents

TABLE OF CONTENTS

Abstract
Acknowledgements
Essay One: The Relationship Between Actions
Reported by "By"-Sentences7
Part One
1 The Question8
2 Parthood Theory12
2.1 Challenging Parthood Theory14
2.2 Parthood Theory Responds
2.3 Assessing Parthood Theory's Response
3 Trope Theory
3.3 Assessing Trope Theory's Response25
3.4 Conclusions about Trope Theory26
4 Identity Theory26
4.1 Challenging Identity Theory27
4.2 Identity Theory Responds to
the First Challenge28
4.3 Identity Theory Responds to
the Second Challenge
4.4 Identity Theory Responds to
the Third Challenge32
4.5 Assessing Identity Theory's Response35
5 Conclusions, so far
Dart Two. You to analyze them
Part Two: How to analyze "by"
1 "By" as the Relation of Identity on Events38
2 "By" as Syncategorematic: Forming
Complex Verbs out of Simpler Verbs40
3 "By" as Expressing the Part-Whole
Relation on Events42
4 "By" as Syncategorematic:
Causal Verbs43
5 "By" as a Relation on Facts48
Conclusion57
Appendix A59
Bibliography60
Essay Two: The Ontology of Action and Ethics61
1 What is Intrinsic Value?
2 Acts Do Not Have Intrinsic Value:
An Argument64
The state of the s
Value of Actions
4 Can the Act Have the Disvalue That It Needs?71
5 Instrinsic Value, Utilitarianism.

		and the Mafia Case76
	6	Conclusions about Intrinsic Value
		and Utilitarianism77
	7	Instrinsic Value and Moral Theory79
	8	Ontology and Ethics80
	9	Ethics and the Claim That There
		Are No Actions83
	10	Horgan's Project85
	11	Horgan's Project and Ethics88
	12	What Has Horgan Shown?90
	Con	clusion92
	Bib	oliography95
Essay	Th.	ree: Responsibility and Autonomy96
-	1	The Choice View
	2	Wolf's Critique of The Choice View:
		The Reason View
	3	Amy, Rachel, and the Ability to
		Do Otherwise107
	4	Amy and the Desirability of
		the Ability to do Otherwise110
	5	Do We Need the Praiseworthiness Clause?112
	6	Conclusion118
	Bib	aliography 122

ACTIONS AND EVENTS:

A STUDY IN ONTOLOGY AND ETHICS

by

TRACY LYNN ISAACS

Submitted to the Department of Linguistics and Philosophy on May 1, 1992 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Philosophy

ABSTRACT

The philosophy of action is about agents and actions. As such, it has both a metaphysical and an ethical dimension. My dissertation is divided into three papers.

The first is wholly metaphysical, concentrating on the ontology of actions. I explore the relationship between actions reported by a certain class of "by"-sentences and argue that the relationship is identity.

The second paper concerns the bearing that ontological conclusions about actions have on ethics. I argue that, except for the claim that there are no such entities as actions, ontological conclusions about actions do not make a substantive difference to ethical theory.

In the third paper, I concentrate on the ethical issue of an agent's responsibility for her actions. I argue that an agent is responsible only if she could have done otherwise, as long as "could have done otherwise" is interpreted as meaning "would have done otherwise if she had so chosen."

Thesis Supervisor: Judith Jarvis Thomson

Title: Professor of Philosophy

ACKNOWLEDGEMENTS

Many people helped me in many ways throughout the four years that I have been at MIT.

I am especially grateful to Judith Jarvis Thomson, my thesis supervisor, for her excellent philosophical guidance and her much appreciated encouragement. I have learned very much from her since my arrival at MIT. Robert Stalnaker and Richard Cartwright have also been invaluable members of my thesis committee. Their comments on many drafts helped me both to formulate my ideas and to express them more clearly.

I also owe thanks to the following people for friendship, philosophical insights, advice, and in general for helping to get me through the many ups and downs of being a graduate student: David Brink, Lenny Clapp, Ronald de Sousa, Tim Dilligan, Diane Jeske, Rob Stainton, Ed Stein, and Daniel Stoljar.

My parents, Norma and Ray Isaacs, have contributed ongoing and incomparable emotional support and encouragement, as well as very generous financial assistance. Lots of love and many thanks to them for everything.

ESSAY ONE

THE RELATIONSHIP BETWEEN ACTIONS REPORTED BY "BY"-SENTENCES

PART ONE

form

1 The Question

Suppose

is true. Then it would be natural to conclude that there was an event that was a killing of Caesar by Brutus and that there was an event that was a stabbing of Caesar by Brutus. Both of these events are actions of Brutus's. In what relation does the event that is the stabbing stand to the event that is the killing?

Given the truth of a certain class of sentences with the

(2) X does alpha BY doing beta in which X is the agent and "does alpha" and "doing beta" are replaced with action verbs, it is natural to conclude that there is an event that is X's doing alpha and there and event that is X's doing beta, and that these events stand in some relation to

The scope of my discussion is going to be restricted in two ways. First, I am going to limit the verbs under discussion to action verbs as we commonly construe them without addressing what sets acts apart from other events and exactly which verbs fall into the set of action verbs. Second, there are some action sentences which do not offer themselves up for the kind of analysis that I am suggesting. For example, "By staying up all night and cramming, I passed the exam" needs a different sort of analysis from "Brutus killed Caesar by stabbing him." I shall not be addressing the former, and do not know how to characterize them except to say that there are going to be some exceptions that do not fit any of the proposed analyses. Now, this is not to say that I am going to dismiss all of the difficult cases as There is a large literature surrounding a certain paradigm of controversial sentences, such as "Brutus killed Caesar by stabbing him" that require an analysis. They are by no means "the easy cases." These are the cases with which I shall be concerned. However, I am not going to propose any systematic way of delimiting the set of sentences in question.

one another. I shall consider three theories, each of which yields an answer to the question:

beta when X does alpha by doing beta?

Two of the theories, Parthood Theory and Trope Theory, though they yield an answer to this question, were not developed with this question in mind. Identity Theory, by contrast, developed out of a question about the relation between actions reported by

"by"-sentences.

In what relation does X's doing alpha stand to X's doing

Once we have seen what each of these theories says generally about this issue, I shall consider specific ways in which each of the theories might analyze "by" in order to capture the relationship that it proposes between the agent's doing alpha and the agent's doing beta.

The first theory, Parthood Theory, was designed to give an account of the part-whole relation as it applies to acts and events. In spelling out the details of the relationship between events and their parts, Parthood Theory yields an answer to our question. It says that X's doing beta may be identical with X's doing alpha or it may be a proper part of the X's doing alpha. What Parthood Theory decides in each case depends upon whether the "by" statement is timespanning or synchronous. A statement is timespanning if, as Jonathan Bennett puts it, "what makes it true that X [does alpha] by [doing beta] is the fact that because X [does beta] he caused something to occur later" (Bennett, p.

194). 2 Statements like "Brutus killed Caesar by stabbing him" and "he poisoned the inhabitants by operating the pump" are timespanning. By operating the pump, the pumper caused it to be the case that the inhabitants later became poisoned. By stabbing Caesar, Brutus caused it to be the case that Caesar later died. By contrast, the truth of synchronous "by" statements is not dependent upon the agent's causing something to occur later on. For example, "by crossing the finish line first he won the race" is synchronous because his crossing the finish line first did not cause him to later win the race, it made it the case at that very moment that he won the race. Parthood Theory says that in timespanning cases, X's doing beta is a proper part of X's doing alpha and as such, X's doing beta does not take as much time as X's doing alpha. In synchronous cases, X's doing beta is identical to X's doing alpha and as such, they have exactly the same temporal extent. With respect to our original example, Parthood Theory claims that the killing and the stabbing are different actions, and that the stabbing is a proper part of the The stabbing is complete before the killing is because the killing, unlike the stabbing, is not complete until Caesar's It results from this theory that actions include as parts death. much more than bodily movements. They can and often do extend beyond the movement of the body to include some of the movement's effects. Judith Jarvis Thomson's account of Parthood Theory is

² All references to Bennett are from Jonathan Bennett's Events and Their Names (Cambridge: Hackett Publishing Co., 1988).

the version of it that I shall concentrate on.

The second theory under consideration will be Bennett's Trope Theory of events. Trope Theory is a detailed general theory of events. It provides an account of what events are. Although its primary aim is much more ambitious, in spelling out the theory of events, it yields an answer to our question. Trope Theory gives the same general answer as Parthood Theory does. That is, Trope Theory says that for timespanning cases, X's doing alpha and X's doing beta are different events because they are instances of different properties. It asserts an identity in synchronous cases. However, the reason Trope Theory gives for this answer is very different from Parthood Theory's, for Trope Theory disagrees with Parthood Theory that X's doing alpha and X's doing beta have different temporal extents. With respect to our original example, Trope Theory says that Brutus's killing of Caesar and Brutus's stabbing of Caesar are co-temporal yet different events. When I present the detailed account of the theory, it will become clear how it supports this claim.

The third theory, and the one that I shall be defending is the Identity Theory. It was initially developed by Elizabeth Anscombe to answer the question about "by" with which we are concerned. This theory says that when X does alpha by doing beta, then X's doing alpha is the same entity as X's doing beta. With respect to our original example, "the killing of Caesar by Brutus" and "the stabbing of Caesar by Brutus" are different ways of describing the same entity. It turns out on Identity Theory

that every action is identical to a bodily movement. As well as Anscombe, advocates of the theory include Donald Davidson and Jennifer Hornsby.

Let us assess each theory for its strengths and weaknesses.

2 Parthood Theory

Parthood Theory says that, in timespanning cases, if someone does alpha by doing beta then the action that is X's doing alpha is different from the action that is X's doing beta. The claim is that X's doing beta is a proper part of X's doing alpha but is not identical with it. Thus, when Brutus kills Caesar by stabbing him, Brutus's stabbing of Caesar is a proper part of Brutus's killing of Caesar. A Parthood Theorist denies an identity in this case, claiming that the killing has parts that the stabbing does not have. For example, the killing includes the death of Caesar as a part, whereas the stabbing does not. Not only does this indicate that they have different parts, but it also indicates that they have different temporal extents. Since the death of Caesar, which is a part of the killing, occurs after the stabbing, Parthood Theory says that the killing lasts longer than the stabbing does. Since the killing and the stabbing have different parts and different temporal extents, they must be different events.

Parthood Theory has two features which make it seem an attractive way of answering the question. First, it agrees with our usual say of speaking about the causal relations between

actions and the actions and events that they cause. We never say, for example, that killings <u>cause</u> deaths. We shall see that Identity Theory does make this claim, and needs to explain this oddity. Since a death is a part of a killing for Parthood Theory, killings do not cause deaths, they include them, as noted above, as proper parts. This is not to say, of course, that a killing can <u>never</u> cause a death on the parthood account.

Brutus's killing of Caesar might have so shocked Marcus that the killing <u>caused</u> Marcus to have a heart attack and die. If this were the case, then it would be true that Brutus's killing of Caesar caused Marcus's death, but it certainly would not, if we were following Parthood Theory, be true that it caused Caesar's death.

The second appealing feature of Parthood Theory is that it accounts for parts of actions in exactly the same way that it accounts for parts of other events. Actions and events are not treated differently in their metaphysical analyses on the parthood account. Since actions are events, it makes sense to think that they can be complex entities, composed of other events and other actions. Parthood Theory spells out the complexity of actions in some detail, providing an account of how the parts relate to the whole. For example, a very rudimentary analysis of Brutus's killing of Caesar might look something like this:



The killing is composed of events causally related in a specific

way which makes them parts of the whole action "Brutus's killing Caesar." Since it attempts to spell out the part-whole relation with respect to actions, Parthood Theory seems a promising route to take if we are concerned with the complexity of actions.

2.1 Challenging Parthood Theory

The theory faces one major challenge to which I do not believe it can provide a convincing reply. This is an argument to the effect that the temporal claim it makes in timespanning cases is false. If Parthood Theory's account of the times of actions is wrong, then so also is its account of the parts of actions. For this account of the times of actions depends for its plausibility upon the way that the theory breaks actions down into parts. As a result, the challenge is a very serious one.

Both Identity Theory and Trope Theory reject the temporal claim using the No Further Effort argument (Bennett names it "No Further Effort" in Bennett, p. 196). Using the example of Brutus's killing of Caesar, this argument says:

Actions are over when the agent is no longer doing anything relevant to the action.

Once Brutus has stabbed Caesar he is no longer doing anything relevant to the killing of Caesar.

Therefore, once the stabbing is over, the killing must be over.

The argument can be generalized to all timespanning cases of "by" statements. Whenever an agent X does alpha by doing beta, the No Further Effort argument says that X's doing alpha is over when X's doing beta is because the agent no longer needs to do

anything relevant to his doing alpha. The intuition behind the argument is that once there is no further effort required from the agent to complete the action, the action is complete. This argument also accounts for our intuitions regarding synchronous cases. In these cases, no further effort is required on the part of the agent in order for the action to be complete.

I believe that this intuition is a very strong one, and that rejecting the intuition and the No Further Effort argument leads to some conclusions that are not only unusual, but are also implausible. In particular, Parthood Theory's way of understanding the times of actions yields the odd result that actions can continue long after the agent whose actions they are is dead. Thus, Brutus might die immediately after stabbing Caesar but before Caesar is dead, yet Parthood Theory would say that Brutus's action of killing Caesar continues nonetheless. It seems very implausible to claim that Brutus's actions continue after he is dead, and this result gives us a good reason to accept the No Further Effort argument. It is much more plausible to claim that Brutus's actions are all over once he is dead. Naturally, things that Brutus has done may cause events that happen after he dies and these events may warrant new ways of describing his action, but he certainly cannot do anything further once he is dead; his actions must be complete.

2.2 Parthood Theory Responds

In <u>Acts and Other Events</u>, Thomson presents an example in an attempt to undermine the No Further Effort argument.

In the example, Judy cleans the kitchen floor by pouring Stuff on it, waiting for Stuff to dissolve the dirt as it dries, and then sweeping up the residue. According to Thomson, Judy is cleaning the floor from the time she pours on Stuff until the time it finishes evaporating, and throughout the time she is sweeping up the dust. Thomson illustrates Judy's cleaning the floor as follows (rectangles indicate time-stretches; points indicate time-points):



It seems very plausible on any theory to say that Judy's cleaning the floor is not complete until the end of T". According to Parthood Theory, Judy is cleaning the floor from the beginning of T until the end of T', and throughout T". Stuff's dissolving the dirt while Judy waits for it to dry is as much a part of Judy's cleaning the floor as her pouring Stuff on the floor and her sweeping up the dust are.

Since we find it plausible that Judy's cleaning continues beyond Judy's pouring Stuff, Thomson urges that we should find it equally plausible in the following case that Diane's cleaning

Acts and Other Events, Judith Thomson (Ithaca: Cornell University Press, 1977).

continues beyond her pouring her product. Diane lives down the street from Judy and she uses a newer product - Super-Stuff.

Super-Stuff not only dissolves the dirt, it also disintegrates it. Thus, once Diane pours Super-Stuff on the floor, she does not have to do anything more towards cleaning the floor. Her action is illustrated as follows:



She pours Super-Stuff on the floor at T, Super-Stuff does its work between T and T'. T' marks the time that the last of Super-Stuff evaporates, and, according to Parthood Theory, the completion of Diane's cleaning the floor. Advocates of the No Further Effort argument would, of course, claim that Diane's cleaning of the floor is complete at the end of T, once she has finished pouring Super-Stuff on the floor.

The force of the two examples is supposed to be that if we are going to claim that Judy's cleaning the floor is going on while Stuff is doing its work, then we must agree with Parthood Theory's claim that Stuff's doing its work is part of Judy's cleaning. If we say that, it appears that, analogously, Diane's cleaning the floor is going on while Super-Stuff does its work: we have no basis upon which to deny that Super-Stuff's doing its work is part of Diane's cleaning. If anything, Super-Stuff appears to play a larger part in Diane's cleaning her floor than

Stuff does in Judy's cleaning her floor. This undermines the No Further Effort Argument because it looks as if Diane's action is not over when she has no more effort to put into it.

2.3 Assessing Parthood Theory's Response

This example does not give us a convincing reason for giving up the No Further Effort argument. There is no dispute that if we say that Stuff's dissolving and disintegrating the dirt is a part of Judy's action, then we need to say that Super-Stuff's dissolving and disintegrating the dirt is a part of Diane's action. However, what I and other advocates of the No Further Effort argument deny is the claim that anything Stuff or Super-Stuff does is a part of any of the actions of Judy and Diane. One good reason for denying this comes out when we consider what would happen if the agents were to die immediately after pouring their products. Diane's floor would become clean, Judy's would not. Now unless we want to say that an agent's actions can continue after she is dead, Diane had better be done with her cleaning when she no longer has anything to do for her floor to become clean. What Stuff and Super-Stuff do subsequently are consequences of actions of Judy and Diane. It is what Judy and Diane themselves do, not what the products they use do, that counts as part of their action. This is not to say that Judy's cleaning the floor is complete once she has poured Stuff on it. It cannot be complete because Judy still has more work to do after Stuff does its work.

