AN ARGUMENT FOR CONSEQUENTIALISM

Walter Sinnott-Armstrong
Dartmouth College

The most common way to choose among moral theories is to test how well they cohere with our intuitions or considered judgments about what is morally right and wrong, about the nature or ideal of a person, and about the purpose(s) of morality. Another kind of intuition is often overlooked. We also have intuitions about principles of practical and moral reasoning, such as those captured by deontic logic. In order to be principles of reasoning rather than substance, these principles must be consistent with all substantive moral theories. But consistency is not enough. We want the deeper kind of coherence that comes only with explanation. A moral theory that simply reports the principles behind common moral reasoning but cannot explain why these principles are so common or so plausible is inferior in this respect to another moral theory which not only includes the principles but also explains why they are true. Why is the explanatory theory better? Because we want a moral theory to help us understand moral reasoning, and such understanding is gained only when our principles are explained. Without such understanding, our intuitions do not seem justified, and we cannot know whether or how to extend our principles to new situations. These are reasons to prefer a moral theory that explains our principles of moral reasoning.

This preference for explanation provides a new method for choosing among competing moral theories. I will illustrate and apply this method in this paper. First, I will argue that a certain principle holds for reasons for action in general and for moral reasons in particular. Next, I will argue that this principle of moral reasoning cannot be explained by deontological moral theories or by traditional forms of consequentialism. Finally, I will outline a new kind of consequentialism that provides a natural explanation of this principle of moral reasoning. Its explanatory power is a reason to prefer this new version of consequentialism.
1. General Substitutability

My principle can be introduced by a non-moral example from everyday life. I have a cavity, and cavities become painful when they are not filled, so I have a reason to get my cavity filled. I can’t get my cavity filled without going to a dentist, so I have a reason to go to a dentist. Arguments with this form are very common.²

They are also incomplete. Suppose that no dentist will fill my cavity without an appointment, and I don’t have an appointment. Then I don’t have any reason to go to a dentist. Why not? Going to a dentist would be a waste of time, since it would not enable me get my cavity filled. Of course, going to a dentist is never sufficient by itself to get my cavity filled, since I also must stay there long enough, promise payment, etc. Nonetheless, going to a dentist often enables me to get my cavity filled in the sense that, if I go to the dentist, I can do other things which will together be sufficient for me to get my cavity filled. In general, I will say that doing Y enables an agent to do X if and only if Y is part of a larger course of action that is sufficient for the agent to do X, and the agent can do the other acts that make up what is sufficient for X. Now, when going to a dentist both enables me to get my cavity filled and also is necessary to get my cavity filled, then a reason to get my cavity filled does generate a reason to go to a dentist.

It is crucial not to overestimate this claim. Although I have some reason to go to a dentist, this reason still might be overridden. I might have an overriding reason to leave for Australia. Also, my reason to go to a dentist and my reason to get my cavity filled need not be distinct in any way that would allow me to add the force of two reasons. Even if these reasons are the same, I still have a reason to go to a dentist. That is all the above argument claims.

It is also important that this argument does not require logical impossibility. It is logically possible for me to get my cavity filled without going to a dentist. My wife might know how and be willing to fill my cavity, but she doesn’t and isn’t. So my particular situation makes it causally impossible for me to get my cavity filled except by going to a dentist. That kind of causal impossibility is enough for the above argument to be valid.

The most general principle that warrants arguments of this form is this:

(GS) If there is a reason for A to do X, and if A cannot do X without doing Y, and if doing Y will enable A to do X, then there is a reason for A to do Y.

I will call this ‘the general principle of substitutability’ (or just ‘general substitutability’), since it specifies conditions when ‘Y’ can be substituted for ‘X’ within the scope of the operator ‘there is a reason’. I will also call Y a ‘necessary enabler’ of X. Some more conditions might be needed,³ but some
principle along these lines must be accepted in order to explain why the above argument is valid.

It might seem that a stronger principle is true. Suppose I get another cavity, and two dentists are equally good and available. Going to the first dentist will enable me to get my cavity filled, but it is not necessary, since the other dentist is also available. Do I have a reason to go to the first dentist? This question is hard to answer because reasons are relative to a set of alternatives. If the question is whether I have a reason to go to the first dentist as opposed to going to no dentist at all, the answer is that I do have a reason of this kind. However, if the question is whether I have a reason to go to the first dentist as opposed to the other dentist, then I do not have any reason of this kind (since the dentists are equal). It is this latter, stronger kind of reason that is most accurately called a reason to go to the first dentist, and I will discuss this kind of reason. In general, I will say that there is a reason to do X only if there is a reason to do X as opposed to all relevant alternatives to X. When we consider reasons of this kind, (GS) is the strongest principle that is plausible.4

(GS) can be supported by several arguments. First, (GS) is confirmed when we can apply it again to the same example. If I have a reason to go to a dentist, and I can't go to a dentist without getting in my car, and getting in my car will enable me to go to a dentist, then I have a reason to get in my car. It might seem more natural to say, 'You ought to get in your car' or just 'Get in the car. We have to get going.' But what justifies these utterances is that I have a reason to get in the car.

More confirmation comes from reasons with different sources. I have an aesthetic reason to play a certain note at a certain time, for I am playing a piece of music that will be ugly if I don't play that note then. I can't play the right note without moving my fingers just so. Moving my fingers just so will enable me to play the right note, since I do have the right instrument in my hands. Therefore, I have a reason to move my fingers just so. Similarly, if I have a religious reason to go to church, and getting out of my chair is necessary and will enable me to go to church, I have a reason to get out of my chair. It need not be imprudent or ugly or immoral for me to stay in my chair, but religious truth or belief can still give me a reason to get out of my chair.

General substitutability also works for negative reasons. For example, I have a reason not to work all day, because I have a reason to get my cavity filled today, I cannot do so if I work all day, and not working all day would enable me to get my cavity filled. This argument is valid, so the above principle remains true when we replace 'Y' with '¬Y'.

The principle also holds when we replace 'X' with '¬X'. If I have a reason not to anger my boss, and not leaving early is necessary and will enable me to avoid angering my boss, then I have a reason not to leave early (even

An Argument for Consequentialism / 401
if this reason is overridden). Similarly, if I have a reason not to anger my boss, going to my boss's party is necessary and will enable me to avoid angering her, then I have a reason to go to her party. These arguments remain valid throughout these variations in the source and structure of reasons, and this confirms the general principle of substitutability.

