

Violations of normative invariance: Some thoughts on shifty oughts¹

by

KRISTER BYKVIST

Jesus College, Oxford

Abstract: It seems paradoxical to say that an action's normative status – whether it is right, wrong, or obligatory – depends on whether or not it is performed. In this paper, I shall argue that in itself this dependency is not paradoxical. I shall argue that we should not reject a normative theory just because it implies this kind of dependency. Not all dependencies of this kind are bad, or at least not bad enough to warrant wholesale rejection. Instead, we should reject a theory when this dependency makes it a poor guide to action, in particular, when the dependency makes it impossible for agents to comply with the theory.

Keywords: normative theories, guide to action, normative invariance, ought-implies-can, ought-implies-can satisfy

1. Introduction

A: Should I tell him the truth?

B: That depends.

A: On what?

B: Whether or not you will tell him the truth.

A: Thanks! Great advice.

Whether we ought to perform an action may depend on many things, but as this dialogue suggests, we are strongly inclined to think that the performance of the action cannot be one of the determining factors. More generally, it seems paradoxical to say that an action's normative status – whether it is right, wrong or obligatory – depends on whether or not it is performed. In this paper, I shall argue that in itself this dependency is not paradoxical. In particular, I shall argue that we should not reject a norma-

¹ An ancestor of this paper was presented at the conference *Choice and Value*, Corpus Christi College, Oxford, June 2, 2001. I have benefited in particular from the suggestions and criticisms by Gustaf Arrhenius, John Broome, Erik Carlson, Roger Crisp, Jonathan Dancy, Olav Gjelsvik, Richard Holton, Brad Hooker, Derek Parfit, Wlodek Rabinowicz, John Skorupski, and Philip Stratton-Lake.

tive theory just because it implies that the normative status of an action may depend counterfactually on its own performance. Not all counterfactual dependencies of this kind are bad, or at least not bad enough to warrant wholesale rejection. To argue this point, I will compare different theories that for different reasons imply this kind of dependency and point out when this dependency is something we should be worried about.

The plan of the paper is this. In section 2, I will formulate *Normative Invariance*, the principle that bans normative variance, and provide a list of theories that for different reasons violate this principle. The reasons are summarized in section 3. In section 4, I will discuss and compare different explanations for why it is bad to violate this principle, and my conclusions are stated in section 5. My main conclusion will be that *Normative Invariance* is not a sacred principle. We should not reject a theory just because it violates this principle. We should reject a theory when the way it violates the principle makes the theory a poor guide to action.

2. Normative Invariance and its violators

To the best of my knowledge, the idea that the normative status of an action should be independent of its performance was first presented by Erik Carlson in his monograph about consequentialism.² He, in turn, got the idea and the name ‘Normative Invariance’ from Wlodek Rabinowicz. A similar idea is expressed by H. A. Prichard, who says that ‘the existence of an obligation cannot possibly depend on actual performance of the action’.³ Carlson gives the following formulation of normative invariance.

Normative Invariance (NI): An action’s normative status does not depend on whether or not it is performed.

Now, as it stands, the principle is ambiguous, since the notion of dependency employed in the principle can be given different interpretations. The principle is uncontroversial if we read it as saying that the mere fact that an action is performed or unperformed cannot be a *right-making*

2 Carlson (1995), Ch. 6, section 6.3. Unlike Carlson who only talks about invariance of moral status, I will take the principle to apply to non-moral norms as well.

3 Prichard (1986), p.37.

characteristic. Under this reading, the principle will exclude obviously implausible theories such as the theory that says that the only thing that makes an act right is that it is performed, and its mirror-image that says that the only thing that makes an act right is that it is not performed. NI is, however, supposed to exclude more. It is supposed to exclude any theory that makes the normative status of an action *counterfactually dependent* on the performance of the action. NI can be given a slightly more precise formulation that will bring this out. Let N stand for any kind of normative status, including moral rightness, wrongness, and obligatoriness. Then NI amounts to the following claim:

If A is N, then,

if A were performed, then A would be N, and

if A were not to be performed, then A would be N.⁴

I will now present a list of theories that violate NI. The first ones are a bit eccentric, but bear with me; I will later present some NI-violators that seem to be quite plausible normative theories. The reason I bother to discuss the crazy ones is that it enables us to find out exactly when and why a violation of NI is a problem.

2.1 *Indebted procreators*

Consider a theory according to which obligations are always owed to people. More exactly, if I have an obligation to do something, I have an obligation *to someone* to do this. So, obligations are relations between an agent, a person the obligation is owed to, and an action. Now, add to this the assumption that we have an obligation, owed to a person, not to create her just in case she will lead a miserable life. Then we get a theory I will call '*the indebted procreator theory*'.⁵ This theory will violate NI. To see

4 I assume here that it makes sense to say that an action has normative status even when it is not performed. I take this to be in line with how we talk about actions and their normative status. For instance, we can meaningfully say that a person ought to do something he will not do, or ought to have done something he did not do. It is true that we often say that an unperformed action *would* be right rather than that it *is* right. But the subjunctive mood is not used in this context to imply that we are making a hypothetical evaluation; it is only used to imply that we are making an evaluation of a hypothetical (unperformed) action.