The question we need to address is: what makes it plausible to say in this case that Judy is not done right after she pours Stuff, while Diane is done cleaning after she pours Super-Stuff? Is it, as Parthood Theory maintains, implausible to claim that Judy's cleaning is not complete after she pours Stuff, whereas Diane's cleaning is complete after she pours Super-Stuff? No, we may plausibly make this claim because there is a crucial difference between Judy's cleaning the floor and Diane's cleaning the floor. It seems correct to say that Judy is not finished cleaning the floor in this case, not because Stuff is doing something, but rather because Judy still has something to do. By contrast, Diane has nothing left to do once she has poured Super-Stuff. Advocates of No Further Effort agree with Parthood Theory that Judy's action is not over as soon as she has poured Stuff on the floor.

On the face of it, the intuition that these examples rely on is the intuition that a floor-cleaning is not over until the floor is clean. But if that is the point being made, then there is no need for two separate cases to make the point. The claim that Diane's floor is not clean as soon as she pours Super-Stuff is uncontroversial. Super-Stuff still has some work to do. The controversial question is whether the action that is Diane's cleaning the floor is complete once she has poured Super-Stuff. The reason for starting with the Stuff example and subsequently introducing the Super-Stuff example is that in the Stuff example no one denies that Judy's cleaning the floor is not over as soon

as Judy has poured Stuff. But advocates of No Further Effort are quick to agree to this claim because Judy still has work to do, not because they concede that Stuff's doing its work is part of Judy's cleaning the floor. It is the fact that Judy still has some work to do, not the claim that the cleaning is not over until the floor is clean, that gives the Stuff example its initial force. We can see that this is the case by looking at a variation of this example in which Judy's task is simplified. Suppose that she needs only to sweep the floor; there is no Stuff involved at all. She sweeps one half, takes a rest, and then sweeps the other half. Any advocate of No Further Effort would say that Judy's sweeping the floor in incomplete after she has swept the first half because she still has some work to do; Judy has not finished doing everything that she needs to do to get the floor clean until she sweeps the other half of the floor. original example, Stuff has very little to do with Judy's action of cleaning the floor other than the fact that it gives Judy a reason to take a break. Judy's cleaning the floor consists in what Judy has to do; Stuff's cleaning the floor consists in what Stuff has to do. Judy's cleaning the floor is not the same as Stuff's cleaning the floor. Judy's cleaning the floor is not over once she pours Stuff because Judy still has work to do. Diane's cleaning the floor is over once she pours the Super-Stuff because Diane has nothing left to do that is relevant to her cleaning the floor. The rest is up to Super-Stuff, and its doing its work is not a part of Diane's doing hers.

Both Parthood Theory and theories that accept No Further

Effort need to claim that Judy's action is divided. Parthood

Theory has it divided during the time between the last of Stuff's dissolving and Judy's sweeping up the dust. The No Further

Effort account has it divided during the time between Judy's pouring and Judy's sweeping.

I believe that Parthood Theory has not provided an adequate argument to counter the No Further Effort argument. In fact, the plausibility of the claim that Judy's cleaning the floor is not over once she pours Stuff depends on the very intuitions behind the No Further Effort argument. In light of this weakness of Parthood Theory, I think we should set aside its claim that in timespanning cases of "by"-statements, the agent's doing beta is a part of the agent's doing alpha and as a result, different from it.

In considering Trope Theory and Identity Theory, both of which use the argument from no Further Effort to establish the times of actions, we shall see that No Further Effort is compatible with two very different accounts of the relationship between X's doing alpha and X's doing beta when X does alpha by doing beta.

3 Trope Theory

With respect to the question posed at the outset, Trope
Theory says that Brutus's killing Caesar and Brutus's stabbing

Caesar, just like any timespanning case of an agent's doing alpha by doing beta, are co-temporal yet different events. What remains unclear, however, is the theoretical machinery that allows Trope Theory to make this claim seem like a plausible one. Before we can assess the claim, we need to have a look at the theory.

According to Trope Theory, events, and hence actions, are tropes. A trope is an instance of a property. Bennett provides an example:

My house is a concrete particular that has whiteness and other properties; whiteness is an abstract universal that is possessed by my house and other particulars; and the whiteness of my house is a trope, an abstract particular. It is unlike my house in that all that there is to it is whiteness, and it is unlike whiteness in that it pertains only to my house. (Bennett, p. 15-16).

As an instance of a property, a trope is both abstract and particular. It is abstract because it is a property but particular because it is just one instance of the property. While the example of "the whiteness of my house" gives us some idea of what a trope is, event tropes can be much more complex entities.

Initially, based on Jaegwon Kim's theory that an event is the instantiation of a property by a substance at a time, Bennett says that an event is the instantiation of a complex property P* at a spatiotemporal zone. He interprets this as meaning that each event is an <u>instance</u> of a complex property that occurs at a certain spatiotemporal zone. The zone will usually be delimited by a substance and a time, but Bennett leaves it open that this

may not always be the case. As an <u>instance</u> of a property, an event is a trope. Every event is to be understood as an instance of a <u>complex</u> property P* that constitutes it. According to the theory, the whole intrinsic truth about a given event e is that it is an instance of the complex property that constitutes it, i.e. that it is an instance of the constitutive property P*. Given the complexity of events, we cannot know exactly what property P* constitutes any given event e. But for illustrative purposes, Bennett suggests the following example:

suppose that e is an instance - in a certain pebble S at a time T - of the property of falling with an acceleration of 32ft/sec² while rotating .68 times per second on an axis at right angles to the line of the fall (Bennett, p. 94)

All of the properties of the fall outlined above are included in the property that constitutes the event. That is, the whole intrinsic truth about this event is that it is "an instance - in a certain pebble S at a time T - the property..." and so on. The parts of the complex trope that is constituted by P* are themselves tropes, each constituted by simpler properties P. Since the constitutive property P* is so complex, it need not be the case that every term used to refer to the event refers to every part of the event. For example, "this stone's fall" and "this stone's dangerous fall" can refer to the same event. The latter simply provides more information about it.

Different events can occupy the same zone if they are differently constituted, that is, if the constitutive properties are different. In short, more than one trope can occupy the same spatiotemporal zone. This is what makes it possible for Trope

Theory to say that Brutus's killing Caesar and Brutus's stabbing Caesar are different events even though the No Further Effort argument shows them to be co-temporal. They have different constitutive properties, yet they occupy the same spatiotemporal zone. According to Trope Theory, it is generally the case that in timespanning cases, X's doing alpha and X's doing beta are co-temporal but not identical.

Trope Theory is attractive because it provides a detailed account of what events are and, if correct, would thereby prove to be theoretically useful for much more than giving an answer to the question at hand. In addition to being so ambitious in its goals, Trope Theory is consistent with the No Further Effort Argument. Hence, it is not subject to some of the temporal difficulties of Parthood Theory.

3.1 Challenging Trope Theory

Trope Theory faces one serious and fundamental difficulty. That is, we have good reason to question whether we should posit the existence of tropes at all. In using tropes to characterize events, we are explaining one problematic entity in terms of another that is even more troublesome. Why should we think that instances of properties form a distinct and fundamental ontological category? Ontological simplicity demands that we have good reasons before introducing a new entity into our ontology. Bennett has not been successful at providing us with good reasons.

3.2 Trope Theory Responds

Bennett is well aware that someone might consider tropes to be a gratuitous and pointless addition to our ontology. Using a non-event trope to illustrate, he outlines how the objection might go:

This stone is a particular substance; its shape is a universal property, <u>flatness</u>. The friends of tropes are trying to introduce a third item that is particular rather than universal but is a property rather than a substance, namely, the flatness of this stone. This is a gratuitous, pointless addition (Bennett, p. 90).

Bennett maintains that if we run a parallel objection in terms of a property that we think of as construing an event, the objection loses its force:

This stone is a particular substance; its way of moving is a universal property, <u>falling</u>. The friends of tropes are trying to introduce a third item, namely, the fall of this stone...(Bennett, p. 90).

At this point, Bennett claims, the objection peters out.

3.3 Assessing Trope Theory's Response

It certainly does look as if the objection peters out at this point. However, Bennett's construal of why it peters out is somewhat misleading. When rephrased in terms of events, the objection appears to lose its force because it is not the existence of events, such as this stone's fall, that is at issue. What is at issue is whether the stone's fall is a trope. That is, is an event an instance of a complex property at a zone? By reconsidering the objection in terms of a "property that is more

naturally thought of as constituting an event" (Bennett, p. 90),
Bennett is already assuming that events are tropes. The
challenger of Trope Theory need not deny the existence of events.
What she wants, and what she has not been given, is an argument
to the effect that events are tropes.

3.4 Conclusions About Trope Theory

The upshot of this discussion is that Trope Theory rests on the dubious assumption that there are tropes. If we have no good reason to suppose that there are entities that are tropes, then we have no good reason to suppose that the best way to explain actions and events is in terms of them.

The most attractive feature of Trope Theory is that it can give a plausible account of the times of actions because it is consistent with the No Further Effort argument. In addition, it allows for some leniency in the way an event in described. The next theory that we shall consider, Identity Theory, has these features and an additional one besides - simplicity.

4 Identity Theory

Identity Theory is appealing above all for its simplicity.

Unlike the other theories, Identity Theory is much more modest in its goals. It does not spell out the part-whole relationship in events, nor does it give a detailed account of what event are.

But it is often the complexity of Parthood Theory and Trope

Theory that leads them into difficulties that the more modest

goals of Identity Theory make it easy to avoid.

Identity Theory gives us the same answer in both synchronous and timespanning "by" statements. For any case in which an agent X does alpha by doing beta, X's doing alpha is X's doing beta. With respect to the original example, Brutus's killing Caesar is the same entity as Brutus's stabbing Caesar. "Brutus's killing Caesar" and "Brutus's stabbing Caesar" are just two different ways of describing it. Thus, like Trope Theory, Identity Theory does not tie the reference of an event description directly to the way in which the event is described. It falls out of Identity Theory that every action is identical to a bodily movement.

4.1 Challenging Identity Theory

Identity Theory, although attractive, faces several challenges. The first is a temporal one, familiar from the discussion of Parthood Theory. Since Identity Theory claims that Brutus's killing Caesar is identical with Brutus's stabbing Caesar, the killing and the stabbing must have the same timespan. Once the stabbing is over, the killing is over, even if Caesar is not yet dead. However, according to Parthood Theory, the killing took longer than the stabbing. Parthood Theory says that it is

^{*} At least, every action that is done by doing something that is a bodily movement is a bodily movement. Mental activities, such as adding in one's head or making a decision, are not bodily movements in the appropriate sense. I think that a case can be made that they should not count as actions in any case, but I shall not attend to that issue here.

not true to say that Brutus <u>killed</u> Caesar before Caesar is dead because the killing was not complete until Caesar's death. If the killing, but not the stabbing, includes the death, that would explain why the killing took longer than the stabbing.

The second challenge Identity Theory faces is that it does not seem correct to say, for example, that the killing <u>caused</u> the death. Yet if Brutus's killing Caesar and Brutus's stabbing Caesar are the same entity, and if, as we all agree, the stabbing caused the death, then it must be the case that the killing caused the death. According to Parthood Theory, to say that killings cause deaths is odd enough that it should at least give us pause. We simply do not speak of killings causing deaths. This problem exists for many of the timespanning "by" statements. Parthood Theory agrees with our usual way of talking. Identity Theory does not.

The third challenge to Identity Theory concerns necessary properties. With respect to timespanning cases in particular, the agent's doing alpha and the agent's doing beta usually have different necessary properties. Killing, but not stabbings, necessarily result in deaths; poisonings, but not pumpings, necessarily result in people being poisoned. Yet if Brutus's stabbing and Brutus's killing are the same entity, then they ought to have the same necessary properties.

4.2 Identity Theory Responds to First Challenge

In order to meet the first objection - the temporal

challenge - Identity Theory has to make a convincing case for the claim that Brutus's killing Caesar does not last longer than Brutus's stabbing Caesar, but rather takes exactly the same amount of time. Identity Theory is consistent both with the claim that it sounds odd to say that Brutus killed Caesar before Caesar is dead, and with the claim that we might be unjustified in saying that Brutus killed Caesar until after we know Caesar is going to die. But the fact that something sounds odd or that we might be unjustified in saying something does not entail that it According to Identity Theory, the reason that we tend is false. not to say that Brutus killed Caesar before Caesar dies is that the description of Brutus's action as "a killing" remains epistemically unavailable to us until Caesar dies. We are not justified in calling it a killing until we are sure that it will result in the death, but given that it does result in a death the action is a killing all along.

Now, the epistemic response explains why we tend not to call a killing "a killing" until we know that it will result in a death. However, there are other odd cases that it does not seem to explain. Consider the following oddity that does not admit of any epistemic uncertainty:

Brutus killed Caesar in January. Caesar went south to recover and in March he died as a result of the wounds that Brutus had inflicted on him three month earlier.

Granted, it does sound very odd to put it like this. But oddity does not entail falsity. In fact, there are other ordinary language examples that would lead us to conclude that it makes

perfect sense to say that Brutus killed Caesar before Caesar died. Consider, for example, the following scenario:

Brutus stabbed Caesar. Caesar was losing blood very rapidly and we rushed him to the hospital. When we arrived, the doctor told us that there was nothing that she could do to save Caesar. His death was inevitable. At that point we knew that Brutus had killed Caesar.

We would find it exceedingly odd to say that at that point we knew that Brutus was killing Caesar or that Brutus was in the process of killing Caesar or that Brutus's killing of Caesar was taking place. As these two cases demonstrate, conclusions drawn from oddities in ordinary language can often conflict with one another. As a result, it does not seem as if we should reject metaphysical claims purely on the basis of oddities in ordinary language that might result from them.

Identity Theory finds independent support for the claim that action reported in timespanning "by"-sentences are co-temporal from the No Further Effort argument. Recall from our earlier discussion of Parthood Theory that the No Further Effort argument says that acts are over at the time when the agent is no longer doing anything relevant to the action. Since once Brutus has stabbed Caesar he is no longer doing anything relevant to the killing of Caesar, it must be the case if No Further Effort is correct that once the stabbing is over, the killing is over. The intuition behind this argument is that once there is no further effort required from the agent to complete the action, the action is complete. As we have already seen, this is a very strong intuition and figures prominently in the very counterexample

intended to show No Further Effort to be false.

This argument lends support to the epistemic response because it gives us a reason to believe that Brutus's action of killing Caesar could have been over before Caesar's death, despite the apparent oddity in making such a claim. Moreover, accepting the No Further Effort argument gives us a plausible reason to avoid the claim that a person's actions can continue after his or her death. Accepting the temporal challenge means accepting that even if Brutus had died before Caesar did, Brutus's action of killing Caesar continued after Brutus was dead. Advocates of the No Further Effort argument believe that an agent's deeds must be done by the time the agent is dead. Since the agent might die immediately after, for example, plunging in the knife, the action must be complete when the agent no longer needs to make any effort. Whereas the temporal challenge requires that Brutus's killing Caesar lasts up to and includes Caesar's death, according to Identity Theory Caesar's death is in no way a part of the killing. If Identity Theory is correct, the killing is completely distinct from the death.

4.3 Identity Theory Responds to Second Challenge

Recall that the causal objection against Identity Theory says that if, for example, Brutus's stabbing Caesar causes Caesar's death, and the stabbing is thus identical with the killing, then Brutus's killing Caesar causes the death. It sounds very odd to say that killings cause deaths, and Identity

Theory must account for this oddity.

Identity Theory accepts the result that killings must be considered as causes of deaths. Again, although it may sometimes sound odd to say that killings cause deaths, the mere oddity does not serve as solid grounds for a serious objection. Though it might be odd in many circumstances to say that the killing caused the death, it need not be considered false. As we did in addressing the temporal objection, we can point to cases in which it does not seem odd to suggest that the killing caused the death. Consider an instance in which, upon arriving at the scene of a death, one asks, "How did he die?" and receives the response, "He was killed." It does seem as if, in a case like this, what is being asked for is a cause of death. Now to say that he was killed is to give, albeit in very general terms, a cause of death. It is true that it is not as informative a response as could be given. For example, "He was stabbed" might be more informative. But "He was killed" certainly rules out many natural ways in which his death might have been caused, and would not strike one as an incredibly odd response to the question, "How did he die?"