Additional justification comes from a theory of reasons for action. In my view, a reason for an action is a fact about that action. Others deny that there is a reason when the relevant fact is not known. But suppose that both the high road and the low road lead to my house. A bridge is out on the low road, and I do not know this. I might not have any reason not to take the low road, but there still is a reason for me not to take the low road. That is why someone who knows about the bridge can tell me that I ought to take the high road. This suggests that there is a reason for an action when there is the right kind of fact about the action.

But what is the right kind? What makes certain facts reasons? It is hard to say anything very informative here, but one formula does provide some guidance for our intuitions. A fact is a reason when it has rational force in the sense that it can have an effect on what is rational or irrational. Some facts are reasons to do what there would otherwise be no reason to do. They make rationally neutral acts rationally required. I will identify reasons by another effect: a fact is a reason for action if it can make rational an action that would otherwise be irrational.5 For example, it would be irrational to pay a dentist to drill a hole in your tooth if you had no reason to do so, but this act is not irrational when it is necessary and enables you to avoid greater future pain. This ability to make otherwise irrational acts rational is what makes this fact a reason to get your cavity filled.

A reason for an act does not always make the act rational, since the reason is sometimes overridden. Nonetheless, even if a reason is overridden in a particular case, it is still a reason if it has enough force to make an irrational act rational in some other case where the opposing reasons are weaker. If this kind of fact has this force in other cases, it also needs to be weighed along with other reasons in order to determine whether the present act is rational. This force and this need are captured by defining a reason as a fact that can affect rationality in some cases.

This use of ‘can’ might seem to make every fact a reason. If the fact that the wall is blue can make it rational to repaint the wall, this is a reason to paint the wall even if I like the color blue. However, all this really shows is that it is not the color of the wall that makes it rational to paint the wall. The color of the wall cannot be the real reason to paint the wall because it would still be irrational to paint the wall if I did not dislike or suffer some harm because of its color or condition. Thus, there are many facts that do not count as reasons on this theory.

This partial theory of reasons for action supports the general principle of
substitutability. When there is a reason for me to get my cavity filled, the fact that going to a dentist is necessary and enables me to get my cavity filled makes it rational to go to a dentist even if this act would otherwise be irrational, so that fact is a reason for me to go to a dentist. Thus, the ability to make irrational acts rational transfers from acts to their necessary enablers, just as general substitutability claims.

2. Moral Substitutability

Since general substitutability works for other kinds of reasons for action, we would need a strong argument to deny that it holds also for moral reasons. If moral reasons obeyed different principles, it would be hard to understand why moral reasons are also called 'reasons' and how moral reasons interact with other reasons when they apply to the same action. Nonetheless, this extension has been denied, so we have to look at moral reasons carefully.

I have a moral reason to feed my child tonight, both because I promised my wife to do so, and also because of my special relation to my child along with the fact that she will go hungry if I don't feed her. I can't feed my child tonight without going home soon, and going home soon will enable me to feed her tonight. Therefore, there is a moral reason for me to go home soon. It need not be imprudent or ugly or sacrilegious or illegal for me not to feed her, but the requirements of morality give me a moral reason to feed her. This argument assumes a special case of substitutability:

(MS) If there is a moral reason for A to do X, and if A cannot do X without doing Y, and if doing Y will enable A to do X, then there is a moral reason for A to do Y.

I will call this 'the principle of moral substitutability', or just 'moral substitutability'.

This principle is confirmed by moral reasons with negative structures. I have a moral reason to help a friend this afternoon. I cannot do so if I play golf this afternoon. Not playing golf this afternoon will enable me to help my friend. So I have a moral reason not to play golf this afternoon. Similarly, I have a moral reason not to endanger other drivers (beyond acceptable limits). I can't drink too much before I drive without endangering other drivers. Not drinking too much will enable me to avoid endangering other drivers. Therefore, I have a moral reason not to drink too much before I drive. The validity of such varied arguments confirms moral substitutability.

We can also extend the above theory of reasons. Since a reason for action is a fact that can affect the rationality of an act, a moral reason is a fact that can affect the morality of an act, either by making an otherwise morally neutral act morally good or by making an otherwise immoral act moral. As above, a moral reason need not be strong enough to make its act moral in
every case as long as it has that ability in some cases. For example, if I promised to meet a needy student later this afternoon, it is immoral for me to go home now if I have no morally relevant reason to go. Nonetheless, it is not immoral for me to go home now if this is necessary and enables me to feed my child when I have a moral reason to feed her. Thus, this fact about going home now can make an otherwise immoral act moral, so this fact is a moral reason. This supports moral substitutability. When there is a moral reason for me to feed my child, and going home now is necessary and enables me to feed my child, this fact makes it moral for me to go home now even in a situation where this would otherwise be immoral, so this fact is a moral reason for me to go home now. Thus, the ability to make immoral acts moral transfers from acts to their necessary enablers, just as moral substitutability claims.

Despite these arguments for moral substitutability, critics will raise several objections. I will consider only three kinds of objections, but they set the basic patterns for most others.

The first kind of objection claims that some necessary enablers of what I have a moral reason to do seem morally neutral. For example, I promised, so I have a moral reason to cook dinner. I cannot cook dinner without moving some air molecules, and moving some air molecules does enable me to cook dinner. Moral substitutability warrants the conclusion that there is a moral reason for me to move some air molecules. This seems at least odd.

My response is that this conclusion is still true. Its oddness can be explained by pragmatic principles. My moral reason to move some air molecules is very weak and general, since I can move air molecules in many different ways, but only some of these count as cooking dinner. Thus, when one can make the stronger and more specific judgment that I have a moral reason to cook dinner, it is misleading to say only that there is a moral reason to move some air molecules. This weaker judgment suggests that this is all I have a moral reason to do, since, if I have a moral reason to do more, why not say so? Furthermore, if the purpose of the conversation is to direct my action, then it is irrelevant to say that I have a moral reason to move some air molecules, since this judgment fails to tell me how to move them. This judgment also fails to help others determine whether I have done all that I have a moral reason to do, or whether I am blameworthy. The failure to achieve such purposes explains why it is odd to say that I have a moral reason to move some air molecules. But this explanation is purely pragmatic, so it can still be true that I have a moral reason to move some air molecules. And it is true, because my act will move some air molecules if I do what I have a moral reason to do. The truth of this consequence is all that is needed to defend moral substitutability.