5 Brock seems to embrace this theory in Buchanan *et al* (2000), pp. 236–237. He claims that we ought not to create a miserable life because the child who does exist with a life not worth living has a right not to have been brought into existence with such a life. Since non-existent beings cannot have rights, he claims that we do not have any duties to such beings.

this, assume that the choice is between creating Jane or not creating her, that she will lead a miserable life if she is created, and that no other obligations are involved. If I were to create Jane, then since she would lead a miserable life, I would have an obligation, owed to her, not to create her. If I were not to create her, then I would not have any obligation towards her, since I cannot owe anything to someone who will never exist. Thus I have the obligation not to create her only if I create her.⁶

It might sound absurd to say that I have an obligation not to create her when I have created her. For once she is created there is nothing I can do about it. To answer this objection, we must assume that once she is created I no longer have the obligation; rather, I *had* the obligation before she was created. More precisely, we have to assume that at a time before I created Jane I had an obligation, owed to her, not to create her. So what we have here is a *now-for-then* obligation: my present obligation is owed to a future actual person, a person who will exist but does not exist now.

Similarly, we could get a violation of NI by assuming that we may have an obligation, owed to a person, to create her just in case she will lead a happy life. Suppose that I create Jane and that she leads a happy life. Then, since she leads a happy life, I have an obligation to create her. Of course, again we have to say that this obligation precedes her creation. But suppose instead that I do not create her. Then, since she will never exist, I do not have any obligations toward her.

2.2 *World-relative preference-affecting theory*

We can generate NI-violations without assuming that obligations are owed to people. Assume instead that we should only be concerned with satisfying the preferences of actual people. What we ought to do in a world is then exclusively determined by what the people in the world feel or will feel about the available options and their consequences. More generally, what an agent ought to do in a certain world *w* is determined by what the

6 I assume here that it makes sense to talk about a merely possible person, a person who will not exist but would exist if created. This raises many questions about the ontology of possibilities, but I shall not delve into these thorny questions. Nothing in my argument hangs on exactly how you construe possibilities. I only have to assume that it makes sense to talk about them. It is important to note, however, that you can believe in merely possible people without identifying them with non-actual flesh-and-blood people who exist in some distant possible world. One could instead identify merely possible people with certain complexes of actually existing properties.

people in w feel about the options.

World-relative preference-affecting theory: An action is *obligatory* in a world w iff the inhabitants of w prefer its outcome to that of any alternative action.

This world-relative theory will easily generate NI-violations. Suppose that I can create Jane and that the well-being of no one else will be affected by my actions. Suppose also that if I create her, she will lead a happy life and prefer this outcome to that of not creating her. This means that I ought to create her only if I create her. For suppose that I create her. Then Jane will prefer her existence to her non-existence. Since I ought to realize the outcome Jane prefers, I ought to create her. Or, more exactly, I ought, at a time preceding her creation, to create her. Now, suppose I do not create her. Then she will not prefer being created – non-existing people do not prefer anything – and it is not true that I ought to create her.

We can of course imagine an NI-violation in a situation in which Jane would instead lead an *unhappy* life if she were created. In this situation, if I create her, then she will prefer her non-existence to her existence, which means that I ought not to create her. If, on the other hand, I do not create her, she will not prefer her non-existence to her existence, and it is not true that I ought not to create her.⁷

2.3 World-relative prudence

Not only moral theories will violate NI. Theories of prudence may also violate NI. Consider Bricker's example about a would-be philosopher who has been offered but a single job: a teaching position in the Black Hills of north-eastern Wyoming.⁸ He must now choose between moving to Wyoming and giving up professional philosophy to become a taxicab driver in New York City. Moreover, suppose that if he chooses to become a cab driver in New York, he will come to prefer this to living in Wyoming

7 You do not have to assume that preferences are relevant to generate this kind of violation. A similar violation is implied by a theory that says that (a) we ought to make people better off, (b) creating a happy person makes that person better off, and (c) not creating a happy person does not make that person worse off. According to this theory, if we create Jane and she leads a happy life, then we make her better off and no one worse off. Hence, if we create her, then this is what we ought to do. If we do not create her, we do not make anyone worse off, and it is not true that we ought to create her. The idea that value can vary in this way is defended in Holtug (2001). For a more critical discussion, see Parfit (1984), p. 489, and Broome (1999), p. 168.