4.4 Identity Theory Responds to Third Challenge

The third objection to Identity Theory is that it fails to give a coherent account of necessary properties of action. It is often the case, particularly in timespanning when someone does alpha by doing beta, that the actions in question seem to have

different necessary properties. If it is a necessary property of a killing that it results in a death, yet it is not a necessary property of a stabbing that it results in a death, then they cannot be identical. For surely identity requires that all the properties be the same.

Identity Theory construes actions in a way that allows it to explain and avoid this apparent objection. Given the way that Identity Theory construes actions, it is de dicto necessary, but not de re necessary that a killing result in a death. The relevant way in which a killing necessarily results in a death does not make trouble for the identity claim with respect to a particular killing and a particular stabbing. We can make sense of this claim by considering the following parallel to the necessary properties objection from Quine:

Mathematicians may conceivably be said to be necessarily rational and not necessarily two-legged; and cyclists two-legged and not necessarily rational. But what of an individual who counts among his eccentricities both mathematics and cycling? (Word and Object, p. 199)⁵

We do not read this passage and conclude from it that one and the same individual cannot be both a mathematician and a cyclist.

Clearly, an individual can be both. Identity Theory invites us to treat the properties of actions and events in the same way that we treat the rationality and two-leggedness of the mathematician-cyclist. We need not conclude from the claim that killings necessarily result in deaths and that stabbings do not

Richard Cartwright brought this passage from Quine to my attention.

necessarily result in deaths that one and the same action cannot be both a killing and a stabbing.

The difficulty arises only if we fail to consider the difference between de dicto and de re modal statements. The de dicto reading of the claim that killings necessarily result in deaths and stabbings do not necessarily result in deaths says:

- \square (x)(killingx \rightarrow x results in death)
- ~□ (x)(stabbingx→x results in death).
 The de re reading says:
 - (x) (killingx $\rightarrow 0$ x results in death)
 - (x)(stabbingx $\rightarrow \sim 0x$ results in death)

Identity Theory denies the de re reading, affirming only the de dicto reading of necessary properties. If our variables range over actions, the first pair of statements tells us that for any action, in so far as the action happens to be a killing, it results in a death. For example, Brutus's action might happen to be a killing. If it is, it happens to result in a death. But, according to this reading, no action is essentially a killing, thus no action essentially results in a death. It is the de re reading that requires that if an action is a killing then it is essentially a killing and so essentially results in a death. If we deny the de re reading of necessity, it is possible that the very entity, i.e. Brutus's action, that we call "a killing" in this world may not have resulted in Caesar's death in some other

possible world. If Identity Theory is true, there is no need to respond differently in the case of actions from the way we would respond in the case of individuals. Just as our mathematician-cyclist is neither essentially a mathematician nor essentially a cyclist, so our killing-stabbing is neither essentially a killing nor essentially a stabbing. Just as it is a contingent fact that this individual is a mathematician and a contingent fact that he is a cyclist, so it is a contingent fact that Brutus's action is a stabbing and a contingent fact that it is a killing. And so, as the man happens to be rational (qua mathematician), and happens to be two-legged (qua cyclist), Brutus's action happens to result in a death (qua killing) and happens to involve a sharp object (qua stabbing).

4.5 Assessing Identity Theory's Response to Third Challenge

The analogy with the cyclist-mathematician certainly makes it seem more plausible than it initially did that the same action could be both a killing and a stabbing. However, there may be a larger issue at work which would render the analogy illegitimate. That is, someone might want to claim not just that killings,

This argument depends of course on using the language of the actual world to refer to entities in other possible worlds. The argument is not claiming that the entity that we refer to here as "Brutus's killing of Caesar" would be referred to as "Brutus's killing of Caesar" at some other possible world where Brutus's action did not result in Caesar's death. The claim is only that insofar as we can pick out an entity in some other possible world that corresponds to or is (depending on whether you accept counterpart theory or transworld identity) the action that we describe as "Brutus's killing of Caesar," we can describe it as such even at worlds where Caesar fails to die.

insofar as they are killings, necessarily result in deaths, but rather that a killing is essentially a killing. If a killing is essentially a killing then that sets it apart from the mathematician-cyclist who is neither essentially a mathematician nor essentially a cyclist. It is quite conceivable that he could be neither, and so neither necessarily rational nor two-legged. Both Parthood Theory and Trope Theory subscribe to the view that if some action alpha is a killing, then it could not have been anything else. Had things gone otherwise and not resulted in a death, then that action alpha would simply not have occurred. Some other action beta would have occurred instead. The action could have been different in other unessential ways, but if it is a killing then it cannot fail to be a killing in some other possible world.

The dispute comes down to the question of how essentialist we need to be about actions. Identity Theory leaves room for many different ways in which a given action may vary. This may run it into some trouble because where Identity Theory can make a plausible case for Brutus's action not being a killing, it may not be able to make a similar claim for every description of Brutus's action. For example, we said that it was contingently a stabbing. Is it plausible to say that the action that was a stabbing could have been a shooting instead? Or does the bodily movement have to have been the same, making it the essential feature of the action? On the face of it, the bodily movement

Richard Cartwright put this question to me.

looks as if it would be a natural place for Identity Theory to pin down what is essential to an action. However, it is also natural to think that Brutus's stabbing could have been faster or slower than it was. If its speed or angle were different, the bodily movements would have been different. Hence, bodily movement does not seem to be what is essential to the action. In its leniency about the essential features of actions, Identity Theory leaves many questions unanswered. The other two theories are less lenient in this respect, making stronger claims about what is essential to an action.

5 Conclusions, So Far

So far, although we have no definitive argument in its favour, Identity Theory appears to be give us the most plausible response to the question:

If an agent X does alpha by doing beta, what is the relationship between X's doing alpha and X's doing beta?

It is a simple and elegant theory, giving us a straightforward answer. In every case, both synchronous and timespanning, Identity Theory says that the relationship is identity. In addition, although Identity Theory results in some oddities, it has ways of explaining them that make them seem less odd and its simplicity compensates. The other two theories are much more ambitious in their goals, and as a result, far more complex. As we have seen, in an attempt to call the No Further Effort argument into doubt, Parthood Theory needs to appeal to the very

intuitions that make No Further Effort plausible in the first place. Trope Theory gives us a way of drawing a distinction in timespanning cases without having to give up No Further Effort. However, its scope creates other difficulties. In addition, we have no convincing reason for thinking that tropes form a distinct ontological category. Therefore, partly because the other two theories face more serious difficulties and partly because of its own attractive features, Identity Theory is the best theory we have for answering the question with which we began.

PART TWO: HOW TO ANALYZE "BY"

1 "By" as the relation of identity on events

Identity Theory gives a plausible answer to the question we posed at the outset. However, although the theory suggests that "by" is indicative of an identity between events, it has not provided us with an analysis of "by." What we need to consider is what kind of relationship, if any, "by" expresses. It may be indicative of an identity, but can Identity Theory analyze "by" as the relationship of identity between events?

Consider again the case in which Brutus killed Caesar by stabbing him. Can Identity Theory analyze

- (3) Brutus killed Caesar by stabbing him
- as
- (4) (3x)(3y)[killing, (x, Caesar, Brutus) & stabbing

(y, Caesar, Brutus) & By(x, y). In this analysis, "by" is a relation on events.

The difficulty that arises when we analyze "by" straightforwardly as the relation of identity on events is that "by", unlike identity, is not a symmetrical relationship. If we recognize the relation "By(x, y)" and if "x = y" is true, then we also have to allow that "By(y, x)" is true. Alvin Goldman claims, "if an agent S does act A' 'by' doing act A, then he does not do A 'by' doing A'" (Goldman, p. 5). For example, if Brutus kills Caesar by stabbing him, it is not the case that Brutus stabs Caesar by killing him.

Now, the objection is not as straightforward as Goldman's construal of it suggests. For given (4), which must be what Goldman has in mind as the way Identity Theory would render (3), there are no singular terms referring to events. The sentence, as it stands, is an existential generalization that can be made true by any number of events. However, if we make the right assumptions and instantiate singular terms using existential instantiation, the objection does go through.

What this suggests is that either "by" is not a relationship between events, or that Identity Theory is false. If we conclude on the basis of the asymmetry of "by" that Identity Theory is false, then we are no just going to create trouble for Identity Theory. We are also going to create trouble for Parthood Theory. Recall that it is only in timespanning cases that Parthood Theory

See "Appendix A" for the details of the proof.

says there is a difference between the events. In synchronous cases, Parthood Theory, like Identity Theory, asserts an identity. But it is not only in timespanning cases that "by" expresses an asymmetric relation (if it expresses a relation at all). To conclude on the basis of the truth of "he won the race by crossing the finish line first" that "he crossed the finish line first by winning the race" is true is just as unwarranted as it would be to conclude on the basis of the truth of "Brutus killed Caesar by stabbing him" that "Brutus stabbed Caesar by killing him" is true. Given that neither Identity Theory nor Parthood Theory is plausible if we consider "by" to express a relation on events, let us consider how we might analyze "by" as expressing something else.

2 "By" as syncategorematic: Forming complex verbs out of simpler ones

Hornsby, also an advocate of Identity Theory, suggests that rather than relating events to one another, "by" forms complex verbs out of simpler ones. Consider Anscombe's case of the man who poisoned the inhabitants by operating the pump. Rather than understanding "by" as expressing a relation between the actions his operating the pump and his poisoning the inhabitants, Hornsby claims that it creates the complex verb "to poison the inhabitants by operating the pump" from the simpler verb "to poison the inhabitants." Likewise for the synchronous case that we just considered. "By" does not relate his winning the race and his crossing the finish line first. Instead, from the verb

"to win the race" it creates a more complex verb, namely "to win the race by crossing the finish line first." On Hornsby's account of "by", the action that we describe as "his poisoning the inhabitants" and the action that we describe as "his operating the pump" is really all along his poisoning-the-inhabitants-by-operating-the-pump. Similarly, the action we describe as "his winning the race" and the action we describe as "his crossing the finish line first" is really all along his winning-the-race-by-crossing-the-finish-line-first.

If we opt for Hornsby's account of "by", we have no difficulty accounting for its asymmetrical character. This account does not explain, however, why we seem to be able to conclude, from the truth of "he poisoned the inhabitants by operating the pump," that there was an event that was his poisoning the inhabitants and there was an event that was his operating the pump. Consider

- (5) Jones poisoned the inhabitants by operating the pump.
 if (5) is true, the most that we seem justified in inferring from it on the basis of the account of "by" under consideration is
 - (6) (∃x) [poisoning-the-inhabitants-by-operating-thepump (x, Jones)].

However, we want to be able to infer from (5) that

(7) (∃x)(∃y) [poisoning-the-inhabitants (x, Jones) & operating-the-pump (y, Jones)].

The complex predicate account appears to leave "by" essentially unanalyzed.

3 "By" as expressing the part-whole relation on events

As we have already seen, Parthood Theory cannot analyze "by" as a relation on events because of the identity it claims obtains in synchronous cases. This is unfortunate because it might make sense, especially in timespanning cases such as Brutus's killing Caesar by stabbing him, for Parthood Theory to analyze "by" as a relation between events and their parts. Thus, if Brutus killed Caesar by stabbing him, and the stabbing is part of the killing, then Parthood Theory could say that there was a killing of Caesar by Brutus and the stabbing was a part of the killing. This analyzes sentences of the form

- (2) X does alpha BY doing beta
 as (where "<" means "is a part of")</pre>
- (8) (3x)(3y) [doing (x, X) & doing (y, X) & (y < x)]. Besides having the virtue of being straightforward, this analysis allows us to pick out nested events, that is, events that are included in larger events in the way that Parthood Theory says that the killing includes the stabbing. Giving an account of parts of events is something that any theory should be able to do. For even if we disagree with Parthood Theory about the relationship between the killing and the stabbing, we might still want to say that certain events contain other events as their parts. For example, the performance of one concerto has as its

Judith Thomson drew attention to this possibility and its shortcomings in her metaphysics seminar (MIT, spring 1990).

parts the playing of various instruments. These events are nested in the performance of the concerto.

As I said, however, the part-whole relation does not adequately capture "by". First, it is a relation on events and hence does not deal adequately with synchronous cases because of the asymmetrical character of "by." Second, if we take (8) to be the formalization of an English "by"-sentence, then it must translate back into one as well. But surely it is not the case that any events that are related as part to whole can be appropriately placed in a "by"-sentence. For example, we might fill in y with "Brutus's moving his arm one inch," since this is surely a part of Brutus's killing Caesar. But we do not want to say that Brutus killed Caesar by moving his arm one inch. The fact that not every part of an event is "by-related" to the whole indicates that the part-whole relationship is simply not strong enough to stand as an analysis of "by."

4 "By" as syncategorematic: Causal Verbs

We have seen Hornsby's account in which "by", rather than expressing a real relation on events, creates complex verbs out of simpler ones. Thomson proposes another way of analyzing "by" in which it is syncategorematic. On this account, "by" gets analyzed out of the action sentences in which it occurs.

Thomson points to a certain class of verbs - causal verbs - which (i) take "by"-clauses and (ii) accept events as subjects.

For any sentence of the form

we can analyze "by" out as long as the first verb is a causal verb. For example, "kill" is a causal verb because it takes "by"-clauses and accepts events as subjects, e.g. we can say "Brutus killed Caesar" and "the stabbing killed Caesar." This gives us an analysis for causal verbs.

If we analyze "by" in terms of causal verbs, then what we are saying in a sentence of the form "X did alpha BY doing beta," where the first verb is a causal verb, is that both the agent X and X's doing beta did alpha, but that X's doing beta "did the work." For example, in Brutus's case, he would not have killed Caesar if his stabbing Caesar had not resulted in Caesar's death. According to this analysis,

- (1) Brutus killed Caesar by stabbing him should be read as saying
 - (9) (3x)(3y)[kills (x, Caesar, Brutus) &

stabs (y, Caesar, Brutus) & (1 ls (x, Caesar, y)]
Brutus killed Caesar and Brutus's stabbing Caesar killed Caesar,
but the stabbing did the work. According to Parthood Theory, as
we have already seen, Brutus's killing Caesar is not the same
event as Brutus's stabbing Caesar. What makes (9) true is not
that the stabbing is a killing, but that the stabbing is a
killer. In other words, (9) is true because the stabbing brought
about Caesar's death.

This analysis is attractive because it accounts for timespanning and synchronous cases of "by" involving causal

verbs. However, in restricting itself to causal verbs, the analysis is only partial. Thomson recognizes that since there are "by"-sentences in which the first verb is not a causal verb, there are cases that this analysis does not cover. There are two kinds of cases involving non-causal verbs that need attention. Both sorts of cases involve verbs that are non-causal because they fail to satisfy the second criterion for causal verbs. They do not take events as subjects.

The first range of cases to which Thomson draws attention involves action verbs that do not take events as subjects because they seem to involve intentionality. If we accept the preceding analysis of "by" in terms of causal verbs then we must accept that

- (10) Brutus murdered Caesar by stabbing him is intrinsically different from
- (1) Brutus killed Caesar by stabbing him.

 It is not entirely clear why (10) should invite an analysis entirely different from that of (1). Except for the fact that Brutus's action is being referred to differently, (10) and (1) look essentially the same. While it is true that murdering suggests more than killing does, and involves a presupposition that the action was intentional, it is also true that in this case, given that the killing was also a murdering, it was intentional. The same action makes both (1) and (10) true. They appear to have the same logical form. Why then should we suppose that they ought to be analyzed differently from one another?

Indeed, given the analysis, it is difficult to see how to make sense of (10) at all. For recall that in the original case involving the causal verb "to kill", both Brutus and Brutus's stabbing Caesar killed Caesar, but the killing did the work. That is, Brutus killed Caesar only insofar as his stabbing Caesar killed Caesar. This enables us to avoid the implausible conclusion that there are two killing events - one by Brutus and one by Brutus's stabbing Caesar. To infer this about the case seems unjustified, given that there can only be one killing; Caesar can only die one death. In making this claim, however, we are buying into the claim that Brutus does not literally "do the work", his stabbing does. Now if this is the case, what "does the work" when Brutus murders Caesar by stabbing him? Clearly in this case we cannot claim that Brutus murdered Caesar only insofar as his stabbing Caesar murdered Caesar. A stabbing is an event and not the sort of thing capable of murder. however, that we may say of the same case, both that "Brutus killed Caesar by stabbing him" is true and that "Brutus murdered Caesar by stabbing him" is true. Given the present analysis of "by" we are warranted on the basis of the first "by"-sentence in drawing the conclusion that the stabbing is what "did the work", not Brutus. In the second "by"-sentence, however, not only are we unwarranted in making that claim, but we cannot claim that the stabbing "did the work" because stabbings are not the sorts of things that can do the work in murderings. Since "to murder" does not take events as subjects, Brutus must have done all the

work himself. But it certainly makes sense to think that Brutus brings about certain events only insofar as his actions do. And if this is the case, then it should not make a metaphysical difference whether we describe Brutus's action in terms of its being a killing or in terms of its being a murdering. According to the analysis of "by" in terms of causal verbs, however, it does make a metaphysical difference. This putative difference creates an implausibility for the analysis because, given that the same event can be both a killing and a murdering, it is unclear why we should claim that there is a metaphysical difference at all. It is furthermore unclear why we should think this difference, if it exists, needs to be brought out in the analysis of "by."