The second objection is that some necessary enablers of what I have a moral reason to do are morally wrong. Philippa Foot gives an example:
Suppose, for instance, that some person has an obligation to support a dependent relative, an aged parent perhaps. Then it may be that he ought to take a job to get some money. ...But what if the only means of getting money is by killing someone? "It is not the case that the son or daughter ought to kill to get the money."7

If we change 'obligation' and 'ought' into 'moral reason', Foot seems to claim that the child does not have any moral reason to kill, even though the child has a moral reason to get the money, the child cannot get the money without killing, and killing would enable the child to get the money. This would refute moral substitutability.

However, the child does have some moral reason to kill. This moral reason is vastly overridden by the moral reason not to kill, but, if the child kills, it will not be for no reason at all or for a selfish reason. Moreover, the fact that an act is necessary and enables the child to support the parents can make an otherwise immoral act moral in other cases, even if it is not strong enough to remove the immorality in this case. It still would be odd to say only that the child has a moral reason to kill, but what makes this utterance odd is just that it leaves out some very important information: the reason is vastly overridden. Nonetheless, it is still true that the child has some moral reason to kill, so moral substitutability stands.

Foot does make another suggestion that deserves a response. In place of moral substitutability, she says in effect that there is a moral reason to do what is a necessary means of doing what there is a moral reason to do. This weaker principle is supposed explain the validity of arguments like those above. Foot also claims that this principle does not yield any moral reason to kill in her example, because a means must be possible, but here killing is not a 'moral possibility', because it is wrong. I am not convinced, however, that what is morally impossible in this way cannot be a means. More importantly, Foot's principle is too weak to explain the validity of many common arguments. Getting in my car does not seem to be a means of getting my cavity filled, since I do not get my cavity filled by (means of) getting in my car. Other examples are even clearer. Suppose I have a moral reason to go home, but I cannot go home without waking the dog that is sleeping outside my office. Waking the dog does enable me to go home in my sense that it is part of a larger course of action that is sufficient for me to go home, where I can do the rest of what is sufficient. Therefore, I have a moral reason to wake the dog (even if it is odd to describe the reason in this way). However, waking the dog is not a means of going home. Thus, Foot's principle cannot explain why this argument is valid. We need a principle as strong as moral substitutability in order to explain the validity of all of the arguments that are valid.

A final objection asks which reasons are moral. Even if I have a moral reason to feed my child, and even if this gives me a reason to go home, a
critic still might deny that my reason to go home is *moral* in nature. However, my reason to go home seems moral, both because its source is a moral reason, and also because its role is to affect what is moral or immoral. Furthermore, if it is not a moral reason, what kind of reason is it? My reason to go home is derived by substitutability from a moral reason (my reason to feed my child). This chain of reasons need not have any source apart from morality, and, if it does not, my reason to go home is not aesthetic or religious or prudential or legal. Opponents might respond that my reason to go home is a ‘practical’ or ‘instrumental’ reason. However, this is just another way of saying that it is derived by substitutability, and this does not show that it is not also a moral reason. Instrumental moral reasons must still count as moral reasons, since otherwise we could not explain how they differ from instrumental religious reasons, instrumental aesthetic reasons, and so on. My reason to go home is stronger and different in character if it derives from a moral reason to feed my child than if it derives from a prudential reason to take a nap. The differences between my reasons to go home in these two cases would be hidden if both reasons were described solely as ‘instrumental’ without recognizing that the first is moral and the second is prudential. That is why we need to count reasons derived from moral reasons via substitutability as moral reasons themselves. Finally, we can also argue more positively for the moral nature of my reason to go home. This reason meets every necessary condition in every plausible definition of morality. Morality is usually defined by its form, content, or force. Regarding form, my reason to go home is universalizable, since anyone in similar circumstances would also have a reason to go home. My reason to go home also has the content essential to morality, since it concerns social relations and harm to others. Even if morality must be supreme in force (which I doubt), my reason to go home is just as likely to be supreme as my reason to feed my child, since my reason to go home is not overridden in this situation unless my reason to feed my child is also overridden. This leaves no reason to deny that my reason to go home is moral in nature.

There are, of course, many other possible objections to moral substitutability, and I cannot answer all of them here. Nonetheless, we already have several arguments for moral substitutability. It is confirmed by numerous and varied examples. It coheres well with a general theory of reasons for action. And the main objections have been met. So I conclude that the principle of moral substitutability is true.

3. Why and How to Explain Substitutability

If moral substitutability is so obvious, why does it need to be explained? The answer lies in its unusual features. Moral substitutability relates a moral
operator ('there is a moral reason') to a non-moral operator ('can'). Moreover, the non-moral operator ('can') represents contingent facts and not just necessary relations, such as logical impossibility or act identity. Thus, moral substitutability takes us from a moral judgment about one action to a moral judgment about a different action that is only contingently related.

Similar inferences are not allowed for many other properties. For example, even if I cannot chop vegetables without taking a knife out of a drawer, and even if taking out a knife does enable me to chop vegetables, my taking out a knife still might be quick when my chopping vegetables is slow. Thus, quickness and many other properties of acts do not transfer to necessary enablers of those acts. This makes it odd that moral and other reasons for acts do transfer to necessary enablers of those acts. This oddity creates the need to explain moral substitutability.

How can a moral theory explain moral substitutability? The most direct way is if moral substitutability follows from the substantive principles in the moral theory. Moral theories present basic moral principles about what morally ought or ought not to be done, or about what is morally right or wrong, etc. Regardless of terminology, these substantive principles tell us in effect what moral reasons there are. For example, if the rule 'Keep your promises' or the principle 'You ought to keep your promises' is basic to a moral theory, then the fact that an act is an act of keeping a promise is a moral reason to do that act according to that theory. In such basic rules or principles, each moral theory specifies which properties of acts constitute moral reasons to do or not to do acts with those properties.

Moral theories can then explain moral substitutability by picking out the right kinds of properties as moral reasons. Suppose a moral theory implies that an agent has a moral reason to do act R because act R has a property P. Now suppose that act N is a necessary enabler for act R. Moral substitutability implies that there is also a moral reason to do act N, but the question is: why? If act N also has property P, this makes it clear why there is also a reason to do act N. However, if act N does not have property P, this moral theory does not explain why the agent has any reason to do act N. Schematically,

\[
\begin{align*}
P' & \quad P \\
& \quad \mid \quad | \\
N & \quad \leftarrow R
\end{align*}
\]

Thus, a moral theory can explain why moral substitutability holds if its basic substantive principles imply that all moral reasons are or are due to properties such that, if one act has such a property, then any act which is a necessary enabler for the first act also has such a property. Moral substitutability then follows from the substantive principles of the moral theory.