8 Bricker (1980), pp. 397–398.

– country living would not be for him! If he chooses to live in Wyoming, however, he will come to prefer ‘the lowing of cattle to the honking of horns, would soon prefer a rodeo to a symphony’, and, on the whole, prefer living in Wyoming to living as a cab driver in New York.

Now assume that in line with the world-relative preference-affecting theory presented earlier an agent has a prudential obligation to realize the outcome that she actually prefers and will prefer the most, i.e., the outcome that best satisfies her actual preferences (present and future).

World-relative prudence: An action is *obligatory* for agent S in w iff its outcome is overall preferred by S in w to that of any alternative action.

The situation is then the following (assuming that he has at the time of choice no preference either way). If he moves to New York, he will prefer to be a cab driver in New York, and thus he ought now to go to New York. If he moves to Wyoming, he will instead prefer the life of a philosopher in Wyoming, and ought now to move to Wyoming. So, what he ought to do depends on what he will do.

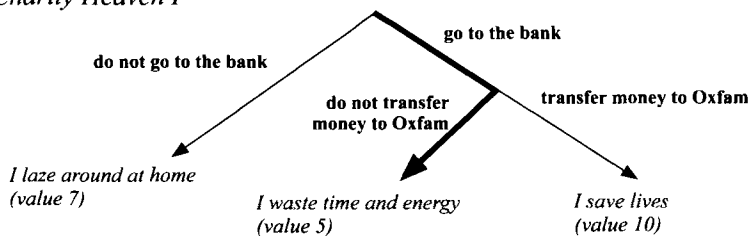
The reason why these world-relative theories violate NI is simple: the shift in normative status is explained by the fact that the normatively relevant preferences depend counterfactually on the performance of the action. Admittedly, both world-relative theories and the ‘indebted procreator’ theory seem controversial even setting aside the fact that they violate NI. Even if one think that preferences are prudentially relevant, one might reasonably deny that they are relevant in the sense dictated by world-relative theories. Similarly, one might deny that duties are owed to people in the way described by the ‘indebted procreator’ theory. I do not want to deny that there are other ways you can be concerned with discharging your duties to people, or satisfying your own or other people’s preferences – ways that will not lead to violations of NI. But I am not going to discuss these possible refinements. My question here is instead whether the mere fact that a theory violates NI should prompt us to refine it. Things would be easy if only implausible theories violated NI. But, as I will show in the next section, there are theories that have much going for them but nevertheless fail to satisfy NI.

2.4 *Consequence-sensitive theories*

I take it that any reasonable moral theory assesses actions partly by the values of their consequences. Consequentialists claim that nothing but

the values of consequences matter for the assessment of actions, whereas deontologists claim that options and constraints matter too. We are allowed but not always required to maximize goodness, and we are only allowed to maximize goodness given that no constraints are violated, e.g., given that we do not tell any lies, break any promises, kill any persons and so on. But how do we decide whether an action would maximize value? In other words, how do we decide whether the outcome of an action is better than that of any *alternative* action? The crucial problem here is of course how to select the relevant alternatives. One simple solution would be to say that *all* actions that are immediately performable in the situation are relevant alternatives.⁹ But, as Carlson has pointed out, this *option-promiscuous* theory will violate NI.¹⁰ To see this, consider the following example.

Charity Heaven I



The thick line represents the fact that I will actually go to the bank, but refrain from transferring money to Oxfam. This means, of course, that if I *were* to go the bank, I *would* not transfer money to Oxfam (assuming strong centering for counterfactuals, i.e., ‘A and B’ entails ‘If A were the case, then B would be the case’). It is important to stress that I will in no way be forced to refrain; it is just that I will as a matter of fact freely choose to refrain. Perhaps my stinginess will be triggered once I am in the bank. This means, of course, that I will *not* perform the compound act of going to the bank and transferring money to Oxfam. Now,

⁹ Immediate performability is the ability at a certain time to perform an action that starts at that time. Note that an action that extends into the future can still be immediately performable at the present moment. The crucial thing is that the action starts at the present moment, not that the act is completed at this moment.

¹⁰ Carlson (1995), p. 100.

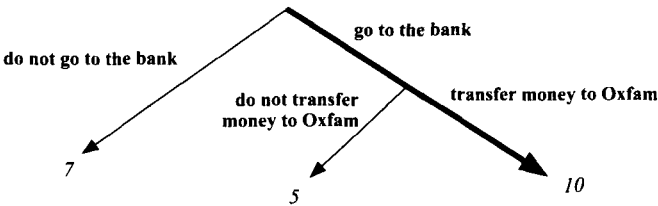
if all immediately performable actions are relevant alternatives, then the consequentialist would say that the compound act of going to the bank and transferring money to Oxfam is obligatory, since the outcome of this action is better than that of any alternative action, as can be seen from the following chart.