The second range of cases involves non-causal action verbs in synchronous "by"-sentences. For example, consider the verb "to walk." For any case of an agent X's walking, we may say that "X walked by putting one foot in front of the other several times in a row" is true. We may do the same for "to kick," "to eat," "to dance," and so on. For any of these non-causal action verbs, there is a sentence involving a synchronous case of "by" that describes exactly what the agent did when she walked, ate, danced, or kicked. Since this analysis addresses only causal verbs, there is a whole range of cases like this that it leaves unanalyzed.

The drawback of the analysis of "by" in terms of causal verbs is that it leaves a whole range of cases unanalyzed because

they involve non-causal verbs. Thomson recognizes this feature of the analysis but does not consider it to be a drawback. She thinks that a separate analysis is required for non-causal verbs because they involve intentionality. It is not entirely clear to me, however, why the fact that some action verbs take events as subjects and some do not should be brought out as a difference reflected in the analysis of "by."

5 "By" as a relation on facts

Bennett does not believe that an analysis of "by" tells us much about the relation between events, since "by" cannot be construed as a relation on events and actions. Nor does he claim that "by" is syncategorematic. Rather, Bennett provides an explicit analysis of "by" as a relation on facts.

Although we saw some difficulties in getting a grip on tropes, since Bennett's account is about facts, it might give us the alternative that we are looking for. Recall that according to Trope Theory an event is a trope, viz., the instantiation of a complex property P* at a zone. Whenever we name an event, we refer to the entire complex property being instantiated, not just the part of the property explicitly picked out by the name. We capture this attractive feature of the proposal with Identity Theory as well. That is, Identity Theory is also lenient with respect to the way the description or name of an event specifies the event that it does. Following an example of Bennett's, "Leibniz's journey at T" names an instance of a property

instantiated at T of which journeying is a proper part. But the event includes other properties as well, e.g. being on a train, being a travelling at a certain speed, being a going between certain cities, and so on. Every event has a unique corresponding fact which is "the fact that complex property P* is instantiated at a certain zone" (Bennett, p. 128). This fact is called the companion fact of the event. It will become clear that we do not need to construe events in terms of tropes in order to make sense of this analysis of "by."

Bennett analyzes "by" as a relation on facts. In any statement of the form "X does alpha by doing beta" the first clause of the sentence "entails something to the effect that the subject of the clause - the 'agent' in the 'by'-statement - was instrumental in its coming to be the case that P_a, where P_a is a fact or state of affairs systematically associated with [the agent's doing alpha]" (Bennett, p. 214). For example, if Brutus killed Caesar by stabbing him, then Brutus was instrumental in bringing about the fact that Caesar died. An agent is instrumental in its becoming the case that P_a if something about the agent's conduct has P_a as a consequence. The second clause of the sentence tells us what it is about the agent's conduct that contributes to its becoming the case that P_a. With respect to Brutus, it is the fact that he stabbed Caesar that makes

P., the fact that Caesar dies, is what Bennett calls "the end fact." Of course, Brutus was instrumental in bringing about

many facts and states of affairs by stabbing Caesar, so how do we know which one of these is supposed to be the end fact Pa?

Bennett outlines his notion of an end fact:

To the extent that we understand a sentence "[Noun phrase][verb phrase]," we know which sentences it entails of the form "[Noun phrase] is instrumental in its becoming the case that [S]." Of all the values of S that make this true, the one that has a place in [the] analysis is the strongest noninstrumental S for which the entailment holds. The fact expressed by this sentence is what I shall call "the end fact" (Bennett, p. 216).

Consider, for example, the sentence "Alex boiled the water." If it is true, it entails all of the following facts:

Something happened to the water
The water's physical state altered
The water boiled
Someone made the water boil
Alex made the water boil

of these facts entailed by "Alex boiled the water" only the third one is the end fact. The first two are weaker than is required of the end fact; they do not give us enough information. The last two are instrumental facts, that is "they speak of a state of affairs not merely as obtaining but as being brought about" (Bennett, p. 216). The end fact must be a noninstrumental entailment, not an instrumental one. Bennett gives several other examples of sentences and their end facts: "He felled the tree / the tree falls; He nets the fish / the fish comes to be in a net; He reduces the price of silver / silver becomes cheaper" (Bennett, p. 216). Bennett admits that "that is the best [he]

can do to explain [his] concept of the end fact."10

In addition to the limits on what can stand in as the end fact, there are certain constraints "K" on the role that the agent's conduct plays in its becoming the case that P. The analysis must include these constraints since different facts about an agent's conduct might have the same end fact. For example, "Brutus killed Caesar" and "Brutus brought it about that Caesar died" both have the end fact, "Caesar died." However, the sentences impose different constraints K on Brutus's conduct in bringing it about that Caesar died. In the first case, Brutus must have done the killing himself. In the second case, it might be that Brutus paid Marcus to do the deed.

With these details in mind, let us look at Bennett's analysis of "by".

Where "X does alpha..." means

Some fact about X's conduct has a K role in its becoming the case that P....

The addition "...by doing beta" means

... namely the fact that X does beta.

Thus, "X does alpha by doing beta" means

The fact that X does beta has a K role in its becoming the case that P. (Bennett, p. 217)

This analysis construes "by" as meaning "...namely the fact

requires some attention. I do not want to get into it here, however, because I think that Bennett's example give us a clear enough general idea of what the end fact of a sentence is supposed to be. Note that end facts are not limited to "by" sentences; every action or event sentence has one.

that..." and acting as an operator that turns the fact name, "X's doing beta" into its parent sentence, X did beta. In order to use "by" in this way, the initial clause of the "by"-sentence, for example, "Brutus killed Caesar," has to be reconstrued as meaning that something about Brutus's conduct had a K role in its becoming the case that Caesar died. Now the analysis of "by" can work. "By stabbing him" becomes "namely, the fact that Brutus stabbed him." A complete analysis of "Brutus killed Caesar by stabbing him" yields the following:

Some fact about Brutus's conduct had a K role in its becoming the case that Caesar died, namely the fact that Brutus stabbed him.

Bennett provides further details in his presentation of the analysis. In particular, he spells out the K constraints on the role of the agent's conduct in bringing about the end fact. I shall briefly mention them before evaluating Bennett's proposal. Bennett makes two main points. First, the behavioral fact about the agent that brings about the end fact P_a must be seen as the sole input into P_a's obtaining, e.g. the fact that Brutus stabbed Caesar must be considered under the circumstances to be the fact that alone contributed to Caesar's death. We do not want to allow that certain intervening coincidences brought about the end fact. For example, if as Brutus stabbed Caesar, all of a sudden the earth began to shake, the building started to crumble, and a huge column toppled over, struck Caesar on the head, and killed him, then the fact that Brutus stabbed Caesar does not bring it

about that P_a. Brutus does not kill Caesar at all in this circumstance, the column does. Second, for any instrumental verb to apply correctly in a given circumstance, it must be the case that the causal chain resulting in the end fact P_a "runs wholly through the will of the agent." For example, as we said earlier, "Brutus killed Caesar" and "Brutus brought it about that Caesar died" both have the end fact, "that Caesar died." However, in the second case it might be that Brutus hired Marcus to kill Caesar. In that case, Marcus killed Caesar; Brutus did not. Bennett raises this point to illustrate that our intuitive use of instrumental verbs, and thereby, the K role that an agent's conduct can play in bringing about a certain end fact, is constrained by the idea that there should be no intervening agency involved.

This analysis of "by" is the most complicated one that we have considered so far. Its main virtue is that it provides a uniform analysis of "by" in every case. Unlike the previous account, we need not worry about the distinction between causal and non-causal verbs. Nor do we need to concern ourselves with the distinction between synchronous and timespanning cases; it treats both cases in exactly the same way. Since the analysis is not in terms of events, but in terms of facts instead, this account of "by" is compatible with any theory of actions and events that we opt for.

There remain some questions that we need to consider. It is especially important to look at the functions that the end fact

P. and the constraints K on the role of the agent's conduct serves in the analysis.

First, let us examine the function of the end fact in this account of "by." The end fact of "Brutus killed Caesar by stabbing him" is the fact that Caesar died. That Brutus killed Caesar is supposed to indicate that something about Brutus's conduct played a certain role in bringing about the fact that Caesar died. The fact that Brutus stabbed Caesar is supposed to explain what it was that Brutus did to bring about the fact that Caesar died. Given the complexity of the analysis, is the addition of the end fact really necessary? What we need is an analysis of "by" and yet before we can get that we are required to analyze the verb in the first clause of the "by"-sentence. In this example, we need to analyze "kill" in terms of dying before we can begin to analyze "by". But, would it suffice to say:

Some fact about Brutus's conduct played a K-role in its becoming the case that Brutus killed Caesar, namely the fact that Brutus stabbed Caesar.

If the analysis did not lose anything in the process, we would be able to avoid altogether the business of deciding what the end fact is by taking this route. The claim that Brutus killed Caesar by stabbing him seems to give us, in a remarkably straightforward fashion, the fact that Brutus killed Caesar. Does this analysis allow "by" to function as a relation between the fact that Brutus killed Caesar and the fact that Brutus stabbed Caesar in just the way that Bennett wants it to? We shall come back to answer this question in a moment.

The second complication in the analysis is the K constraints on the role of the agent's conduct in bringing about the end fact. The main reason for imposing these constraints on the role of the agent's conduct in bringing about the end fact is that different facts about the agent's conduct might have the same end fact. If we dispense with the end fact, we might be able to dispense with these vague K constraints on the agent's conduct as well. In explaining the need for the K constraints, we saw that "Brutus killed Caesar" and "Brutus brought it about that Caesar died" both have as their end fact the fact that Caesar died. They impose different constraints, however, on the role of Brutus's conduct in bringing about this result. If we dispense with the end fact from the analysis, we might analyze "Brutus killed Caesar by stabbing him" as

Some fact about Brutus's conduct has a role in its becoming the case that Brutus killed Caesar, namely the fact that Brutus stabbed Caesar.

We could try analyzing the case in which Brutus brought about Caesar's death by paying Marcus to do the deed, as

Some fact about Brutus's conduct has a role in its becoming the case that Brutus brought about Caesar's death, namely, the fact that Brutus paid Marcus to stab Caesar.

Because the difference between "Brutus killed Caesar" and "Brutus brought it about that Caesar died" shows up in the analysis if we leave out the end fact, perhaps we do not need to specify it in the K-role either.

These two modifications to the analysis of "by" simplify it a great deal and make it somewhat more appealing as a result,

that is, if they work. However, neither of them is going to work. The end fact has to be a non-instrumental fact, such as the fact that Caesar died, rather than an instrumental fact, such as the fact that Brutus killed Caesar or that Brutus brought about Caesar's death. The reason for this requirement is that the end fact needs to be a consequence of the agent's conduct and so cannot include the agent's conduct. For example, the fact that Caesar died is a consequence of Brutus's stabbing Caesar whereas the fact that Brutus killed Caesar is not, on any of the views we have considered, a consequence of Brutus's stabbing Caesar. Nor is it possible for us to do away with the Kconstraints altogether. If we cannot do away with the end fact, we need the K constraints to make discriminations and impose some limitations where different facts about the agent's behaviour bring about the same end fact. It cannot be the case that the fact about the agent's conduct plays just any role in bringing about the important result. For example, the relevant fact about Brutus's conduct is not the fact that he borrowed a knife, nor is it the fact that he took LSD, it is the fact that he stabbed Caesar. For Brutus, did not kill Caesar by borrowing a knife, nor did he kill Caesar by taking LSD, he killed Caesar by stabbing him. We need to impose some constraints on the fact about Brutus's conduct that played a role in bringing about Caesar's death.

The drawback of the analysis is that, although it leans very heavily on the K constraints, it offers only sketchy details

about what they might involve. Until we know, despite the complexity of the account, we do not have a very informative picture of what the role of "by" is. In order to make this account of "by" workable, it needs to be much more detailed than it is.

CONCLUSION

None of the various analyses of "by" brings us much closer to deciding which of the three theories considered in Part One gives us the best answer to the initial question. The analysis of "by" as a relation on events will not work for any of the theories. All three theories claim that in synchronous cases in which X does alpha by doing beta, X's doing alpha is X's doing beta. In these cases, interpreting "by" as a relation on events fails to capture its asymmetry. Two of the proposals suggested that "by" is syncategorematic. The first maintained that "by" forms complex verbs out of simpler ones. This proposal, however, leaves "by" essentially unanalyzed. The second provided a comprehensive account of "by" when it occurs in sentences whose first verb is a causal verb. The limited range of cases to which this account applies is a drawback, especially since it is not clear why the difference between causal and non-causal verbs should come out as a difference in interpreting "by." Finally, we saw an analysis of "by" as a relation on facts. Simplifying Bennett's account does not help and we are still left with unspecified K-constraints on an agent's behaviour. Unless we can get a clearer picture of what there are, it is not clear what role "by" plays. In addition, analyzing "by" as a relation on facts takes us far afield from the question with which we began at the outset. For this analysis says nothing about the relationship between the agent's doing alpha and the agent's doing beta.

Given that each of these ways of analyzing "by" has some difficulties, and that aside from analyzing "by" as a relation on events, none of them favours any one of the three theories we have been considering, it does not look as if analyzing "by" is going to get us any closer to understanding the relationship between X's doing alpha and X's doing beta when X does alpha by doing beta. Based on the considerations in Part One about the strengths and weaknesses of the theories, I favour Identity Theory over Parthood Theory and Trope Theory. Its simplicity is not compromised by the fact that it lacks the ambition of the other two theories, although it is partly explained by it. The fact that Trope Theory and Parthood Theory attempt to do much more than Identity Theory leads them into complexities that ultimately create difficulties for them.

Appendix A

Proof for the claim that if Identity Theory is true, "by" cannot be a relation on events. If it were, we would be warranted in inferring the undesirable conclusion that if X does alpha by doing beta, then conversely, X does beta by doing alpha.

Formalize

(1) Brutus killed Caesar by stabbing him

as

(2) (3x)(3y)[killing (x, Caesar, Brutus) & stabbing (y, Caesar, Brutus) & By (x, y)].

Given the assumption that Brutus has killed only once, viz. when he killed Caesar, and Brutus has stabbed only once, viz. when he stabbed Caesar, and Identity Theory's claim that

- (3) Brutus's killing Caesar = Brutus's stabbing Caesar
 Using existential instantiation on (2), we get
 - (4) killing (alpha, Caesar, Brutus) & stabbing (beta, Caesar, Brutus) & By (alpha, beta).

Making the proper substitutions based on the identity in (3), from (4) we can get

(5) stabbing (beta, Caesar, Brutus) & killing (alpha, Caesar, Brutus) & By (beta, alpha)

which we can generalize to give us the unwanted result

(6) (3x)(3y)[stabbing (x, Caesar, Brutus) & killing (y, Caease, Brutus) & By (y,x)].

BIBLIOGRAPHY

- ANSCOMBE, ELIZABETH. Intention. Oxford: Basil Blackwell, 1958.
- BENNETT, JONATHAN. Events and Their Names. Cambridge: Hackett Publishing Company, 1988.
- DAVIDSON, DONALD. "Agency" in Essays on Actions and Events.
 Oxford: Clarendon Press, 1980. pp. 43-61.
 ----"The Individuation of Events" in Essays on
 Actions and Events. pp. 163-87.
- GOLDMAN, ALVIN. A Theory of Human Action. Princeton: Princeton University Press, 1970.
- HORNSBY, JENNIFER. Actions. London: Routledge and Kegan Paul, 1980.
- QUINE, WILLARD VAN ORMAN. Word and Object. Cambridge: The MIT Press, 1960).
- THOMSON, JUDITH JARVIS. Acts and Other Events. Ithaca: Cornell University Press, 1977.

ESSAY TWO

THE ONTOLOGY OF ACTION AND ETHICS

In what ways might action theory be thought to make a difference to moral theory? Perhaps our ontological claims put constraints on what we can or cannot say in ethics. If so, then it would be significant for the moral theorist to know what these constraints are. While the broad issue I am interested in is how the ontology of action affects ethics, I shall focus this discussion by addressing a very specific topic - the intrinsic value of action - and postpone until later the more general discussion. I shall argue that although the ontology of action does constrain ethical theory to some extent, it does not typically do so in any substantive way. The only ontological claim that would require a shift in the way we view many of our ethical claims is the claim that there are no actions at all. We have no convincing reason, however, to think that the extreme claim that there are no actions is true.