It might seem that this kind of explanation is not needed, because another
explanation is readily available. Moral substitutability can be derived from general substitutability. Doesn’t that explain moral substitutability? Not really. General and moral substitutability set abstract constraints on reasons and moral reasons. We still need to know whether and why these constraints are actually met by the moral reasons picked out by each moral theory. Suppose a substantive moral theory implies that there are moral reasons to do acts of types A, B, C, and D, but no more. Moral substitutability, however, implies that there are also moral reasons to do acts of type E, F, G, and H, since these are necessary enablers for A-D. The substantive moral theory itself cannot explain why there are these additional moral reasons to do E-H. General substitutability still implies moral substitutability, but neither of these principles stands in any explanatory relation to the substantive moral theory itself. We could just add the principle of moral substitutability to the original theory, and the new, extended theory would imply moral reasons to do acts of type E-H. However, the original, substantive principles would still give us no understanding of why the extension is needed. The substitutability principles were just tacked on in an ad hoc fashion, so they are still not connected in any significant way to the substantive part of the theory. This lack of coherence and explanatory value makes this extended theory inferior to another moral theory whose substantive moral principles imply and thus explain moral substitutability.

The same burden arises outside morality. General substitutability applies to prudential reasons, aesthetic reasons, etc., so a substantive theory of prudence, aesthetics, etc., also needs to specify the content of its reasons in such a way that we can understand why those reasons obey the constraints of general substitutability. Although this explanatory burden is general, this does not make it any easier to carry. Each theory in each area will lack coherence unless its particular substantive principles bear some explanatory relation to the relevant principles of substitutability. This general requirement has important implications in all of these areas, but I will focus on moral reasons.

4. Kinds of Moral Reasons

In order to determine which moral theories can explain moral substitutability, we need to distinguish two kinds of moral reasons and theories: consequential and deontological. These terms are used in many ways, but one particular distinction will serve my purposes.

A moral reason to do an act is *consequential* if and only if the reason depends *only* on the consequences of either doing the act or not doing the act. For example, a moral reason not to hit someone is that this will hurt her or him. A moral reason to turn your car to the left might be that, if you
do not do so, you will run over and kill someone. A moral reason to feed a starving child is that the child will lose important mental or physical abilities if you do not feed it. All such reasons are consequential reasons.

All other moral reasons are non-consequential. Thus, a moral reason to do an act is non-consequential if and only if the reason depends even partly on some property that the act has independently of its consequences. For example, an act can be a lie regardless of what happens as a result of the lie (since some lies are not believed), and some moral theories claim that that property of being a lie provides a moral reason not to tell a lie regardless of the consequences of this lie. Similarly, the fact that an act fulfills a promise is often seen as a moral reason to do the act, even though the act has that property of fulfilling a promise independently of its consequences. All such moral reasons are non-consequential. In order to avoid so many negations, I will also call them 'deontological'.

This distinction would not make sense if we did not restrict the notion of consequences. If I promise to mow the lawn, then one consequence of my mowing might seem to be that my promise is fulfilled. One way to avoid this problem is to specify that the consequences of an act must be distinct from the act itself. My act of fulfilling my promise and my act of mowing are not distinct, because they are done by the same bodily movements. Therefore, my fulfilling my promise is not a consequence of my mowing. A consequence of an act need not be later in time than the act, since causation can be simultaneous, but the consequence must at least be different from the act.

Even with this clarification, it is still hard to classify some moral reasons as consequential or deontological, but I will stick to examples that are clear.

In accordance with this distinction between kinds of moral reasons, I can now distinguish different kinds of moral theories. I will say that a moral theory is consequentialist if and only if it implies that all basic moral reasons are consequential. A moral theory is then non-consequentialist or deontological if it includes any basic moral reasons which are not consequential.

5. Against Deontology

So defined, the class of deontological moral theories is very large and diverse. This makes it hard to say anything in general about it. Nonetheless, I will argue that no deontological moral theory can explain why moral substitutability holds. My argument applies to all deontological theories because it depends only on what is common to them all, namely, the claim that some basic moral reasons are not consequential. Some deontological theories allow very many weighty moral reasons that are consequential, and these theories might be able to explain why moral substitutability holds for some of their moral reasons: the consequential ones. But even these theories
cannot explain why moral substitutability holds for all moral reasons, including the non-consequential reasons that make the theory deontological. The failure of deontological moral theories to explain moral substitutability in the very cases that make them deontological is a reason to reject all deontological moral theories.

I cannot discuss every deontological moral theory, so I will discuss only a few paradigm examples and show why they cannot explain moral substitutability. After this, I will argue that similar problems are bound to arise for all other deontological theories by their very nature.

The simplest deontological theory is the pluralistic intuitionism of Prichard and Ross. Ross writes that, when someone promises to do something, 'This we consider obligatory in its own nature, just because it is a fulfillment of a promise, and not because of its consequences.'12 Such deontologists claim in effect that, if I promise to mow the grass, there is a moral reason for me to mow the grass, and this moral reason is constituted by the fact that mowing the grass fulfills my promise. This reason exists regardless of the consequences of mowing the grass, even though it might be overridden by certain bad consequences. However, if this is why I have a moral reason to mow the grass, then, even if I cannot mow the grass without starting my mower, and starting the mower would enable me to mow the grass, it still would not follow that I have any moral reason to start my mower, since I did not promise to start my mower, and starting my mower does not fulfill my promise. Thus, a moral theory cannot explain moral substitutability if it claims that properties like this provide moral reasons.

Of course, this argument is too simple to be conclusive by itself, since deontologists will have many responses. The question is whether any response is adequate. I will argue that no response can meet the basic challenge.