<i>Alternatives (immediately performable)</i>	<i>Outcome-values</i>
my act of not going to the bank	7
my act of going to the bank	5
my compound act of going to the bank and transferring money to Oxfam	10
my compound act of going to the bank and not transferring money to Oxfam	5

Even a deontologist can accept that I ought to go to the bank and transfer money if we assume that I will not violate any deontological constraints by doing this, and that it will not be especially costly for me to do it.

But note that this action would not have been obligatory if I had gone to the bank and transferred money to Oxfam. For in this situation my act of going to the bank and my act of going to the bank and transferring money to Oxfam would have the *same* total consequences. (I am assuming here that the total consequence of an action is the whole possible world or the whole future that would come about if the action were performed.)

Charity Heaven II



Both actions would be right as can be seen from the following chart.

<i>Alternatives</i>	<i>Outcome-values</i>
my act of not going to the bank	7
my act of going to the bank	10

my compound act of going to the bank and transferring money to Oxfam	10
my compound act of going to the bank and not transferring money to Oxfam	5

Hence, going to the bank and transferring money to Oxfam is obligatory only if it is not performed – a clear violation of NI.

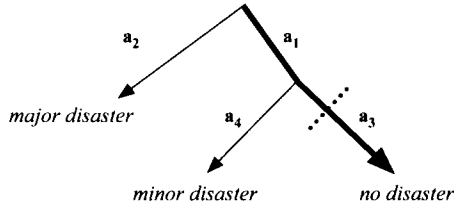
The diagnosis of this NI-violation is that the normative status of an action varies with its performance because the *outcomes* of the relevant alternatives depend counterfactually on the performance of the action. If I will go to the bank but not transfer money, then going to the bank will lead to the outcome of value 5. If I will go to the bank and transfer money, then going to the bank will lead to the outcome of value 10.

One could, of course, attribute this result to the notion of alternative actions employed. More specifically, one could claim that the root of the problem is that alternative actions are allowed to be *mutually compatible*. In the example, going to the bank is compatible with going to the bank and transferring money to Oxfam. But one could claim that a pair of actions should not be seen as alternatives if they are not mutually exclusive. If we respect this constraint, we can avoid normative variance in the charity case. For instance, we would avoid it if the set of alternatives consisted only of *maximal* actions, i.e., actions that are performable but not proper parts of anything else the agent can do.¹¹ On this account, the relevant alternatives in the charity case are the act of not going to the bank and the compound actions. Since the outcomes of these actions do not depend on which of them is performed, their normative status will not depend on whether or not they are performed.

¹¹ Alternatively, one could go for *minimal* actions instead, i.e., those actions that once begun cannot be stopped short of completion. However, I doubt that this will give us an acceptable theory. First of all, this would mean that only very basic actions, such as my raising my arm, can be assigned normative status. This seems to be too restrictive. Secondly, non-minimal actions that are clearly within my present control will then be treated on a par with the actions of other people. The agent will let his predictions of his future actions influence his deliberation. Of course, this might be a justified attitude to take when the future self exists in the remote future and are beyond the control of the present self. But it is surely an odd attitude to take when the future self exists in the near future, one minute ahead say. For more on the problems with minimalism, see Carlsson (1996), pp. 130–133.

This approach is not fully satisfactory, however. Consider the following schematic example.¹²

Disaster I



The thick line represents the fact that the agent will first do a_1 and then later do a_3 . This means, of course, that if I *were* to do a_1 , then I *would* do a_3 , and that a_1 -and- a_4 will not be performed. Again, it is important to stress that the agent will in no way be forced to do a_3 ; it is just that he will as a matter of fact freely choose to do a_3 . However, the dotted line represents the fact that the compound act a_1 -and- a_3 is not performable by the agent in his *present* situation, although a_3 *will* be performable once he has reached the future choice point.

One could, of course, wonder if this kind of case is possible. If a_1 would be followed by a_3 and the agent can do a_1 , must we not also say that he can do a_1 -and- a_3 ? Not if we make the reasonable assumption that an agent's abilities to act might grow with time. I take it that normative theories are primarily concerned with intentional actions. I therefore assume that our ability to act at a certain time should be understood in terms of what we can do *intentionally* at that time. This means that an act is performable at a time only if the agent can at this time form an intention to do it. Since one cannot form an intention to do a certain action if one does not have the concepts necessary to grasp the action, an act is performable only if the agent can grasp the act. This explains why a_1 -and- a_3 might not be performable in the original situation. For suppose that the agent at the original choice point cannot comprehend a_3 . Imagine, for instance, that a_3 is the act of formally deducing a theorem in predicate logic and that at the original choice point the agent has no idea about what it means to prove a theorem in predicate logic, because he lacks the concepts of 'proof', 'theorem', and 'predicate logic'. Then since he cannot comprehend a_3 ,

¹² This is a slightly modified version of an example presented Carlson (1996), p. 122.

he cannot at this time form the intention to do a_3 , or any act that includes a_3 . Consequently, a_1 -and- a_3 is not performable at this time.¹³ However, we may assume that the agent will have gained a firm grasp of the crucial logical concepts involved in predicate logic at the time he reaches the next choice point (perhaps a_1 = taking a philosophy course). At this point he does understand predicate logic and can derive the theorem and thereby avoid a minor disaster. It is not often you can make the world a better place by doing predicate logic!