1 What is intrinsic value?

In an earlier paper, I discussed three theories of action:
Identity Theory, Parthood Theory, and Trope Theory. I was
interested in how each handled the relationship between an
agent's doing alpha and an agent's doing beta if the agent does
alpha by doing beta. For example, suppose that Brutus killed
Caesar by stabbing him. How is Brutus's killing Caesar related
to Brutus's stabbing Caesar? Identity Theory says that the
relation between the killing and the stabbing is identity.
Parthood Theory denies an identity, claiming that the stabbing is
a proper part of the larger event that is the killing. Trope

Theory claims that every event is a trope and that the killing and the stabbing are different because they are different tropes.

One way in which the ontology of action might be thought to conflict with a conclusion that we would want to accept in moral theory emerges as follows. If we accept Identity Theory as our theory of action² then we have accepted the first step in an argument whose conclusion is that human actions have no intrinsic value. I shall show that contrary to our intuitions, no ethical theory requires actions to have intrinsic value.

Intrinsic value is to be understood as the value a thing has independently of what it causes or what causes it. We can focus the issue about intrinsic value by concentrating on a recent treatment of it by Judith Thomson. She defines intrinsic value in the following way:

A thing has positive intrinsic value in amount D just in case it is good in amount D but not because of what it causes or what causes it; and a thing has negative intrinsic value in amount D just in case it is bad in amount D but not because of what it causes or because of what causes it (Thomson, The Realm of Rights, pp. 131-32).

Given this definition of intrinsic value, the conclusion of the argument is that no action has any value in itself; the only value any action has derives from its causes or its consequences.

¹ A trope is an instance of a complex property at a spatiotemporal zone. I discuss tropes and Jonathan Bennett's Trope Theory of action at length in the first paper of my thesis.

² I defend this theory of action in my first paper.

³ All subsequent references to Thomson will be from The Realm of Rights (Cambridge: Harvard University Press, 1990).



The Libraries Massachusetts Institute of Technology Cambridge, Massachusetts 02139

Institute Archives and Special Collections
Room 14N-118
(617) 253-5688

There is no text material missing here. Pages have been incorrectly numbered.

Now, it seems very plausible to say that human acts do have intrinsic value. Murderings, stabbings and torturings are just the sorts of things that we believe to be bad, not just because of what they cause, but in some way intrinsically bad - bad in themselves. Likewise, we think that telling the truth, completing a university degree, and saving a drowning child are good, not because of what they cause or bring about, but rather intrinsically good - good in themselves. Since we have these plausible intuitions that acts do have intrinsic value, we need to take very seriously an argument whose conclusion is that they do not have intrinsic value.

2 Acts do not have intrinsic value: an argument

The first premise in the argument is Identity Theory. This theory says,

FIRST PREMISE (Identity Theory) In any situation in which an agent does alpha by doing beta, her doing alpha is her doing beta.

One of the implications of this theory of action is that every action turns out to be identical with a bodily movement. For example, consider once again Brutus's stabbing Caesar. We know that Brutus kills Caesar by stabbing him. Therefore, if Identity Theory is true, the killing is the stabbing. Furthermore, Brutus

I leave open the question of how to characterize mental acts like "adding in your head" or "making a decision." If these are mental actions, I don't think we want to say that they are bodily movements, so we would have to characterize them in some other way if IT is true. If IT is true, then any action that is performed by performing a bodily movement is a bodily movement.

stabs Caesar by moving his arm in a certain way. (Of course, he has a knife in his hand; if he did not have a knife in his hand, his moving his arm would not be a stabbing.) Therefore, again according to Identity Theory, the stabbing is the arm moving. Given Identity Theory and the transitivity of identity, the killing is the arm moving. Similarly, for any action, Identity Theory tells us that it is a bodily movement. The way we describe the action depends upon how much information about the bodily movement we want to give. For example, in order to provide the information that the bodily movement caused a death, we might describe it as "a killing." In order to provide the information that it was done with a sharp object, we might describe it as "a stabbing."

The second premise in the argument to the conclusion that human actions have no intrinsic value is the No-Value Thesis about bodily movements. This thesis says,

SECOND PREMISE (No-Value Thesis) Bodily movements lack value. (Thomson, p. 133.)

The No-Value Thesis gets its motivation from viewing intentions in terms of the following theory of intention:

MOTIVATION FOR SECOND PREMISE (Causal Theory of Intention) What makes an act be intentional is its being caused by an intention. (Thomson, p. 133)

For example, Brutus's killing of Caesar is intentional if it is caused by an intention to kill Caesar (Thomson, p. 133). If the Causal Theory of intention is true, then intentions cause intentional actions, and whether or not an act is intentional is simply a function of whether it was caused by an intention. On

this view, the intention is not something that accompanies an intentional action. Nor is it something that is going on in the agent's mind while she is performing an intentional action. It is simply a part of the history of the intentional action. Now, one might want to claim that bodily movements that are intentional actions have intrinsic value in virtue of their being intentional. However, the intrinsic value of an action cannot come from the action's causes or its consequences, whatever those causes or consequences may be. Thus, if the Causal Theory of intention is true, a bodily movement's being caused by an intention cannot give it intrinsic value. This makes the No-Value Thesis seem very plausible.

The argument says that: (1) All actions are bodily movements; (2) Bodily movements lack intrinsic value; therefore all actions lack intrinsic value. Thus, there is a thesis in the ontology of action, namely, Identity Theory, that figures in an argument for a conclusion that is of interest to moral theory. The moral theorist can either accept the conclusion that actions have no intrinsic value or reject at least one of Identity Theory, the Causal Theory of Intention, or the No-Value Thesis.

Now it might be the case that Identity Theory is false, or that the Causal Theory of Intention is false, thereby undermining the No-Value Thesis. Rather than addressing the merits of accepting or rejecting each of these theses, however, it is more interesting to consider whether the conclusion that actions have no intrinsic value is as troublesome for moral theory as it might

look. If all three theses were true, would there be an impact on moral theory? We have already noted why it seems intuitively plausible to say that our actions do have intrinsic value. Is there any additional moral motivation for saying that they do and for thinking that, therefore, the claim that they do not have intrinsic value makes trouble for moral theory?

3 A Utilitarian argument for the intrinsic value of actions
Consider a case that Thomson calls Transplant:

Here is Bloggs, who is a transplant surgeon, an extraordinarily good one - he can transplant anything at all successfully. He has five patients who need parts and will soon die if they do not get them; two need one lung each, two need one kidney each, and one needs a heart. Here is a young man in excellent health; he has the right blood type and can be cut up to supply parts for the patients who need them. The surgeon asks the young man whether he would like to volunteer his parts, but the young man says "I deeply sympathize, but no." If the surgeon proceeds despite the young man's refusal, he saves five lives instead of one. (Thomson, p. 135.)

Now the question is: ought Bloggs to proceed to cut up the young man and use his parts to save the five despite the young man's refusal? On the face of it, I think we want to say that no, Bloggs ought not to proceed in this case, and that his doing so would be a blatant violation of the requirements of morality. I agree fully with Thomson's view that any moral theory that yields the conclusion that Bloggs ought to proceed in this case is "in dire need of revision" (Thomson, p. 135).

The focus of the discussion is going to be a particular issue, namely, intrinsic value. I am going concentrate to on its

There is a certain idea with a long histon moral theory - it has enjoyed quite wide acceptance - which we shall call, as Thomson does, the "Central Utilitarian Idea." It says that "one ought to do a thing, whatever it may be, if and only if more good (even just a little more good) would come of doing it than of not doing it" (Thomson, p. 124). The Central Utilitarian Idea has been popular for several reasons, one of which is that it seems plausible, on the face of it, that we ought always to maximize goodness.

As it stands, however, the Central Utilitarian Idea is vague and calls for interpretation. Under one interpretation - Consequentialist Act Utilitarianism - the Central Utilitarian Idea yields the conclusion that Bloggs ought to cut up the young man and use his parts to save the five. Before I state it, let me clarify some terms. Let the consequence-set of an agent X's doing alpha be the set that contains every event that is both discrete from X's doing alpha and occurs if X's doing alpha occurs. These are all of the consequences of X's doing alpha. Furthermore, let us suppose that the consequence-set of X's doing alpha maximizes value if and only if it contains more value than any of the consequence-sets of X's doing anything that is among X's available alternatives to doing alpha. X's doing alpha maximizes value because the value of X's doing alpha outweighs

⁵ This is an interpretation of consequences supplied by Judith Thomson. There are other ways of construing consequences, but there is no need to go into them here.

the value of X's doing anything that is an alternative to alpha. Consequentialist Act Utilitarianism says,

CONSEQUENTIALIST ACT UTILITARIANISM X ought to do alpha if and only if the consequence-set of X's doing alpha would "maximize value" (Thomson, p. 132).

For Consequentialist Act Utilitarianism, an action maximizes value if and only if its consequence-set maximizes value. It is important to notice that, according to this theory, the value, if any, of the action itself is irrelevant in deciding whether or not that action, rather than one of its alternatives, ought to be done. The only thing of importance is the value of the consequence-set.

why does Consequentialist Act Utilitarianism say that Bloggs ought to cut up the young man and use his parts in Transplant?

Let us suppose that each death in this case has the same amount of negative intrinsic value, in so far as it is a death. Let us suppose further that there is no question that Bloggs would save the five if and only if only he transplanted the parts. And finally, let us suppose that the young man's parts are the only ones available. Now, since Bloggs's cutting up the young man and using his parts to save the five would have just one death in its consequence-set, whereas any alternative would have five deaths in its consequence-set, Bloggs would maximize value by cutting up the young man and using his parts to save the five.

It is hardly controversial to claim that any moral theory that yields the conclusion that Bloggs ought to cut up the young man and use his parts to save the five is unacceptable. Our

moral intuitions tell us that Bloggs surely ought not cut up the young man and use his parts without the young man's consent, even if his not cutting up the young man and using his parts will result in the death of the five other patients. Consequentialist Act Utilitarianism, therefore, is a moral theory that yields false conclusions. The theory itself must, therefore, be false.

Is there any way of retaining the Central Utilitarian Idea, but interpreting it so that it does <u>not</u> yield the conclusion that Bloggs ought to cut up the young man and use his parts to save the five? Recall that we initially introduced Transplant as a case that might give us reason to think that actions have intrinsic value. Consequentialist Act Utilitarianism, in attending only to the value of the consequence-set of an agent's doing something, does not take into account the intrinsic value that an action might have. By supposing that actions do have intrinsic value and by taking that value into account when we are determining the best course of action, we may be able to claim that Bloggs ought not cut up the young man and use his parts to save the five because that complex action does not maximize value.

Non-Consequentialist Act Utilitarianism is a reading of the Central Utilitarian Idea that allows us to attend to more than just the consequence-set of an agent's doing something. It permits us to include the value of the action itself as well by taking into account the value of the act-plus-consequence-set of an agent's doing something. Let the act-plus-consequence-set of

an agent X's doing alpha be the set that contains (i) everything that the consequence-set of X's doing alpha contains, and (ii) agent X's doing alpha. Non-Consequentialist Act Utilitarianism interprets the Central Utilitarian Idea as follows:

NON-CONSEQUENTIALIST ACT UTILITARIANISM X ought to do alpha if and only if the act-plus-consequence-set of X's doing alpha would maximize value. (Thomson, p. 132.)

Accepting this reading of the Central Utilitarian Idea puts us in a position to make a case for the conclusion that Bloggs ought not cut up the young man and use his parts. For it might be the case that the disvalue (or "negative value") of the act-plus-consequence-set of Bloggs's cutting up the young man and using his parts to save the five is so great that Bloggs's action will fail to maximize value. If this is the case, the claim that human actions have intrinsic value seems to have a role within a value-based moral theory like Utilitarianism because it can help us achieve the desired result in Transplant that Bloggs's action is morally impermissible.

4 Can the act have the disvalue that it needs?

If we opt for Non-Consequentialist Act Utilitarianism to avoid the trouble with Transplant, from where are we to suppose that the action gets its negative value? The first thing to

Notice that if actions do not have intrinsic value, then Consequentialist Act Utilitarianism and Non-Consequentialist Act Utilitarianism yield everywhere the same conclusions because the action has no value to contribute to the value of the act-plus-consequence-set of X's doing alpha.

recall is that if Identity Theory is true and actions are mere bodily movements, then we are on our way to the conclusion that actions have no intrinsic value. If we retain Utilitarian principles, then what we are in need of in order to make our point that Bloggs ought not cut up the young man and use his parts to save the five is a theory of action in which actions can have intrinsic value. One way of getting what we need is to claim that actions are more than bodily movements or sets of bodily movements. If actions are complex entities involving other actions and events as parts, such as they are in Parthood Theory, then we can make a plausible case for their having intrinsic value.

If actions are such complex entities, then given any part of an action that has disvalue, that part will contribute to the overall value of the complex action. Bloggs's complex action is his cutting up the young man and using his parts to save the five. Among the parts of that action, his using the parts to save the five contributes positive value to the overall complex action, while his cutting up the young man and removing his parts contributes disvalue to the overall complex action. In order to achieve the result that Bloggs's cutting up the young man and using his parts to save the five is morally impermissible, we need to show that the cutting up the young man and removing his parts (the earlier part of Bloggs's complex action) has enough

⁷ See chapter one for an extensive discussion of the Parthood Theory of action.

disvalue to outweigh the positive value in Bloggs's using his parts to save the five (the later part of Bloggs's complex action) and its consequences (five lives saved). We might reason to this conclusion in the following way, suggested by Thomson,

...consider the subparts of that complex act. One subpart is the surgeon's cutting the young man up and removing his parts. Since by hypothesis the young man did not consent to the surgeon's doing this, the surgeon's cutting the young man up and removing his parts is the surgeon's committing battery and theft. A very terrible battery and theft, since the battery is cutting up and the theft is of body parts. If any acts have negative value, this battery and theft has negative value, immense negative value (Thomson, p. 136).

The battery and theft needs to be <u>so</u> intrinsically bad that the act-plus-consequence-set of the complex act of cutting up the young man and using his parts to save the five will have more disvalue than the act-plus-consequence-sets of its alternatives.

Let us consider how much disvalue we require to make this the case. Any alternative that involves not cutting up the young man and using his parts to save the five will result in four more deaths than cutting him up and using his parts will. Thus, in order for the act-plus-consequence-set of Bloggs's complex action to have enough disvalue for it to be the case that Bloggs ought not cut up the young man and use his parts to save the five, his cutting up the young man and removing his parts (that is, his committing battery and theft) must have at least slightly more than four times the disvalue of a death. Only then will the act-plus-consequence-set of the complex action have more disvalue than the act-plus-consequence-sets of any alternative.

More than four times more disvalue than the disvalue of a death is an incredibly large amount of disvalue for Bloggs's committing battery and theft to have. Admitted?" Bloggs's cutting up the young man and removing his parts is a particularly extreme battery and theft. Nonetheless, can even this battery and theft have more than four times the disvalue of a death?

Even if we grant that this battery and theft has enough disvalue, this solution is not perfectly satisfactory. In the case as described, the cutting up and removing the parts (the battery and theft) has to be an action with more than four times more disvalue than the disvalue of a death, and even that might be difficult to maintain. But what if it were the case that Bloggs could save fifteen people by cutting up the young man and removing his parts? If the act-plus-consequence-set of any alternative to Bloggs's cutting up the young man and removing his parts contains fifteen deaths, then Bloggs's cutting up the young man would have to be an action with greater than fourteen times the disvalue of a death. (The death of the young man in the actplus-consequence-set of cutting him up and removing his parts "cancels out" one death in the act-plus-consequence-sets of alternatives to cutting him up and removing his parts.) Indeed, for any n deaths in the act-plus-consequence-sets of alternatives, Bloggs's cutting up the young man has to be an action with more than n-1 times the disvalue of a death in order for us to conclude, on Non-Consequentialist Act Utilitarian grounds, that Bloggs ought not cut up the young man and use his

parts to save the patients.

So even if we can do it for Transplant when there are only five patients awaiting transplants, the <u>n</u>-case description of Transplant creates trouble. It looks very difficult to attribute enough disvalue in cases with large numbers of patients awaiting transplants to outweigh the positive intrinsic value of Bloggs's cutting up the young man and using his parts to save the patients. Yet this is exactly what we need to do in order to claim, on Utilitarian grounds, that Bloggs ought not cut up the young man and use his parts.