A deontologist might respond that his moral theory includes not only the principle that there is a moral reason to keep one's promises but also another principle that there is a moral reason to do whatever is a necessary enabler for what there is a moral reason to do. This other principle just is the principle of moral substitutability, so, of course, I agree that it is true. However, the question is why it is true. This new principle is very different from the substantive principles in a deontological theory, so it cries out for an explanation. If a deontologist simply adds this new principle to the substantive principles in his theory, he has done nothing to explain why the new principle is true. It would be ad hoc to tack it on solely in order to yield moral reasons like the moral reason to start the mower. In order to explain or justify moral substitutability, a deontologist needs to show how this principle coheres in some deeper way with the substantive principles of the theory. That is what deontologists cannot do.

A second response is that I misdescribed the property that provides the moral reason. Deontologists might admit that the reason to mow the lawn
is not that this fulfills a promise, but they can claim instead that the moral reason to mow the lawn is that this is a necessary enabler for keeping a promise. They can then claim that there is a moral reason to start the mower, because starting the mower is also a necessary enabler for keeping my promise. Again, I agree that these reasons exist. But the question is why. This deontologist needs to explain why the moral reason has to be that the act is a necessary enabler for fulfilling a promise instead of just that the act does fulfill a promise. If there is no moral reason to keep a promise, it is hard to understand why there is any moral reason to do what is a necessary enabler for keeping a promise. Furthermore, deontologists claim that the crucial fact is not about consequences but directly about promises. My moral reason is supposed to arise from what I said before my act and not from consequences after my act. However, what I said was 'I promise to mow the grass'. I did not say, 'I promise to do what is a necessary enabler for mowing the grass.' Thus, I did not promise to do what is a necessary enabler for keeping the promise. What I promised was only to keep the promise. Because of this, deontologists who base moral reasons directly on promises cannot explain why there is not only a moral reason to do what I promised to do (mow the grass) but also a moral reason to do what I did not promise to do (start the mower).

Deontologists might try to defend the claim that moral reasons are based on promises by claiming that promise keeping is intrinsically good and there is a moral reason to do what is a necessary enabler of what is intrinsically good. However, this response runs into two problems. First, on this theory, the reason to keep a promise is a reason to do what is itself intrinsically good, but the reason to start the mower is not a reason to do what is intrinsically good. Since these reasons are so different, they are derived in different ways. This creates an incoherence or lack of unity which is avoided in other theories. Second, this response conflicts with a basic theme in deontological theories. If my promise keeping is intrinsically good, your promise keeping is just as intrinsically good. But then, if what gives me a moral reason to keep my promise is that I have a moral reason to do whatever is intrinsically good, I have just as much moral reason to do what is a necessary enabler for you to keep your promise. And, if my breaking my promise is a necessary enabler for two other people to keep their promises, then my moral reason to break my promise is stronger than my moral reason to keep it (other things being equal). This undermines the basic deontological claim that my reasons derive in a special way from my promises. So this response explains moral substitutability at the expense of giving up deontology.

A fourth possible response is that any reason to mow the grass is also a reason to start my mower because starting my mower is part of mowing the grass. However, starting my mower is not part of mowing the grass, because I can start my mower without cutting any grass. I might start my mower hours
in advance and never get around to cutting any grass. Suppose I start the mower then go inside and watch television. My wife comes in and asks, ‘Have you started to mow the lawn?’, so I answer, ‘Yes. I’ve done part of it. I’ll finish it later.’ This is not only misleading but false. Furthermore, mowing the grass can have other necessary conditions, such as buying a mower or leaving my chair, which are not parts of mowing the grass by any stretch of the imagination.

Finally, deontologists might charge that my argument _begs the question_. It would beg the question to assume moral substitutability if this principle were inconsistent with deontological theories. However, my point is _not_ that moral substitutability is _inconsistent_ with deontology. It is not. Deontologists can consistently tack moral substitutability onto their theories. My point is only that deontologists cannot _explain_ why moral substitutability holds. It would still beg the question to assert moral substitutability without argument. However, I _did_ argue for moral substitutability, and my argument was independent of its implications for deontology. I even used examples of moral reasons that are typical of deontological theories. Deontologists still might complain that the failure of so many theories to explain moral substitutability casts new doubt on this principle. However, we normally should not reject a scientific observation just because our theory cannot explain it. Similarly, we normally should not reject an otherwise plausible moral judgment just because our favorite theory cannot explain why it is true. Otherwise, no inference to the best explanation could work. My argument simply extends this general explanatory burden to principles of moral reasoning and shows that deontological theories cannot carry that burden.

Even though this simple kind of deontological theory cannot explain moral substitutability, more complex deontological theories might seem to do better. One candidate is Kant, who accepts something like substitutability when he writes, ‘Whoever wills the end, so far as reason has decisive influence on his action, wills also the indispensably necessary means to it that lie in his power.’14 Despite this claim, however, Kant fails to explain moral substitutability. Kant says in effect that there is a moral reason to do an act when the maxim of not doing that act cannot be willed as a universal law without contradiction. My moral reason to keep my promise to mow the grass is then supposed to be that not keeping promises cannot be willed universally without contradiction. However, not starting my mower _can_ be willed universally without contradiction. I can even consistently and universally will not to start my mower when this is a necessary enabler for keeping a promise. The basic problem is that Kant repeatedly claims that his theory is purely _a priori_, but moral substitutability makes moral reasons depend on what is empirically possible. Kantians might try to avoid this problem by interpreting universalizability in terms of a less pure kind of possibility and ‘contradiction’. On one such interpretation, Kant claims it is contradictory to will universal promise
An Argument for Consequentialism / 413

breaking, because, if everyone always broke their promises, no promises would be trusted, so no promises could be made or, therefore, broken. There are several problems here, but the most relevant one is that people could still trust each other’s promises, including their promises to mow a lawn, even if nobody ever starts his mower when this is a necessary enabler for keeping a promise. This might happen, for example, if it is common practice to keep mowers running for long periods, so those to whom promises are made assume that it is not necessary to start one’s mower in order to mow the lawn. This shows that there is no contradiction of this kind in a universal will not to start my mower when this is a necessary enabler for keeping a promise. Thus, this interpretation of Kant also fails to explain why there is a moral reason to start the mower. Some defenders of Kant will insist that both of these interpretations fail to recognize that, for Kant, certain ends are required by reason, so rational people cannot universally will anything that conflicts with these ends. One problem here is to specify which and why particular ends have this special status. It is also not clear how these rational ends would conflict with universally not starting mowers. Thus, Kant can do no better than other deontologists at explaining why there is a moral reason to start my mower or why moral substitutability holds.