What would maximalism say in this situation? Well, of the maximal options a_1 -and- a_4 and a_2 , a_1 -and- a_4 is the one that will have better consequences. So, maximalism would tell us to do a_1 -and- a_4 . This is counterintuitive, since there is a way to avoid a disaster, namely by doing a_1 . A reasonable consequentialist should therefore say that the agent ought to do a_1 and that a_1 -and- a_4 and a_2 are wrong. A deontologist should agree if we assume that a_1 does not violate any deontological constraints and does not ask for any great sacrifice on the agent's part. If this is true, then we should not think that alternatives must be mutually exclusive. The situation could be structured thus:

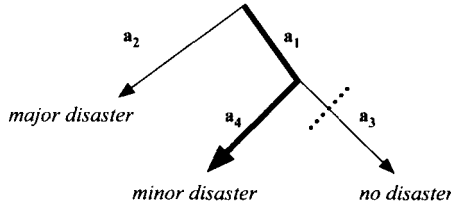
<i>Alternatives</i>	<i>Outcomes</i>
a_1	no disaster
a_1 -and- a_4	minor disaster
a_2	major disaster

According to this theory, an agent's set of alternatives should consist of all maximal actions plus those non-maximal actions that would be followed by actions of his that are beyond his present control. Let us call this theory *quasi-maximalism*.¹⁴ Unfortunately, this theory will also violate NI. To see this, suppose that a_1 -and- a_4 is performed. Then we have the following situation:

13 Note that if I do a_1 intentionally and the later do a_3 intentionally it does not follow that I do a_1 -and- a_3 intentionally. If I intentionally brush my teeth in the morning and two days later, intentionally have coffee with a friend I happen to run into, it does not follow that I intentionally performed the compound action of brushing my teeth and later having coffee with my friend. For I might not have intended to perform this compound action. So we cannot argue from the fact that I will do both a_1 and a_3 intentionally to the conclusion that I will do a_1 -and- a_3 intentionally and thus that I must be able to perform this compound action intentionally.

14 I defend this theory in Bykvist (2002).

Disaster II



Here there is no way to avoid a disaster. We have to choose between a major and a minor disaster. So, a_1 -and- a_4 must be right, since it will bring about less badness. (a_1 is not an alternative action anymore.) Suppose that if a_1 -and- a_4 had not been performed, a_1 would still have been performed but would later have been followed by a_3 . Then we have a violation of NI. For, in this situation, if a_1 -and- a_4 had not been performed, then a_1 would have been an alternative whose outcome is better than that of a_1 -and- a_4 , and a_1 -and- a_4 would therefore have been wrong.

The diagnosis of this NI-violation is, of course, that the normative status of an action varies with its performance because the set of *alternative* actions depends counterfactually on the performance of the action. If a_1 -and- a_4 is not performed, then a_1 is an alternative action. But if a_1 -and- a_4 is performed, then a_1 is not an alternative action.

So, it seems to be difficult to avoid violating NI if you accept a plausible consequence-sensitive theory.¹⁵

2.5 Conditional analysis of can (iffy can)

This far we have only seen violations of NI where the normative status is shifted from one kind to another, e.g., from obligatory to merely right, or from right to wrong. But there are also cases where whether an action has *any* normative status depends on whether or not it is performed. Wlodek Rabinowicz gives the following example (cited in Carlson).¹⁶ Assume that a simple action A is performable by a person in a situation only if it holds that if he wanted or tried to perform A, then he would succeed. Assume that I try to lift a 100 pound weight and succeed. Then, assuming

15 The consequence-sensitive theory defended by Carlson (1996) will also violate NI, and so will the theory defended by Åqvist (1969). This is shown in Carlson (1996), pp. 105–107, 113–115.

16 Carlson (1995), p. 100, fn. 30.

that all other conditions for performability are in place, my lifting this weight is performable, and may very well have some determinate normative status, e.g., rightness. But this fact does not guarantee that I would have had the ability to lift the weight had I *not* lifted the weight. For the closest world in which I do not lift the weight might be a world in which I try to lift the weight but fail. This means that this action would not have had any normative status if it had not been performed, since it would not have been performable. I assume here that an action has normative status in a situation only if it is performable in this situation.

So, in this case we get a violation of NI because whether an action is *performable* depends on whether or not it is performed. An iff can give rise to a *shifty* normative status.