Our discussion so far indicates that any moral theorist who is concerned to defend a version of the Central Utilitarian Idea should show an interest in the claim that actions have intrinsic value. In particular, she should fend off any claims to the effect that they do not, which suggests that she should be wary of Identity Theory, or the Causal Theory of Intention and the No-Value Thesis. We have seen that the only way to get a morally acceptable answer in Transplant using the Central Utilitarian Idea is to suppose that actions have intrinsic value and to include that value when deciding among alternatives. Non-Consequentialist Act Utilitarianism provides a reading of the Central Utilitarian Idea that permits us to do this. But we have seen that there is some trouble in finding enough disvalue in Bloggs's action to make it impermissible.

[•] In cases involving extremely large numbers, such as thousands, millions, or billions the moral intuitions about what is permissible become less clear.

5 Intrinsic value, Utilitarianism, and the Mafia case

Suppose that we could find enough disvalue in Bloggs's cutting up the young man and removing his parts to handle Transplant. How are we to handle the following variation on Transplant, called "Mafia"? In Mafia an organ bank had all of the necessary parts for the five patients, Bloggs has successfully performed all five transplants, and the five are busy recovering from their operations. "[T]he Mafia now tells Bloggs, 'We will kill those five patients of yours unless you cut this young man up and remove his parts, and thereby kill him'"(Thomson, p. 141). Let us complete the picture by adding to this scenario the fact that the Mafia will kill the five by cutting them up and removing their parts.

Now we ask: May Bloggs cut up the young man and remove his parts to save the five? The answer in this case seems just as clear as in Transplant: Bloggs's cutting up the young man and removing his parts is morally impermissible. The mere fact that the Mafia is threatening his five recovering patients does not make it morally permissible for Bloggs to cut up the young man and remove his parts. But how are we going to get enough disvalue into Bloggs's cutting up the young man and removing his parts to outweigh the disvalue of the Mafia's cutting up the five and removing their parts? We cannot do it by appealing to the fact that Bloggs's action is an immensely disvaluable battery and theft because there will be four additional, equally disvaluable

batteries and thefts if Bloggs does not cut up the young man and remove his parts. In short, Bloggs will maximize value by cutting up the young man and removing his parts. Since Non-Consequentialist Act Utilitarianism calls upon us to maximize value, it yields the morally unacceptable conclusion that Bloggs ought to cut up the young man and remove his parts under the threat from the Mafia to kill the five recovering patients.

Mafia indicates that even if we suppose that actions have intrinsic value, there are cases in which Non-Consequentialist Act Utilitarianism morally requires actions that any acceptable moral theory should deem impermissible.

6 Conclusions about Utilitarianism and intrinsic value

This discussion began with the suggestion that Transplant might demonstrate one way in which ontological conclusions about actions can conflict with moral theory. It looked as if moral theory, particularly moral theories based on the Central Utilitarian Idea, could benefit from the assumption that actions have intrinsic value. However, Mafia shows that the real problem for the Central Utilitarian Idea lies outside the issue of the intrinsic value of action. As we have seen, even if we opt for an interpretation of the Central Utilitarian Idea that leaves room for the intrinsic value of action, it only gives a morally acceptable answer in the simple case of Transplant, but fails to give us an acceptable answer in the Mafia case.

An adequate discussion of why it is morally impermissible

for Bloggs to cut up the healthy young man in either case is beyond the scope of this paper, the main concern of which is to see if there is any relationship between the ontology of action and moral theory. For the same reason, I shall not attempt to address in any detail what the source of trouble for the Central Utilitarian Idea is, but I will mention briefly where I think it lies. It is a frequent objection against Utilitarianism that it fails to take rights into consideration, thus on occasion prescribing apparently unjust courses of action. In the case of Bloggs and the young man, given that the young man has not consented to be cut up in either Transplant or Mafia, he has the right not to be cut up. By failing to respect that right, that is, by blatantly violating it, Bloggs is undertaking a morally impermissible act if he cuts up the young man for whatever reason. This seems to be the core of the problem in both Transplant and Mafia.9

At the outset of this paper we looked at three theses which together have the result that actions, whether intentional or not, have no intrinsic value. We then looked at one way in which this conclusion might conflict with what is required or desirable for ethics. While Transplant seems to indicate that advocates of the Central Utilitarian Idea could avoid false conclusions by supposing that actions have intrinsic value, Mafia shows that even with that supposition the Central

[•] For a detailed philosophical discussion of the nature of rights, see The Realm of Rights by Judith Jarvis Thomson.

Utilitarian Idea yields false conclusions. As a result, it seems not to make much difference to any reading of the Central Utilitarian Idea whether or not actions have intrinsic value. When we suppose that they do and that that value should count when deciding among alternatives, we manage only to limit the cases in which the Central Utilitarian Idea prescribes actions that are morally impermissible, we do not eliminate them.

Are any other moral reasons to think that actions have intrinsic value?

7 Intrinsic value and moral theory

It is my view that moral theory does not require that actions have intrinsic value; it is enough for ethics that actions can get their value from their causes and/or their consequences. I cannot think of any way in which moral theory would suffer if actions did not have intrinsic value. We have already seen that the intrinsic value of actions makes little difference even for the value-based theory of Utilitarianism. If it matters so little for a value-based theory, then it should matter even less for any other sort of ethical theory in which value plays a secondary role. Now there may be a theory in which value plays a smaller role and in which it is necessary for actions to have intrinsic value. If there is one, I do not know what it is, nor do I have any idea what it would look like.

To the extent that I believe that ethical theory does not depend upon the claim that actions do have intrinsic value, I

believe that ethics is compatible with theories of action that yield the conclusion that they do, as well as with theories of action that yield the conclusion that they do not have intrinsic value. This is not to say, of course, that a theory like Identity Theory, in conjunction with other theses that together yield the result that they do not have intrinsic value, does not constrain moral theory. It is only to say that this constraint is not one that seriously interferes with the task of a moral theory.

8 Ontology and ethics

This paper began with a promise to address a much more general question than the one about intrinsic value. In asking whether the moral theorist ought to be concerned with ontological conclusions about the nature of actions, I was inquiring whether anything in the ontology of action has to do with anything in ethics. I shall argue that none of the ontological issues about what acts are constrains what can be said in ethics in any substantive way. In addition, I shall discuss a proposal that there are no actions at all and consider what the impact such a conclusion would have on ethical theory.

Since ethics seems to have actions as a large part of its subject matter, it is natural to think that the ontology of action might bear on ethics. An analogy between the impact that the ontology of numbers has on mathematics and the impact that the ontology of action has on ethics will help us understand what

is at issue. 10 For just as actions form a central part of the subject matter of ethics, so do numbers form a central part of the subject matter of mathematics.

In the philosophy of action and in the philosophy of mathematics, there are two basic ontological questions to ask:

Are there any actions/numbers?

- (1) What are actions/numbers?
- Are there ways of responding to these questions that are going to constrain the practices or claims of the moral theorist or the mathematician? The kind of constraints they place will depend on how we view the role of the metaphysics.

With respect to the first question, we are not likely to think that a particular analysis of number is going to have a detectable impact on mathematics. It seems reasonable to take a conservative approach to the ontology of numbers and aim for a metaphysical analysis that does not conflict with accepted laws of mathematics. Whatever we say about what numbers are, odd numbers will still be odd, even numbers will still be even. If that turned out not to be the case, then we would have good reason to reject the proposed analysis, not to revise our mathematics. Similarly, a conservative approach to the ontology

ethics and mathematics, particularly in the way each is done.

Many mathematicians might be happy to be working with constructs and abstract systems constrained only by logical rules and principles. Moral philosophers are constrained by more than logic when they are developing ethical systems. The main analogy is between the place of numbers in mathematics and the place of actions in ethics.

of action will not have a substantive impact on ethical theory. Part of the business of an analysis of action is, among other things, to account for, or at least not conflict with, our basic ethical intuitions. For example, the theories of action that have so far been considered - Trope Theory, Parthood Theory, and Identity Theory - do not force any strong revisions on ethical theory. They are generally conservative insofar as whether an action is a trope, a bodily movement, or a more complex entity made up of other actions and events is not going to have an impact on its basic ethical status as right or wrong. Indeed, if a metaphysical theory of action yielded wrong ethical conclusions, that is, yielded conclusions that conflict with basic ethical intuitions, then we would have a reason to reconsider the metaphysical analysis unless it gave us a very convincing reason to reject our moral intuitions.

Before moving on to (2), note that while mathematics does not require any specific notion of number, it does need to presuppose that numbers have certain properties. Similarly, there are certain suppositions in moral theory about the properties of actions. One common presupposition in moral theory is that actions have causal efficacy. Almost every moral theory requires that they do. Some Utilitarian theories, for example, claim that actions are right or wrong in proportion to the goodness or badness of their consequences. If the theory works with a notion of consequences requiring that events or states of affairs that are consequence of an action, must have been caused

by that action, then the theory requires that actions can enter into causal relations. Not only Utilitarianism, but most ethical theories assume that what our doing alpha or beta <u>causes</u> is at least relevant to whether we ought or ought not do alpha or beta.

9 Ethics and the claim that there are no actions

What would a negative response to (2) mean for mathematics and for ethics? As with (1), our response to (2) is going to depend on what the ramifications of a negative response might be. The mathematician might be able to endure a negative response to (2) without having to revise her theories because she is constructing logical systems. Moral theory, on the other hand, is not something that we want to be simply the construction of logical systems. We would like our claims to be about something real. Logical consistency is a necessary but not a sufficient condition of a correct moral theory. The claim that there are no actions seems to be a claim about the ontology of action that could make a difference to ethics.

There are two ways of spelling out the claim that there are no actions. First, to say that there are no actions might be to say that actions do not form a <u>basic</u> ontological category. Holding this view, one would give a reductive or supervenience analysis of actions in terms of more basic entities, for example, the properties and spatiotemporal zones of Jonathan Bennett's

¹¹ I do not mean to be understating the task of the mathematician.

Trope Theory. Second, to say that there are no actions might mean that actions do not form a <u>real</u> ontological category.

Terence Horgan endorses this eliminative account of action, claiming that actions and events have no place in our ontology at any level. The first claim, that actions are not basic, does not have a negative effect on moral theory. Supervenience does not deny that we are referring to <u>something</u> when we speak of "my writing this sentence" or "my killing your cat for fun." If actions are reducible to something more basic, then our ethical claims about actions may in fact be claims about something more basic, but that does not make them in any way illegitimate. Nor does it require us to rethink the way we view our ethical claims.

The second claim, that actions are not real, is the one that looks as if it might spell trouble for ethics. In "The Case Against Events," Horgan takes himself to be doing much more than giving a supervenience account or a reductive explanation of events. He is trying to show that there is no theoretical need to have events in our ontology at all. He argues that "since their elimination yields an important simplification of ontology, we should banish them from existence" (Horgan, p. 28). It is unclear exactly how we are to interpret Horgan's claim that there are no events. He seems very careful to avoid talk of "reduction" in favour of talk of "elimination." He seems to be claiming that events have something of the status of witches.

Terence Horgan, "The Case Against Events," in The Philosophical Review, LXXVII, No. 1, January 1978, pp. 28-47.

That is, insofar as they do not exist, any claims that we might make about them are false. He does not claim, however, that there are no other ways of explaining the issues that we typically explain by reference to events. Horgan does not offer a uniform account of how to reduce events to other more basic entities. For this reason, his account as he sees it is different from a reductive analysis. There are two separate questions about Horgan's project. First, does he successfully show that "we should banish event from existence?" And second, suppose it has been shown that there are no events, is it the case that moral theory needs to dispense with them?

10 Horgan's project

Horgan proceeds by pointing to a number of areas of philosophy that have traditionally relied on the notion that actions and events are particular, unrepeatable entities. He says,

the mind-body problem is often regarded as the problem of the relation between mental events and physical events; discussions of scientific explanation usually assume that the entities explained are events; causation is usually treated as a relation between events;...and actions, a species of event, are the very subject matter of action theory. (Horgan, p. 28.)

Beginning with causal relations and proceeding through the other philosophical areas in which events figure prominently, Horgan suggests ways of discussing the issues without making references to events. His main concern is to avoid using event-designating singular terms in philosophical discourse. For each of the areas

to which he draws attention, he offers an interesting way of avoiding such terms. In addition to causal relations, Horgan addresses the issue of describing and redescribing actions and events in the course of explaining them, how to discuss the mind-body problem without referring to mental and physical events, and how to preserve logical entailments in action-sentences involving adverbial modification without following Davidson's suggestion that we quantify over events.

In order to address the issue about ontology and ethics it is not necessary to give a detailed account of Horgan's procedure in each case. He does not provide a systematic way of replacing event-designating singular terms with other linguistic forms, but rather deals with each "event-ridden" area of philosophy differently.

He notes that in philosophy of mind, for example, some theories identify mental events with physiological events. If we are to make sense of these theories at all, then events must be particulars. Horgan maintains that we need not suppose that events are particulars for we could instead identify mental attributes with physical attributes, and carry on the discussion on that level. In addition, he suggests that the tired issue of the relationship between mental events and physical events can just fall by the wayside because events do not exist (Horgan, p. 44).

He deals with adverbial modification much differently.

Davidson has suggested that in order to preserve certain logical

entailments in action sentences with adverbial modifiers, we need to quantify over events. For example, if the logical form of a sentence reflects its logical entailments, then the logical form of "Sebastian strolled through the streets of Bologna at 2 a.m." should reflect the entailment "Sebastian strolled through the streets of Bologna". Davidson suggests that we may do this by quantifying over events so that the logical form of the first sentence is "There is an event x such that Sebastian strolled x, x took place in the streets of Bologna, and x was going on at 2 a.m." (cf. Davidson, "The Individuation of Events" and Horgan, p. 44). Taking Davidson's proposal as a challenge to give an alternative account of the logical form of such statements, Horgan suggests following Romane Clark. 13 Clark expands first order quantification theory specifically to deal with all types of predicate modifiers, including adverbs. According to Horgan, Clark "represents these as operators which, when prefixed to a predicate, yield a new predicate" (Horgan, p. 44).

without going into the details of this proposed expansion of first order quantificational logic, it is easy to see that it is a very different proposal from the one that Horgan makes for dealing with events in the mind-body problem. What he says about causation and explanation are different still. It is for this reason that we can take Horgan to be giving something other than a reductive account of events. He wants to deny their existence

Romane Clark, "Concerning the Logic of Predicate Modifiers" in Nous IV, 1970, pp. 311-35.

altogether, concluding

that there is no apparent theoretical need to posit events. Of course it is still possible that events will prove indispensible for philosophical purposes other than those we have considered here, or for theoretical purposes within physics or some other branch of science. But unless and until one of these possibilities becomes actual, which I doubt will happen, the most reasonable course is to invoke theoretical parsimony and deny the existence of events. (Horgan, p. 47.)

11 Horgan's project and ethics

Horgan's suggestions for doing without events are interesting. But he has not exhausted the aspects of philosophy in which there is a reliance on events, and since he has not offered a systematic procedure, he is going to have to individually address every additional area that is brought to his attention. One significant oversight in Horgan's project is his failure to consider the impact that an event-less ontology might have on ethics. Ethical theory relies in many ways on actions and events. Indeed, agents and their actions are the central objects of moral evaluation in any ethical theory. We evaluate agents for their goodness or badness, and we evaluate acts in terms of their rightness or wrongness. Furthermore, many of our evaluations take into account what agents and their acts cause.

In an ontology devoid of actions, we need to explain how moral judgments can be true. If there are no actions, then it looks as if there can be no right or wrong actions. And if there are no right or wrong actions, then how can ethics tell us that we ought or ought not perform them? For example, consider the

moral judgment "I ought not to have killed your cat for fun." On the face of it, this sentence seems to ascribe a particular property, i.e. wrongness, to a particular action of mine, i.e. killing your cat for fun. One natural way to represent the logical form of this ought-sentence is "there is an action x such that x is a killing of your cat for fun by me and x is wrong."

In order to make sense of the moral judgment in an ontology that does not include actions we need to give the logical form of this ought-sentence in a way that does not quantify over actions.

Based on Horgan's way of dealing with adverbial modification, there is a response to this challenge in the offing. Rather than quantifying over actions and assigning the property "wrongness" to them, we could treat "ought" as an operator. Instead of giving the logical form of "I ought not have killed your cat for fun" as we did above, we give it as "it ought not be the case that I kill your cat for fun." This way, we do not talk about an agent's actions, but about what the agent did. So, we need not say that my killing your cat for fun is wrong, since we can say that it is wrong that I killed your cat for fun. In this way, we avoid any singular terms that designate actions while still accounting for individual moral judgments.