Of course, there are many other versions of deontology. I cannot discuss them all. Nonetheless, these examples suggest that it is the very nature of deontological reasons that makes deontological theories unable to explain moral substitutability. This comes out clearly if we start from the other side and ask which properties create the moral reasons that are derived by moral substitutability. What gives me a moral reason to start the mower is the consequences of starting the mower. Specifically, it has the consequence that I am able to mow the grass. This reason cannot derive from the same property as my moral reason to mow the lawn unless what gives me a moral reason to mow the lawn is its consequences. Thus, any non-consequentialist moral theory will have to posit two distinct kinds of moral reasons: one for starting the mower and another for mowing the grass. Once these kinds of reasons are separated, we need to understand the connection between them. But this connection cannot be explained by the substantive principles of the theory. That is why all deontological theories must lack the explanatory coherence which is a general test of adequacy for all theories.

I conclude that no deontological theory can adequately explain moral substitutability. I have not proven this, but I do challenge deontologists to give a better explanation of moral substitutability. Deontologists are very inventive, but I doubt that they can meet this challenge.
6. Against Sufficient Consequentialism

The failure of deontological theories to explain moral substitutability might seem to support the alternative: consequentialism. However, this does not follow unless consequentialists themselves can explain moral substitutability. And many traditional consequentialists cannot explain moral substitutability any better than deontologists can.

Many consequentialists claim that there is a moral reason to do what is sufficient to cause or maximize the good. For example, Smart formulates act utilitarianism as the claim that 'the only reason for performing an action A rather than an alternative action B is that doing A will make mankind (or, perhaps, all sentient beings) happier than will doing B.'\textsuperscript{15} We often say that an action makes something happen when the action is sufficient for the result, even if it is not necessary. This makes it natural to read Smart's principle as ascribing moral reasons on the basis of sufficient conditions. His theory is then a version of 'sufficient consequentialism', which in general is the claim that there are moral reasons for what causes or maximizes happiness or any other good.

Sufficient consequentialism fails when it is applied to examples. Suppose I can give a surprise birthday party for Susan, and this will make her happy. Telling the guests about the party is necessary and enables me to have the party. Moral substitutability then implies that, if there is a moral reason to have the party, there must also be a moral reason to tell the guests about it. However, the act of calling guests is not sufficient to make Susan happy, since she does not even know that I am calling the guests. In general, even if an act is sufficient for some good, the necessary enablers of the act need not be sufficient for the good. If they are not, and if moral reasons depended on what is sufficient for the good, then there would not be any moral reason to do the acts that are necessary enablers. That is why sufficient consequentialists cannot explain moral substitutability.

Defenders of sufficient consequentialism might respond that this argument depends on too narrow a view of what is sufficient. Telling the guests about the party really is sufficient \textit{in the circumstances}, since these include the facts that the guests will come and I will get Susan to the party and do whatever else is necessary for the party. They might even add that, if I know I will not get Susan to the party, then I do not have any reason to call the guests. But this is not right. I do have a moral reason to call the guests even if I will not get Susan to the party. What makes it seem otherwise is only that the gathering will be a disaster or at least a waste if the guest of honor, Susan, does not show up. This gives me a much stronger reason not to call the guests. But I still have \textit{some} moral reason to call the guests, because calling them is a necessary enabler for making Susan happy, and that fact can make an immoral act moral. This comes out clearly if we imagine a case where there
is no cost to calling the guests even if Susan does not show up. Maybe the guests would not mind at all coming to my house for a surprise party even if Susan does not show up. Then I have a reason to call the guests, because this keeps me in a position to make Susan happy if it turns out that she can come. It keeps my options open at no cost. Thus, I have a reason to call the guests, even when doing so is not sufficient in the circumstances for making Susan happy, as long as doing so is a necessary enabler for keeping Susan happy. This reason cannot be explained by sufficient consequentialism.

Another problem for sufficient consequentialism is that it implies that an agent has a moral reason to do whatever is a sufficient condition of what she has a moral reason to do, since the sufficient conditions of an act are sufficient for whatever the act is sufficient for. However, a single act can have many sufficient conditions, and the agent might be allowed to choose among them. Then the agent does not have a reason to do one particular sufficient condition, even if the agent does have a reason to do the disjunction. For example, suppose it would make Susan happy to have a surprise party at my house, but it would also make her just as happy to have a surprise party at Mark's house. Then I do not have a reason to have a surprise party at my house as opposed to Mark's house. I do not even have a reason to have the party at my house, since, as I said, I am concerned with reasons to do an act as opposed to all of its relevant alternatives, and having the party at Mark's house is a relevant alternative to having the party at my house.

Defenders of sufficient consequentialism might respond that, if I had the party at my house, I would not have acted for no reason. But all this shows is that I have a reason to have the party at my house as opposed to not having any party at all. I might also have a disjunctive moral reason to have a party either at my house or at Mark's house. But neither of these amounts to a reason to have a party at my house, even though having it at that particular location would be sufficient for Susan's happiness. Thus, sufficient conditions of the good do not always yield moral reasons. So sufficient consequentialism fails.16

7. For Necessary Enabler Consequentialism

All of this leads to necessary enabler consequentialism or NEC. NEC claims that all moral reasons for acts are provided by facts that the acts are necessary enablers for preventing harm or promoting good. All moral reasons on this theory are consequential reasons, but there are two kinds. Some moral reasons are prevention reasons, because they are facts that an act is a necessary enabler for preventing harm or loss. For example, if giving Alice food is necessary and enables me to prevent her from starving, then that fact is a moral reason to give her food. In this case, I would not cause her death even
if I let her starve, but other moral prevention reasons are reasons to avoid causing harm. For example, if turning my car to the left is necessary and enables me to avoid killing Bobby, that is a moral reason to turn my car to the left. The other kind of moral reason is a promotion reason. This kind of reason occurs when doing something is necessary and enables me to promote (or maximize) some good. For example, I have a moral reason to throw a surprise party for Susan if this is necessary and enables me to make her happy. Because of substitutability, these moral reasons for actions also yield moral reasons against contrary actions. There are then also moral reasons not to do what will cause harm or ensure a failure to prevent harm or to promote good.