3. Time to take stock

This completes my list of NI-violators. They all allow that whether or not an action has certain right-making characteristics may depend counterfactually on whether or not the action is performed. The following right-making characteristics have been mentioned:

- (a) The existence of the person to whom we owe an obligation (the indebted procreator theory).
- (b) The values of outcomes (world-relative person-affecting and prudence).
- (c) The outcomes of the available options (option-promiscuous consequence-sensitive theories).
- (d) The composition of the set of relevant alternatives (quasi-maximalist consequence-sensitive theories).
- (e) The ability to perform an action. (conditional analysis of can)

4. What is the problem?

4.1 *The argument from action-guidingness.*

As the dialogue presented in the beginning of this paper brings out, a theory that violates NI is a poor *guide* to action. One might take this to be a decisive argument for NI for the following reasons. When you use a

theory as a guide to action, you use the theory in your deliberations about what to do. On the basis of this deliberation you then make up your mind and decide what to do. But if your theory violates NI, then in order to decide whether an action has a certain normative status (or any normative status at all) you have to know whether or not you are going to perform it. But there is no point in deliberating about whether to perform an action if either you believe that you will perform it, or you believe that you will not perform it. If you believe that you will perform the action, the issue is settled for you, and there is no point in deliberating about it further. If you believe that you will not perform the action, the action is no longer a serious possibility, i.e., something that is compatible with what you believe (even if it might be something you can do); so again there is no point in deliberating about whether to perform it.¹⁷

The claim here is *not* that once you have formed the belief (or disbelief) that you are going to do A, then you are no longer able to deliberate about whether to do A. If you give up the belief or the disbelief, you may start deliberating again. The claim is rather that *while* you are in the grip of the belief or disbelief that you will do A, it is not possible for you to deliberate about whether to do A. Or at least, this is not possible if you are *rational*. For rational agents, belief or disbelief about what they are going to do excludes wondering about whether to do it. The rational wondering whether to do A cannot cohabit in one's mind with either the belief or the disbelief that one is going to do A.¹⁸

It seems clear that NI-violators are sometimes poor guides to action. But does this make them unacceptable? I doubt it. First of all, note that it is not a good objection to a normative theory that it *sometimes* fails as a guide to action. All plausible theories do, since their application often requires information the agent lacks and might not be able to gather and process in a limited time. We all know that utilitarians have this problem, because they need information about how their actions will affect *future* well-being. But deontologists will face similar problems. For instance,

17 This argument for NI is spelled out in Carlson (1995), pp. 101–102. Bricker (1980), p. 395, touches on a similar argument. The general idea that the prediction of one's actions crowds out deliberation has wide-spread support. See, for instance, Goldman (1970), p. 194, and Taylor (1966), p. 174. Cf. also Ginet (1962), Levi (1997), Schick (1988), and Spohn (1977).

18 In Tennant (1997), pp. 252–254, this is generalized to hold for all rational wonderings no matter what they concern. The rational wondering whether *p* excludes both the belief and the disbelief that *p*.

they may need to know a lot about *past* suffering and wrong-doings in order to know whom to compensate, and how to judge the present distribution of goods and burdens in society. More generally, since any reasonable moral theory must incorporate a principle of non-malevolence, the problem about knowing the well-being of future people will reappear in the deontological framework. Only clearly implausible theories, such as one that tells us to do whatever you feel like doing, will be easy to correctly apply by any agent in any situation. To demand that a normative theory should be usable by anyone in any situation would be to ignore the *theoretical function* of a normative theory, i.e., its function to give an adequate explanation of what makes actions obligatory, right, and wrong. And it seems unavoidable that a theoretically appealing theory will not be usable by fallible humans.¹⁹

However, it is reasonable to claim that a normative theory should at least be usable by agents who do not have any cognitive or motivational shortcomings. A normative theory is not disqualified just because it is not usable by agents who lack necessary information, understanding, and motivation. However, if a theory cannot even be used by agents who are ideal, then it seems to be devoid of all practical relevance, and one might wonder how it could be properly called a *normative* theory - a theory that is supposed to tell you what to do.²⁰ Now, ideal agents will not have any problems about knowing the relevant aspects of the past and the future. So, theories that demand this kind of information can still be usable by ideal agents.

What about the theories that violate NI? All theories I have listed are usable by ideal agents in some situations. There is no difficulty applying the indebted procreator theory and the world-relative theories in situations that do not involve any contingent people or preferences. Option-promiscuous theories can easily be used in situations where no compound actions are involved. Quasi-maximalist theories can easily be applied

19 For more on the distinction between the theoretical and the practical function of morality, see Smith (1988), pp. 89–90.