Even if particular moral judgments can be true compatibly with there being no actions, there remains a difficulty concerning the broader generalizations of moral theory. For example, consider the earlier claim that no action has intrinsic value. Without generalizing over actions, it is difficult to

understand how we might convey the content of this claim. If Horgan had offered a reductive or supervenient account of actions in terms of more basic entities, then we would be able to make sense of these generalizations in those terms. Since he does not do this but proposes instead to eliminate them from our ontology altogether, this route is not available to him. The inability to handle these broader generalizations is not, of course, a crippling obstacle for this proposal about actions. For, if there are no actions, then it would not make sense to speak in terms of their having or lacking intrinsic value. Insofar as Horgan is claiming that there are no actions, there is no entity of the sort that we had supposed of which it is coherent to claim that it has intrinsic value. We shall have to locate intrinsic value elsewhere.

12 What has Horgan shown?

To the extent that Horgan has only addressed a limited range of philosophical contexts in which there is a heavy reliance on events, he has not succeeded in showing that events have no place in our ontology at any level. At most, he has shown that there are ways, although not systematic ones, of avoiding reference to events and actions in some philosophical contexts. His failure to take moral theory into account strikes me as a serious oversight in his project. If we cannot account for the broader generalizations of moral theory without actions, then that gives us an additional reason for supposing that actions are an

important part of our ontology.

Even if there is some viable way of accounting for these generalizations in an ontology devoid of events, has Horgan shown that we have good reason to dispense with them? I do not believe that he has. The mere fact that we can do away with an entity does not, on its own, provide adequate grounds for claiming that we need to eliminate that entity from our ontology. Regardless of whether it is possible for ethics, or any other area of philosophy, to get along without actions, we save ourselves a great deal of metaphysical work by having actions and events in our ontology. I grant that the principle of Occam's Razor is a good heuristic to follow, but we need also to consider explanatory simplicity and commonsense.

Having actions and events as part of our ontology provides us with a systematic way of addressing a whole range of philosophical issues from causation to the mind-body problem. There is no apparent theoretical benefit to be gained from replacing each form of analysis that relies on events with some other more complicated analysis that does not rely on them. Far from showing that events have no place in our ontology on any level, the complexity of Horgan's reformulations shows how much simpler our philosophical tasks are when we do rely on actions

Quine makes this point about concepts: "It is one of the consolations of philosophy that the benefit of showing how to dispense with a concept does not hinge on dispensing with it" in Word and Object (Cambridge: MIT Press, 1960, p. 190). Daniel Stoljar brought this passage to my attention in his helpful comments on this section.

and events.

Conclusion

Horgan has not shown that we do not need actions and events at any level. Indeed, it is difficult to understand what he means in suggesting that they do not exist since he is quite clearly offering something different from a reductive or supervenience account. But, even if Horgan has shown that we can do without actions and events, at least in the limited cases that he addresses, he has not shown that we must. For the purposes of ethics, but also other philosophical areas, we gain theoretical simplicity by supposing that there are actions and events. Embracing the extreme claim that there are no actions would require us to rethink the way that we think about some of our ethical claims. For that very reason, we might question the theoretical benefits of accepting that ontological claim about actions and events.

In general, with respect to the impact that metaphysics has on ethics, neither metaphysics nor ethics should be taken as prior. As the discussion of Horgan's project has shown, it is not necessary to embrace metaphysical conclusions that require a radical shift in our thinking about ethical theory. By the same token, an ethical theory that completely undermined our basic metaphysical intuitions would also be problematic. As long as we take a conservative approach to each area, respecting basic intuitions about them, it is likely that metaphysics and ethics

will not constrain one another in any substantive way.

BIBLIOGRAPHY

- BENNETT, JONATHAN. Events and Their Names. Cambridge: Hackett Publishing Company, 1988.
- DAVIDSON, DONALD. "Agency" in Essays on Actions and Events.
 Oxford: Clarendon Press, 1980. pp. 43-61.
 -----"The Individuation of Events" in Essays
 on Actions and Events. pp. 163-187.
- HORGAN, TERENCE. "The Case Against Events" in The Philosophical Review, LXXXVII, January, 1978. pp. 28-47.
- QUINE, WILLARD VAN ORMAN. Word and Object. Cambridge: The MIT Press, 1960.
- THOMSON, JUDITH JARVIS. The Realm of Rights. Cambridge: Harvard University Press, 1990).

ESSAY THREE

RESPONSIBILITY AND AUTONOMY

It is a commonplace in moral theory to say that responsibility requires autonomy, in some appropriate sense of the two crucial terms in the slogan. Susan Wolf has recently made a new suggestion about how we should understand these terms. I am going to assess her view by contrasting it with another proposal which I shall look at first.

1 The Choice View

This proposal is very familiar. Wolf calls this The

Autonomy View, but since that is the term in the slogan, I shall

call it The Choice View. The Choice View invites us to interpret

the slogan as follows:

A person X is blameworthy or praiseworthy for one of his or her acts - or for an outcome of one of his or her acts - only if X could have done otherwise.

What motivates this view is clear enough. There are many examples of cases in which we feel disinclined to blame or praise an agent for her action because it seems right to think that the agent could not have done otherwise. For example, consider the following two cases.

First, consider Kleo, the kleptomaniac. She has a psychological compulsion to steal. Do we blame her when she steals from the department store? No, we do not. One good reason for thinking that Kleo is an inappropriate candidate for blame is that she could not have done otherwise. Her psychological compulsion makes it the case that she has no choice but to steal from the department store.

Second, consider Helen, the agent under hypnosis. Helen has been hypnotized into being kind, considerate and generous to everyone around her. Do we owe her gratitude and praise for her generosity when she brings us flowers or offers to take us out to dinner? No, we do not. One good reason for thinking that we do not owe Helen our gratitude is that she could not have done otherwise. The fact that she was hypnotized into treating those around her with nothing but kindness, consideration, and generosity makes it the case that she has no choice but to bring us flowers and offer to take us to dinner. If we owe anyone our gratitude, it is the hypnotist.

In each of these cases it seems appropriate for us to withhold praise or blame. We do not hold the agents responsible and the fact that they could not have done otherwise seems to be a good reason to think it is appropriate for us to take this attitude towards their behaviour.

There are at least three ways of interpreting "could have done otherwise." The first way of understanding "could have done otherwise" comes from G.E. Moore. According to his interpretation, an agent could have done otherwise only if the agent would have done otherwise if she had so chosen. This reading of "could have done otherwise" explains why Helen and Kleo are not responsible for their behaviour, as well as explaining certain other problem cases. Kleo is not blameworthy because her psychological compulsion inhibits her from following any other course of action. Someone might be tempted to say that

Kleo could not even have chosen not to steal. But it is easy to imagine that a kleptomaniac like Kleo might decide not to steal and yet be driven to do so nonetheless by her psychological compulsion. Thus, Kleo could not have done otherwise because she would not have done otherwise even if she had so chosen.

Similarly, although we might say that Helen could not, because of the hypnosis, have chosen to act differently, we can also imagine that Helen might choose to act differently and yet, because of the hypnotic suggestion, she finds herself performing these generous acts. Thus, Helen would not have done otherwise even if she had so chosen, and that is why she is not praiseworthy for her behaviour.

Other cases for which this interpretation of "could have done otherwise" accounts include cases of coercion or physical constraint. If Carl is coerced into handing over all of the money in the bank's vault, we do not blame him for his behaviour the way that we would if he simply gave the money away for no good reason. Given that he and all of his colleagues will be shot dead if he does not give over the money, Carl cannot reasonably do otherwise under the circumstances. In a case

One might want to say that in this case, Carl does in fact have a choice and so should be considered responsible. He could have decided not to hand over the money, in which case everyone would have been shot. If Carl would have done otherwise if he had so chosen, that is, would have kept the money if he had so chosen, then he is responsible according to this interpretation of The Choice View. But the important thing to notice here is that he Carl is not blameworthy for his behaviour. If he is at all responsible for handing over the money, then he is praiseworthy for choosing the most reasonable alternative under the circumstances.

involving physical constraints, Philip would not be blameworthy for failing to alert the police, for example, if his reason for failing to do so was that he was gagged and tied to a chair. His circumstances made it impossible for him to get to a telephone. Even if he chose to alert the police, he would not have done otherwise than sit helplessly in his chair. As a result, Philip is not blameworthy for his behaviour.

The second way of understanding "could have done otherwise" says that an agent could have done otherwise only if she was not causally determined to act as she did. It is this sense of "could have done otherwise" that appears to conflict with the doctrine of determinism. For if determinism is true, then it looks as if no agent could ever have done otherwise, and therefore, no agent is ever praiseworthy or blameworthy for her behaviour. This sense of "could have done otherwise" blurs the distinction between the normal case and the problem cases of psyhological compulsion, coercion, and physical constraint.

A third interpretation of The Choice View's condition that the responsible agent "could have done otherwise" will emerge shortly when we examine what Wolf objects to in The Choice View and why she proposes her own view, The Reason View. For the moment, we have two readings of "could have done otherwise." They are:

- (1) An agent X could have done otherwise only if X would have done otherwise if X had so chosen.
- (2) An agent X could have done otherwise only if X was not causally determined to act as he or she did.

Interpretation (1) provides the most plausible reading for The Choice View. It accounts for the problem cases in which we are disinclined to praise or blame an agent for what we would, under normal circumstances, consider praiseworthy or blameworthy behaviour. At the same time, it explains why we do praise and blame agents in the normal case since in the normal case we assume that an agent would have done otherwise if she had so chosen. Only if an agent would have done otherwise if she had so chosen, is she responsible for her behaviour. Interpretation (2) raises a host of issues about the compatibility of determinism and responsibility. We shall not be focusing on those issues here.²

2 Wolf's critique of The Choice View: The Reason View Susan Wolf objects to The Choice View for reasons we shall see in a moment. Her view, The Reason View, interprets the slogan as follows:

Although I shall not directly address them in the body of the present paper, Harry Frankfurt draws attention to certain counterexamples in which an agent could not have done otherwise yet is nonetheless responsible. In "Alternate Possibilities and Responsibility" he focuses on cases of coercion in which the agent had already decided to do the act in question before she was threatened with an unduly harsh penalty unless she did that very thing. Since the penalty is so harsh that any reasonable person would submit to the demands of the coercer, The Choice View would have it that the agent is not blameworthy for her behaviour. In the end, Frankfurt offers an amended version of The Choice View:

An agent X is responsible for one of his or her acts - for an outcome of one of his or her acts - only if X did it only because x could not have done otherwise.

An agent X is blameworthy for one of his or her acts - or for an outcome of one of his or her acts - only if X could have done otherwise. An agent X is praiseworthy for one of his or her acts - or for an outcome of one of his or her acts - only if X does the right action for the right reasons.

The Reason View has two distinctive features. First, it divorces blameworthiness from praiseworthiness, putting different conditions on these two kinds of responsibility. Second, it offers a new account for only one of these. The Reason View, like The Choice View, says that blameworthiness requires that an agent could have done otherwise.

Based on a third interpretation of what it is to have the ability to do otherwise, Wolf maintains that once we develop a conception of what an agent who is autonomous in the sense required for The Choice View is like, "we will come to see it as something we don't particularly want to be like, and, more to the point, as something we don't need to be like in order to make sense of and justify our sense of ourselves as responsible beings" (p. 48). She suggests that a different account be given for praiseworthiness.

The third interpretation of the kind of autonomy required for The Choice View, suggested by Wolf, sees it as a more stringent than the ones we have considered so far. On this interpretation, the ability to do otherwise requires that an agent be in <u>ultimate</u> control of her behaviour and have the ability to make <u>radical</u> choices. Both of these notions need clarification.

According to Wolf, the will of an agent who is in ultimate

control of her behaviour "must be determined by her self, and her self must not, in turn, be determined by anything external to itself" (p. 10). The will is determined by the agent in that the choice that she makes must be "up to" the agent. As responsible beings "we choose to do some things rather than others, and nothing makes us choose" (p. 11). Anything less than ultimate control undermines responsibility on The Choice View because it suggests that something external to ourselves might be making us choose as we do.

Wolf points out that this requirement immediately creates a difficulty for The Choice View because it seems an impossible condition to meet. Not only our environment but also our heredity, both of which are external to ourselves, contribute something to the choices that we make and influence the desires that we have and upon which we frequently act. Wolf is going to sidestep altogether the issue of whether we can even have ultimate control, and claim instead that we neither want it nor need it for responsibility.

closely related to the condition of ultimate control is the ability to make <u>radical</u> choices. Any agent with the ability to do otherwise must, claims wolf, be able to make radical choices. She contrasts these with <u>rational</u> choices made on the basis of reason. A radical choice "must be made on no basis and involves the exercise of no faculty" and as such "there can be no explanation of why or how the agent chooses to make the radical choices she does" (p. 54).

Wolf understands reason to be a faculty that an agent cannot help having, and therefore as an influencing factor that undermines an agent's ultimate control of her actions. Much in the way that an agent cannot help having many of the desires she has and acting upon them, Wolf says

[i]f Reason is similarly a property or faculty that an agent cannot help having (or lacking), and if it generates motives that an agent cannot help acting upon, then an agent who acts in accordance with Reason is likewise not in ultimate control of her actions (p. 52).

Viewing reason, not instrumentally, but normatively, Wolf claims that it is the faculty or set of faculties that we use to recognize good values and to form true beliefs. Our ability to use reason gives us the capacity to reflect upon our values and our desires, to rationally evaluate them, and to reflect upon our involvement in the situations in which we find ourselves. the agent who can make radical choices, insofar as there is no explanation of how she makes the radical choices she does, must be able to act in accordance with reason or not. She must be able to make both rational and irrational choices, and "she must be able to regard the most rational course of action, insofar as there is one, as just one alternative among others" (p. 54). The ability to make radical choices translates into the ability to choose with or against reason with no apparent explanation for which course of action one takes. Reasons and desires are viewed as external factors that threaten the agent's ultimate control over her behaviour.

Keeping in mind what she means by ultimate control and how

that explains her view of radical choice, we may state Wolf's interpretation of "could have done otherwise" as follows:

(3) An agent X could have done otherwise only if X had the ability to make a radical choice, that is, to choose on no basis.

Once Wolf has presented her conception of reason, stated that The Choice View requires ultimate control, and explained why this issues in the ability to make radical choices, she asks the following question: Why should we want to have the ability to do otherwise if it requires that we be able to make radical choices, that is, choices that could be irrational? She says:

To see this as an ability one might want to exercise, one must view the possibility of acting irrationally as potentially desirable. In other words, one must think that irrational action may be as attractive as rational action. If one thinks that acting with Reason may be no better than acting against Reason, then one can see the ability to choose whether to act with or against it as an increase of options that intelligent and perceptive agents might intelligibly want to exercise (p. 56).

Not only does Wolf view the ability to do otherwise as undesirable, but also, insofar as it involves ultimate control, she views it as not necessary for responsibility. Wolf argues that an agent can be praiseworthy for her behaviour without being able to choose against reason, that is, without having the ability to make radical choices, as long as she does the right thing for the right reasons. It is this claim that she offers as a justification for introducing The Reason View with its new necessary condition for praiseworthiness.

Wolf's case rests primarily on an example in which she contrasts two very different agents whom I shall call Amy and

Rachel. Amy, the first agent, is autonomous in The Choice View's sense. She has the ability to do otherwise. As we have seen, Wolf interprets this as the ability to make radical choices. By contrast, Rachel, the second agent, is not similarly autonomous. She does not have the ability to make radical choices. She can only choose in accordance with reason; she cannot choose against reason. Given Wolf's interpretation of what it is to have that ability, Rachel does not have the ability to do otherwise.

Each agent finds herself faced with the following situation: She is walking along a river and sees an unaccompanied child drowning in the water. Each agent recognizes that unless she saves the child, the child will drown. Both Amy and Rachel respond appropriately to the situation by diving into the water and saving the child. Clearly, the rational alternative in this situation is to save the child. The essential difference between Amy and Rachel is that Amy, unlike Rachel, could have rejected the rational alternative and acted irrationally instead. For example, she might have decided to save her hair-do instead of the child. Wolf explains,

The difference between the two cases consists solely in the fact that the first agent, unlike the second, did not have to jump in. Despite the existence of clear and decisive reasons to save the child, the first agent, unlike the second, could have remained on shore. Perhaps the first agent could have thought "I won't do it," whereas for the second agent, given the circumstances, such an idea was unthinkable. But one could equally suppose that he second agent could have had the thought - she simply could not have taken it seriously (p. 59).

Amy is supposed to illustrate that the ability to do

otherwise is not a desirable ability to have. Rachel, by contrast, is supposed to illustrate that The Choice View does not capture the conditions for praiseworthiness since Rachel could not have done otherwise. She could not reject the rational alternative and so lacked the ability to make a radical choice. As a result, The Reason View is proposed in which a new clause for praiseworthiness accounts for agents like Rachel.

3 Amy, Rachel, and the ability to do otherwise

Now, Wolf interprets the ability to do otherwise as the ability to make radical choices. Both agents could have thought not to save the child, but only Amy could have taken the thought seriously. Later, we shall see what it might mean to say that Rachel could not have taken the thought seriously. First, what does it mean to say that Amy could have taken the thought seriously? For it is the fact that Amy could have seriously thought not to save the child that is supposed to suggest that Amy's ability to do otherwise is not something that we should view as a desirable characteristic.