What makes these facts moral reasons is that they can make an otherwise immoral act moral. If I have a moral reason to feed my child, then it might be immoral to give my only food to Alice, who is a stranger. But this would not be immoral if giving Alice food is necessary and enables me to prevent Alice from starving, as long as my child will not starve also. Similarly, it is normally immoral to lie to Susan, but a lie can be moral if it is necessary and enables me to keep my party for Susan a surprise, and if this is also necessary and enables me to make her happy. Thus, NEC fits nicely into the above theory of moral reasons.

NEC can provide a natural explanation of moral substitutability for both kinds of moral reasons. I have a prevention moral reason to give someone food when doing so is necessary and enables me to prevent that person from starving. Suppose that buying food is a necessary enabler for giving the person food, and getting in my car is a necessary enabler for buying food. Moral substitutability warrants the conclusion that I have a moral reason to get in my car. And this act of getting in my car does have the property of being a necessary enabler for preventing starvation. Thus, the necessary enabler has the same property that provided the moral reason to give the food in the first place. This explains why substitutability holds for moral prevention reasons. The other kind of moral reason covers necessary enablers for promoting good. In my example above, if a surprise party is a necessary enabler for making Susan happy, and letting people know about the party is a necessary enabler for having the party, then letting people know is a necessary enabler for making Susan happy. The very fact that provides a moral reason to have the party also provides a moral reason to let people know about it. Thus, NEC can explain why moral substitutability holds for every kind of moral reason that it includes. Similar explanations work for moral reasons not to do certain acts, and this explanatory power is a reason to favor NEC.

Of course, this should come as no surprise. NEC was intentionally structured so that it would explain moral substitutability. But this does not detract from its explanatory force. The point is that moral substitutability remains a mystery.
unless we restrict our substantive theory to moral reasons that obey moral substitutability by their very nature.

The crucial advantage of NEC lies in its unity. Other theories claim that my reason to do what I promised is just that this fulfills my promise or that promise keeping is intrinsically good. However, I did not promise to start the mower, and starting the mower is not intrinsically good. Thus, my reason to start the mower derives from a different property than my reason to keep my promise. In contrast, NEC makes my reasons to keep my promise, to mow the lawn, and to start the mower derive from the very same property: being a necessary enabler of preventing harm or promoting good. This makes NEC's explanation more coherent and better.

A critic might complain that NEC just postpones the problem, since NEC will eventually need to explain why certain things are good or bad, and some will be good or bad as means, but others will not. However, if what is good or bad intrinsically are states (such as pleasure and freedom or pain and death) rather than acts, then they are not the kind of thing that can be done, so there cannot be any question of a reason to do them. This makes it possible for all reasons for acts to have the same nature or derive from the same property. NEC will still have to explain why certain states are good or bad, but so will every other moral theory. The difference is that other theories will also have to explain why there are two kinds of reasons for acts and how these reasons are connected. This is what other theories cannot explain. This additional explanatory gap is avoided by the unified nature of reasons in NEC.18

8. Two Puzzles Solved

Another advantage of NEC is that it can solve two puzzles. The first example is modified from Parfit19 and concerns how we weigh moral reasons. Suppose that I can feed either Ann or Beth but not both. If I feed Beth, this will be sufficient to prevent her death. If I feed Ann, this will be sufficient to prevent the pain of hunger but not to prevent death, because Ann is not starving. This makes my reason to feed Beth seem stronger, since death is worse than the pain of hunger. However, suppose I know that, if I do not feed Beth, someone else certainly will, but, if I do not feed Ann, nobody else will. Now my reason to feed Ann is stronger. Why? Because feeding Beth now is not necessary to prevent any harm, but feeding Ann is a necessary enabler for preventing harm. This shows that the strength of a moral reason for an act depends on what the act is a necessary enabler for, rather than on what the act is sufficient for. This is exactly what NEC would predict.

NEC also solves a puzzle that bothers Feldman.20 Feldman has to decide between working in his garden and going to the dump. Working in his garden
is sufficient for 12 units of utility, and going to the dump is sufficient for only 8 units of utility. Each of these alternatives has a certain prerequisite. Feldman cannot work in the garden without gathering his garden tools, which is sufficient for -1 unit of utility, and he cannot go to the dump without loading his truck, which is sufficient for +1 unit of utility. (It is not clear why Feldman likes to load his truck!) Now sufficient consequentialism implies both that he has a reason to work in the garden (because this is sufficient for more utility than its alternative) and also that it is not true that he has a reason to gather his garden tools (because that act has an alternative that is sufficient for more utility). But this conflicts with substitutability. This result is easily avoided if we turn to NEC. Gathering his garden tools is a necessary enabler for more utility than is loading and starting his truck, since gathering his gardening tools is a necessary enabler for working in his garden (and we are assuming that all else is equal). That is why he has more reason to gather his garden tools, even though that act is sufficient for less utility than its alternative. Thus, NEC again coheres with our intuitions and saves us from the absurdities of sufficient consequentialism.

9. Conclusions

Let's review what has been accomplished. I argued first for principles of substitutability in general and moral substitutability in particular. I argued next that deontological moral theories and sufficient consequentialism cannot explain why moral substitutability holds. My positive conclusion was that necessary enabler consequentialism can easily explain why moral substitutability holds for all of its moral reasons. Since, as I also argued, a moral theory should not only describe but also explain principles of moral reasoning, my conclusions give us a reason to prefer some kind of necessary enabler consequentialism. The reason is that it provides a simple explanation of an obvious principle that cannot be explained as well by contrary theories.

Although I focussed on moral reasons, my conclusions can be extended to non-moral reasons for action. Reasons of prudence, aesthetics, etc., also obey general substitutability. My argument suggests that this cannot be explained by theories of prudence, aesthetics, etc., unless these theories identify these reasons with necessary enablers. The differences among various kinds of reasons for action can still be explained by differences among the goods or harms for which actions are necessary enablers. Prudential reasons concern good and harm only to oneself, and aesthetic reasons concern specifically aesthetic goods, such as beauty or aesthetic pleasure. Despite these differences, all reasons for action share a reference to necessary enablers, since they all obey general substitutability. In this way, necessary enabler consequentialism can be extended to yield a theory about reasons for action
in general, and this wider coherence provides even more support for necessary enabler consequentialism. 21 None of this proves that necessary enabler consequentialism is true or even that it is preferable overall to its competitors. Other theories still might be preferable in other respects, so we need to consider the many objections to consequentialism in order to determine whether the idea of necessary enablers can help to avoid those objections. We also need to specify a theory of value 22 and a theory of how the distribution of goods and harms affects moral reasons and their weights. All of these details need to be worked out before necessary enabler consequentialism can be assessed overall. Nonetheless, I do hope to have shown why this kind of theory meets one general test of adequacy and why other theories need to show how they can meet this test. 23