20 As Holly Smith observes in Smith (1986), p. 342f, the reason why we need normative principles is that 'the criterion of rightness provides the kind of evaluation of choosable events which can form the basis for guiding choices with respect to those events. It is precisely because we need some standard of evaluation to serve this function that we have criteria of rightness in addition to criteria of goodness.'

when your power over your future actions is not constrained by the inability to grasp certain future actions. Finally, theories that assume an iffy can are usable in situations where your ability to act is not at stake.

So, all theories are sometimes usable by ideal agents. However, I think it is clear that the theories that incorporate an iffy can and the quasi-maximalist theory fare better than the others in respect of usability. For note that these theories violate NI only when the agent's ability is at stake or her ability to grasp future actions is limited. In other words, these theories violate NI only when the agent is less than ideal. In contrast, the other theories will violate NI no matter how much you idealize the agent's knowledge, motivation, and ability to act. That they are poor guides to action is therefore more of a failing in them than a failing in the agents.

4.2 *An obligation one cannot violate*

One could claim that not all NI-violations are problematic. The problematic violation is when an action would be obligatory only if it were performed. As the examples discussed in section 2 show, all the theories we have listed imply this kind of normative shift. Why is this shift problematic? Narveson claims that 'an obligation that only exists if it is fulfilled, i.e., which logically cannot be violated, is clearly nonsense.'²¹

Why is this nonsense? One might think that these theories violate the principle that 'ought' entails 'can *refrain*'.

'*Ought*' entails '*can refrain*': If I am obligated to do A, then it must be possible for me to refrain from doing A.

This would be a mistake, however. A theory that violates NI need not have any problems with this principle. In fact, all the theories we have considered satisfy it. They will all imply that you can fail to perform your actual duties. They do, however, violate the principle that 'ought' entails 'can violate' or 'can frustrate':

'*Ought*' entails '*can frustrate*': If am obligated to do A, then it must be possible for me to frustrate this obligation.

How does this principle differ from the 'ought'-entails-'can refrain' prin-

21 Narveson (1996), p. 44.

ciple? To be able to frustrate the obligation to do A is not just to be able to refrain from doing A. It is to be able refrain from doing A *while* being obligated to do A. So, it is more accurate to phrase the principle thus:

If am obligated to do A, then it must be possible for me to refrain from doing A while being obligated to do A.

Now, if an action is obligatory only if it is performed, then it is not possible to frustrate this obligation. For if I *had* refrained from doing what I am in the actual situation obligated to do, then I *would* not have been obligated to do it.

I doubt that the ‘ought’-entails-‘can-frustrate’ principle is valid. For even if I can’t frustrate my actual obligations, I may still be able to act wrongly. To see this, consider the quasi-maximalist consequence-theory. In Disaster case II, this theory will imply that I am obligated to do a_1 -and- a_4 . However, had I not performed this action, then it would have been wrong. So, I can’t frustrate my actual duty in this situation. But this does not make it impossible for me to act wrongly. I can, for instance, do a_2 , which is wrong in the actual situation and would still be wrong if I did it. In other words, I am still able to disobey or frustrate my theory. There is an action available such that if I were to perform it, I would do something that *would* (then) be wrong according to my theory.

One should not conflate these three claims:

- (1) It is not possible for me to refrain from performing the action that is my actual duty.
- (2) It is not possible for me to frustrate my actual duty, (i.e., it is not possible for me to refrain from doing what I ought to do in the actual situation and still be obligated to do it.)
- (3) It is not possible for me to frustrate my normative theory.

What is ‘peculiar’ is not (2), but (1) and (3). And perhaps (1) and (3) are ‘peculiar’ enough to constitute a *reductio*. I will not take a stand on this, however. The important point is that a normative theory is not necessarily disqualified by the fact that it entails (2).

4.3 *A duty one cannot perform.*

One could argue that the problematic NI-violation is when an action would be obligatory only if it were *not* performed. The indebted procreator theory, the world-relative theories, and the option-promiscuous the-

ory entail this kind of violation.²² One could argue that this shift is absurd since it would be a duty that exists only if it is not performed. It is a duty one cannot perform.²³ In other words, these obligations are necessarily elusive. But why can't there be such a duty? It is true that duties must be performable in the sense that 'ought' entails 'can':

'Ought' entails 'can': If I have an obligation to do A, then it must be possible for me to do A.

However, theories that violate NI need not have any problems with this principle. In fact, all the theories we have examined so far satisfy this principle. You can always perform the action that is your actual duty according to these theories. What they have problems with is the principle that 'ought' entails 'can satisfy':

'Ought' entails 'can satisfy': If I have an obligation to do A, then it is possible for me to *discharge* or *satisfy* this obligation.

A more exact formulation of the principle would be the following:

If I have an obligation to do A, then it is possible for me to do A while being obligated to do A.

The NI-violators we are considering entail that it is not possible to do A while being obligated to A. For if I were to do A, which is the action that I am obligated to do in the actual situation, then I would not be obligated to do A.