Wolf has already said that anyone who views the ability to do otherwise as a desirable ability to have must view the option of choosing against reason as a potentially attractive option. Hence, she seems to be suggesting that Amy's ability to make radical choices can only be viewed as desirable if we view the alternative of choosing to save her hair instead of the child as a potentially desirable alternative. Furthermore, since Rachel

could have thought not to save the child but could not have taken the thought seriously, whereas Amy could have taken the thought seriously, we are given to understand that Amy could have viewed the alternative of choosing to save her hair instead of the child as a potentially desirable alternative. It is not simply the case that Amy would have saved the child if she had so chosen, but that given that Amy's choice was a radical one, there is supposed to be no apparent explanation for why she chose the rational course of action over the irrational course of action.

Rachel could not have taken the thought not to save the child seriously. The irrational alternative is simply not somthing that it is within Rachel's power to choose. She does not make radical choices; she makes rational choices. When faced with a clearly rational course of action such as saving the drowning child, Rachel, unlike Amy, cannot choose against reason. She is not in ultimate control of her behaviour because she is under the very strong influence of reason. Hence, on Wolf's view of what it is to be able to do otherwise, Rachel cannot do otherwise and she would not be responsible on The Choice View. But given that Rachel followed the rational course of action because it was the rational course of action, she should, it seems, be a candidate for praise. She is to be understood as a superior moral agent whose grip on reason is so strong that she cannot overcome it. Her reaction to her situation is analagous to someone's getting a grip on a mathematical truth. understand that 2 + 2 = 4, for example, we cannot help but

believe it. Our rational natures demand that we do. Similarly, once Rachel understands what the right thing to do is, her rational nature demands that she do it.

According to The Reason View, an agent such as Rachel can be praiseworthy for her behaviour even if she could not have done otherwise, as long as she does the right thing for the right reasons. Given that The Reason View retains The Choice View's requirement that an agent is blameworthy for her behaviour only if she could have done otherwise, Wolf is proposing an asymmetry between the conditions under which an agent is blameworthy for her behaviour and those under which she is praiseworthy.

then the ability to do otherwise is not a desirable characteristic to have and is not required for praiseworthiness. As a result, an additional clause is required in order to capture the conditions for praiseworthiness. I shall argue that neither of the points that Wolf makes against The Choice View poses a serious challenge to it. Although Wolf is right in thinking that responsibility does not require ultimate control, she is wrong in thinking that a plausible interpretation of The Choice View has ultimate control as a requirement for responsibility. Furthermore, the more plausible reading of The Choice View than the one that Wolf suggests captures the sorts of cases that she has in mind when she introduces the additional clause for praiseworthiness. As a result, the second clause of The Reason View turns out to be superfluous. The Choice View succeeds in

capturing the correct interpretations of autonomy and responsibility.

4 Amy and the desirability of the ability to do otherwise

Wolf's characterization of Amy's ability to do otherwise does make us wonder why anyone should want to have that ability. For implicit in Wolf's portrayal of Amy is the claim that Amy could have taken seriously the thought not to save the child, which suggests that Amy had the capacity to view it as an attractive alternative. Intuitively, we think that anyone who could view letting a child drown as an attractive alternative has a questionable moral character. However, the mere fact that Amy could have failed to save the child need not suggest that she viewed failing to save the child as a desirable thing to do. An undesirable option is an option nonetheless. Indeed, it seems to say something positive about someone's moral character that she chose to do the right thing instead of the wrong thing.

The central difficulty in Wolf's critique of The Choice View is her claim that the ability to do otherwise requires ultimate control and radical choice. Although it is true on any interpretation of The Choice View that the responsible agent could have done otherwise and so could have chosen against reason, we need not take this supposition the further step that Wolf takes it. That is, we have very little reason to think that in order for an agent to be responsible on The Choice View there can be "no explanation of why or how the agent chooses to make

the radical choices she does" (p. 54). There is a large gap between supposing that an agent has the possibility of acting against reason and supposing, as Wolf then does, that such an agent's behaviour is inexplicable. It need not be the case that someone who could have done otherwise must make all of her choices on no basis even when some rational basis exists. An agent might have the option of choosing against reason and yet still regard reason as primary. Wolf is mistaken in suggesting that regarding reason as primary entails that one could not have done otherwise. But in contrasting Amy with Rachel, this is exactly what she is suggesting.

By stipulation, Amy could have done otherwise; she could have saved her hair instead of the child. Amy did not so choose. Nevertheless, we have no basis upon which to claim that her choice cannot be explained. Given her circumstances, Amy recognized that her reasons for saving the child far outweighed her reasons for saving her hair. Far from being unable to explain Amy's behaviour, we can explain it by pointing out that she acted on her strongest reasons. Now if ultimate control and radical choice require that there is no explanation of Amy's behaviour, she must not have been in ultimate control nor made a radical choice. But Wolf has stated that Amy fulfills these very conditions, for she is responsible on The Choice View. Since there clearly is an explanation of why Amy acted as she did and, since by Wolf's stipulation and in contrast to Rachel, Amy could have done otherwise, it appears that "could have done otherwise"

does <u>not</u> require ultimate control nor radical choice. For ultimate control and radical choice require that there can be no explanation of the agent's behaviour. Amy's choice, however, is easy to explain.

Given that "could have done otherwise" does not rule out all explanations of an agent's behaviour, nor does it force us to say that such an agent must be able to find the irrational course of action a desirable or potentially desirable thing to do, Wolf's claim that the ability to do otherwise is an undesirable ability to have loses its force. So while Wolf is right to think that we do not need ultimate control in order to be responsible, she is mistaken in thinking that The Choice View requires responsible agents to have ultimate control and make radical choices of the sort that she outlines. Hence, her first reason for rejecting The Choice View is unconvincing. It remains to be seen whether the second clause of The Reason View is a necessary addition to the analysis of responsibility. If it is, then even if the ability to do otherwise is not undesirable, it fails to capture the conditions for praiseworthiness.

5 Do we need the praiseworthiness clause?

Wolf claims that there are certain agents, such as Rachel, who are praiseworthy for their behaviour even though they could not have done otherwise. These agents cannot do otherwise because they cannot make the radical choices that her interpretation of that ability requires. Unlike Amy, Rachel can

never choose against reason. We have already seen, however, that the ability to make radical choices does not rule out making rational choices. That is, an agent could have done otherwise and still have chosen the rational alternative because it is the rational alternative. Thus, she could have done otherwise and have an explanation for her behaviour. Amy is like that. But Rachel is supposed to be different from Amy. We are to view Rachel as an ultra-rational being whose only serious options are rational options. We need to consider two questions about Rachel in order to understand whether we need to introduce the second clause for praiseworthiness. First, is Wolf's portrayal of Rachel coherent? And second, is Rachel praiseworthy for her behaviour?

Wolf's ideal moral agent, like Rachel, always acts in accordance with right reason. According to her description, this "agent's vision may be so clear that she cannot help seeing whihe action is the right thing, and her virtue may be so sure that, knowing which action is right, she cannot help performing it" (p. 81). In a different example, Gail, the gift-giver, sees a gift for her friend and buys it, not for any occassion, but because she "couldn't resist." As long as Gail did the right thing for the right reasons, Wolf maintains that she deserves praise for her generous act "whether she literally could have resisted performing it or not" (p. 84). With the examples of Gail and Rachel, we can get an intuitive grasp on the kind of agent Wolf is referring to. But it is difficult to understand how these

agents differ from Amy. That is, what is it about Gail and Rachel that makes them unable to choose anything but a rational alternative. The only options that they have, according to Wolf, are rational options because they cannot take seriously thoughts of irrational alternatives. It is only in Wolf's sense of "could have done otherwise" that it makes sense to say that these agents could not have done otherwise. But in Wolf's sense, "could have done otherwise" means "could have chosen an irrational alternative instead of a rational one." Thus, that interpretation of what it means to be able to do otherwise does not take us any closer to understanding how it could be the case that some agent can choose only from rational alternatives. For, by definition, if Rachel cannot do otherwise she cannot choose an irrational alternative. The question is, what could it possibly mean to say that the only options that an agent has are rational options? And if this is the case, is this agent really praiseworthy for her behaviour?

Intuitively, it does seem as if Rachel and Gail and other virtuous agents who always choose in accordance with reason deserve praise. But in these cases, "couldn't resist" and "had no choice" are really just idiomatic expressions whose philosophical clout is questionable. For if Gail <u>literally</u> had no choice but to buy the gift for her friend, I wonder whether we would think her praiseworthy. Surely we would prefer, upon saying to Gail, "Gail, you shouldn't have," to hear something along the lines of "I wanted to" or "I thought it would make you

happy" instead of "I had no choice." What makes the surprise gift-giver worthy of praise is precisely the fact that she did not have to buy the gift for her friend. She chose to buy the gift for her friend.

We are to understand that Rachel's grip of "drowning child + I'm its only hope = I better save it" is so clear that it unquestionably results in her saving the child, in much the same way that her understanding that 2 + 2 = 4 is so clear that it unquestionably results in her believing that 2 + 2 = 4. However, in order for the analogy to go through, we need to make some very strong assumptions about moral motivation. For, on the face of it, beliefs do not have the same voluntary aspect about them that actions do. For one can rationally see what the right thing to do is, and form the belief that e.g. saving the child is the right thing to do, but that does not entail, prima facie, that one will do it. The most that it seems to entail is that one will believe that one ought to do it.

Again, it seems as if the intuitive grip that we have about an agent who "couldn't resist" but who nonetheless deserves praise is not someone who <u>literally</u> could not have done otherwise. That is, Gail and Rachel are ideal moral agents who are so virtuous that we cannot imagine that they <u>would</u> ever do anything but the right thing if they knew what the right thing was. But it is simply a mistake to identify their always

Robert Stalnaker's comments on an earlier draft helped me clarify this point.

choosing in accordance with reason with their being <u>unable</u> to do anything but act in accordance with it. That is, it does not seem as if they <u>could</u> never choose anything but the right thing, but just that they <u>would</u> never choose anything but the right thing.

Our paradigm examples of people who could not have done otherwise, Kleo and Helen, would not have done otherwise even if they had so chosen. It is not just that Kleo, being a kleptomaniac, would most likely steal if the appropriate opportunity arose, it is that she could not help but steal because she is in the grip of a psychological compulsion. Even if Kleo recognized that it was wrong to steal, she might very likely be unable to help herself. Helen has been hypnotized into being kind, considerate and generous to everyone around her. So it is not just that she very predictably would behave as she does, it is not open to her to behave in any other way given the hypnotic suggestion.

If Gail and Helen are in the grip of reason the way that Kleo and Helen are in the grip of their psychological compulsions, they are compulsively rational and good. And if they are compulsively rational and good, their only options are rational ones. But the compulsively good and rational agent is driven to act as she does by some kind of neurosis. Consider Felix Unger's compulsive fanaticism about neatness. That is not a virtue. We are disinclined to praise agents for compulsive behaviour. Thus, this cannot be what Wolf has in mind for Gail

and Rachel. They were not psychologically compelled like Kleo and Helen. In addition, they were clearly not physically constrained like Philip, nor coerced like Carl to act as they did.

The only coherent rendering of Gail and Rachel that preserves the intuition that they deserve praise is that they always make the rational choice when they know what the rational choice is. But this does not set them apart from Amy, who might very well do the very same thing. As we have seen, always making the rational choice cannot coherently entail that rational options are the only options that an agent has. A praiseworthy agent who always acts in accordance with reason could have done otherwise. She would have done otherwise if she had so chosen. And while it is true that she very likely would not choose otherwise, it is not true that she could not.

The initial motivation for The Choice View came from its ability to account for our disinclination to praise or blame in certain problem cases. The most plausible way of interpreting The Choice View has "X could of done otherwise" meaning "X would have done otherwise if X had so chosen." Given this interpretation, the parallel between the problem cases and Rachel and Gail does not hold. For on this more plausible reading of The Choice View Gail and Rachel are reponsible. That is why we have the intuition that they are praiseworthy for their behaviour. If it were true that they could not have done otherwise in that they were hypnotized, or psychologically

compelled to act in the way that they did, we would be disinclined to think them responsible in any case.

Since there is a very relevant sense in which the praiseworthy agent who does the right thing for the right reasons must have been able to do otherwise, the second clause to account for praiseworthiness is superfluous. The Choice View's requirement that a responsible agent could have done otherwise accounts for blameworthiness and for praiseworthiness.

6 CONCLUSION

The preceding discussion indicates why I think that The Reason View, although it appears to be well-motivated, does not offer a better interpretation of the slogan "responsibility requires autonomy" than The Choice View does. The description of Amy in the drowning child example gave us reason to think that, on one interpretation of what it is to have the ability to do otherwise, it is not a desirable ability to have after all. But we have seen that while, on every interpretation, having this ability does entail that one have the option to choose against reason even when a rational basis for choice exists, it need not rule out rational explanations for one's action. Autonomous agents could have done otherwise even though they regard reason as primary.

With this support for The Reason View undermined, we turned our attention to the other motivation for The Reason View. It appeared compelling because it seems that very virtuous

agents, like Rachel and Gail, who in some sense seem to have no choice but to act on reason, should be praiseworthy for their behaviour. As we have seen, however, there is no clear sense in which it makes sense to say that they could not have done otherwise in a way that would incline us to regard them as praiseworthy. The very idea of such agents turned out to be incoherent, except in a very idiomatic sense. And a closer look at what we mean with such idiomatic expressions as "I couldn't resist" and "I had no choice" in the case of the praiseworthy agent revealed that these agents do have options other than the rational one. For it is not that Helen and Gail could not do otherwise in the way that Kleo and Helen could not, but that given their virtuous characters and their understanding of right reason, they consistently choose to do the right thing. Unlike Kleo and Helen, they would have done otherwise if they had so chosen. As a result, The Reason View's second clause introducing a new necessary condition for praiseworthiness is superfluous. On the appropriate reading of The Choice View, it already accounts for the type of cases that motivate that addition.

I want to suggest, then, that The Choice View is, in fact, the correct interpretation of the slogan "responsibility requires autonomy." Autonomous agents are autonomous because they have the ability to make choices. Having the ability to make choices does not mean that the autonomous agent cannot find reasons persuasive in favour of one course of action over others and be influenced accordingly. Agents like Rachel and Gail are

praiseworthy for their behaviour not because their understanding of reasons and their good characters make it the case that they cannot do otherwise in any relevant sense, but because their grip of reasons and their good characters make it the case the they will always choose in accordance with reason.

I would like to suggest also that The Reason View's introduction of the ability to use reason into the conditions for responsibility and autonomy also seems compelling because there does seem to be something right in thinking that our ability to engage our reason in our actions is significant for responsibility. Part of the problem for Kleo and Helen is that their respective psychological conditions render them incapable of engaging their reason. Even if they recognize that there are options, they are unable to follow through on anything but one course of action. But while the ability to act in accordance with reason is significant for responsibility, it is by no means sufficient for it. It is significant because to the extent that we have options, our ability to use reason helps us to recognize what they are. Insofar as we can deliberate over alternatives, we are different from non-rational animals. But it will not do to divorce the ability to act in accordance with reason from the ability to do otherwise. The primary requirement for coherent ascriptions of praise and blame is that the agent could have done otherwise, in that she would have done otherwise if she had so She must have a choice. Kleo and Helen might be able to chosen. deliberate over alternatives. But they are not responsible

because they would not have done otherwise even if they had so chosen. Were they in a position to engage their reason in their actions, they would have been able to do otherwise, and so would have been responsible. Rachel and Gail are in a position to engage their reason. They recognize what the alternatives are and choose the right alternative for their circumstances. But it is not as if they could not do otherwise. They could, but they choose not to do so. For this reason, they are responsible.

My main concern has been to show that The Choice View is the correct interpretation of the slogan "responsibility requires autonomy" and that The Reason View, although it introduces important considerations, does not need its second clause.

Although it may be the case that our reasons do result in action, I have suggested that the claim that a responsible agent could have no choice but to do the right thing is incoherent. Reason leaves it open that we may choose to follow one course of action rather than another because it enables us to recognize that there are alternatives.

BIBLIOGRAPHY

- DAVIDSON, DONALD. "Freedom to Act" in Essays on Actions and Events. Oxford: Clarendon Press, 1980. pp. 63-81.
- FRANKFURT, HARRY. "Alternate Possibilities and Moral Responsibility" in Philosophical Essays, the importance of what we care about. Cambridge: Cambridge University Press, 1988. pp. 1-10.
- WOLF, SUSAN. Freedom Within Reason. Oxford: Oxford University Press, 1990.