Notes

1 Cf Norman Daniels, ‘Wide Reflective Equilibrium and Theory Acceptance in Ethics’, Journal of Philosophy, 1979, pp 256-282 My method is an even wider reflective equilibrium including principles of moral reasoning. None of these ‘intuitions’ requires a special faculty or is supposed to be infallible

2 It might be more common to argue in terms of ‘ought’. I ought to get a cavity filled, and I can’t get it filled except by going to a dentist, so I ought to go to a dentist Although this argument also is valid and needs to be explained, it is not clear whether I ought to do something when there are overriding reasons not to do it That is why I write about reasons, which clearly can be overridden

3 Cf. my Moral Dilemmas, Oxford, Basil Blackwell, 1988, pp 152-4 Since (GS) refers to ‘doing’ X and Y, this principle is restricted to acts What I have a reason to do is usually not a particular act but some act of a relevant kind. Cf Donald Davidson, Essays on Actions and Events, New York, Oxford University Press, 1980, esp essay 6. This complication does not affect my argument, so I will sometimes write as if ‘X’ and ‘Y’ refer to acts

4 Even if some stronger principle were defensible, (GS) would still be true, and its truth would still need to be explained, so my main argument would not be affected

5. Cf Bernard Gert, Morality, New York, Oxford University Press, 1988, p 34


8. I respond to some more objections in Moral Dilemmas, sec. 5.2.

9 Yet another way to explain moral substitutability would be to show that it is analytic However, I have never seen any way to derive this principle from any definition of moral reasons.

10 I assume here that acts are identical when they are constituted by or done by means of the same bodily movements Cf Donald Davidson, Essays on Actions and Events, New York, Oxford University Press, 1980, esp. essay 3 This claim is controversial, but those who reject it can avoid all reference to identity by saying
simply that an act and its consequences cannot be constituted by the same bodily movements There are also other ways to define consequences so that fulfilling a promise is not a consequence of the act that fulfills the promise

On this account, a moral reason to do an act because the act is good in-itself or intrinsically is deontological, so some kinds of perfectionism or eudaemonism are deontological. This is probably not common usage, but it is what I want, because my argument against deontological theories will apply to these kinds of perfectionism, since the reason to do an act that is intrinsically good will be different in nature from the reason to do a necessary enabler of the act

W. D Ross, *The Right and the Good*, Indianapolis, Hackett, 1988, p 44 See also p. 17, and H A Prichard, *Moral Obligation*, New York, Oxford University Press, 1968, pp 6-7 It might seem that what makes Ross and Prichard unable to explain moral substitutability is just that they see rightness and wrongness as non-natural properties, and these need not transfer across natural causal connections. However, my argument will apply to deontologists even if they are naturalists, and some non-naturalistic consequentialists can explain moral substitutability Thus, the issue of naturalism cuts across the issue of how to explain moral substitutability.


14. *Foundations of the Metaphysics of Morals*, trans L. W. Beck, p. 417 of the Akademie edition Kant’s principle is not the same as moral substitutability, since his principle is about willing, and he applies it only to imperatives of skill, which lie outside morality Kant says that his principle is analytic, but his argument for its analyticity is hardly persuasive.

15. J. J. C Smart, ‘An Outline of a System of Utilitarian Ethics’ in *Utilitarianism for and against*, by J J C Smart and Bernard Williams, Cambridge, Cambridge University Press, 1973, p 30 Bentham also suggests sufficient conditions when he refers to whether acts ‘augment or diminish’ or ‘promote or oppose’ happiness Mill uses the words ‘produce’ and ‘promote’ I do not deny that these quotations can be read in terms of necessary enablers, but the most natural reading refers to sufficient conditions Anyway, I am more interested in the plausibility of sufficient consequentialism than in the question of who actually held this position

16. Even if my arguments do not refute sufficient consequentialism, this would not affect my arguments against deontological moral theories

17. NEC is a version of act consequentialism It seems harder for rule consequentialists to explain moral substitutability because moral substitutability generates reasons from the facts of particular situations that might not be shared by other acts that fall under the same general rules It might be disastrous for everyone to break their promises, but it would not be disastrous for everyone not to start their mowers Some rule consequentialists have responses, but there are too many varieties of rule consequentialism for me to discuss the possibilities here

18. This also applies to perfectionists and eudaemonists who claim that certain types of action are intrinsically good and that that is what gives us a reason to do them See footnote 11


20. Fred Feldman, *Doing the Best We Can*, Boston, D Reidel, 1986, pp. 5-7 See also pp 8-11 I have simplified Feldman’s example somewhat It is not clear that the reasons in his example are moral, but that does not affect my point
The coherent package becomes even larger if it is extended to reasons for belief. A reason for a belief can be seen as a fact about the belief which can make an otherwise irrational belief rational. There are two kinds of reasons for belief. reasons for the truth of belief contents and reasons for forming and maintaining belief states. Substitutability seems to work in both cases. First, if there is a reason for the truth of a belief content, there is also a reason for a belief content whose truth is a necessary enabler for the truth of the former belief content (Notice that the 'enabler' clause helps avoid paradoxes of implication.) Second, if there is a reason to form or maintain a belief state, there is also a reason to form or maintain any belief state whose existence is a necessary enabler for the former belief state. All of this adds even more support for NEC.

Some help in value theory might come from another kind of substitutability. If the necessary enablers of what is good are also good (to some extent), this might be inexplicable on desire-based theories of value, because a desire for something does not ensure a desire for its necessary enablers. I can desire to play golf without desiring to pay for it. To develop this additional argument would be beyond the scope of this paper.

For constructive criticisms, I thank Jonathan Bennett, David Cummiskey, Robert Fogelin, Bernard Gert, David McNaughton, Jim Moor, Igor Primoratz, Mark Sainsbury, Geoff Sayre McCord, Dore Scalsas, Sally Sedgwick, Holly Smith, and Timothy Sprigge, as well as audiences at Bates College, Dartmouth College, Johns Hopkins University, the Northern New England Philosophy Association, and the Universities of Aberdeen, Edinburgh, Glasgow, Leeds, St Andrews, York, and North Carolina at Chapel Hill.