Is this a problem? Not necessarily. Note that even if I can't satisfy my actual duty, I may still be able to *comply with* my normative theory. Charity Heaven I brings this out. In this case, an option-promiscuous theory will say that I ought to go to the bank and transfer money to Oxfam. But had I performed this action, it would not have been obligatory, just right. So, in this situation I can't satisfy my actual duty. However, I can still comply with the theory. That is, there is an available action such that if I *were* to perform it, I *would* comply with the theory and act rightly. I would comply with the theory if I were to go to the bank and transfer money to Oxfam. For in this counterfactual situation this is the right thing to do.

22 The indebted procreator theory and the world-relative theories imply that we ought to *refrain* from creating an unhappy person only if we create the person.

23 John Broome suggested this to me.

As in the case of duties that cannot be violated, it is important to distinguish three claims:

- (1) It is not possible for me to perform the action that is my actual duty.
- (2) It is not possible for me to satisfy my actual duty, (i.e., it is not possible for me to do what I ought to do in the actual situation and still be obligated to do it.)
- (3) It is not possible for me to satisfy my normative theory.

What I am suggesting is that (2) in itself is not problematic; (1) and (3) are the culprits. (1) is a violation of the 'ought'-entails-'can' principle. (3) violates the requirement that a normative theory should be *satisfiable*, i.e., possible to comply with. Let us focus a bit more on this satisfiability requirement.

4.4 A theory one cannot satisfy.

I have suggested that the problem is not when a particular obligation fails to be satisfiable; the problem is when the normative theory in question fails to be satisfiable. In other words, every normative theory should meet the following requirement:

Satisfiability: For any agent and any possible situation, there is an action such that if the agent were to perform the action in this situation, then she would conform to the theory.²⁴

What is especially troublesome is therefore a case where, if you did A, A would be wrong and not doing A right, whereas if you did not do A, A would be right and not doing A wrong. For this situation involves unavoidable wrong-doing in the sense that whatever you were to do, you would do something that would be wrong. You are damned if you do, damned if you don't. Or more exactly, you *would* be damned if you *were* to do it, and you *would* be damned if you *were* not to do it. So, what we have here is a kind of normative dilemma, but one that differs from the usual kind. Normally, a dilemma is seen as a situation in which all available actions

²⁴ This principle is similar to 'the principle of separate satisfiability presented in Österberg (1986), p. 127: 'for any agent A and any possible situation, it be logically possible that A conforms to the principle.'

are wrong. This is not the situation here. No matter how you act, there is an available act that *is* right. If A is performed, then refraining from doing A is right; if A is not performed, A is right. But this is cold comfort. For you cannot act in such a way that were you to act in that way you would comply with the theory.²⁵

Why is important that a normative theory is possible to comply with? Again, the answer has to do with the practical role of a normative theory. A normative theory is supposed to guide action and should therefore at least be usable by agents who are ideal in the sense that they lack all cognitive and motivational defects. But if we cannot comply with a theory in a situation, we cannot be guided by it in that situation, since to be guided by a theory is to use that theory when deliberating about what to do and on the basis of this deliberation act in compliance with the theory.

This, I think, is one important reason why the indebted procreator theory and world-relative theories should be rejected. For, as the following examples show, they all imply that in some situations it is impossible for the agent to comply with the theories.

Indebted procreators will have to face this dilemma (assuming that no other duties are involved.). Suppose that my choice is between creating Jane or John (no other option is available), and that each person would lead a miserable life. Now, if I create Jane, then, since she will lead a miserable life, I have an obligation not to create her. So, in this scenario I will fail to do what I ought to do. If I create John, then, again, I will fail to do what I ought to do, for, in this scenario, John will lead a horrible life and I ought therefore not to create him.

Also the preference-affecting theory will generate a dilemma in this situation (assuming that no other people are affected). For if I create Jane, she will prefer her non-existence to her existence, and, since John does not exist, I ought, all preferences considered, not to create Jane. If I cre-

25 As Richard Holton brought to my attention, Yablo (1993), p. 164, makes a similar observation in the context of circular definitions. He shows that there are predicates that are defined in such a way that no matter whether we would pronounce them to be true or not true of something, we would be forced to take our pronouncement back. So whatever we were to do, we would fail to comply with the definition. Yablo notices too that this case differs from the usual moral dilemma. As he puts it, 'Where the usual moral examples involve *coexisting* imperatives such that complying with one means defying the other, in the semantical case I *create* the imperative I defy by complying with its competitor.' (page 164).

ate John, on the other hand, then John will prefer his non-existence to his existence, and since Jane does not exist, John's preferences are the only relevant ones and I ought therefore not to create him.

Finally, here is a dilemma for the world-relative theory of prudence (assuming that no other preferences are affected). Suppose that you know that if you get married, then you will prefer being unmarried to being married. You will adopt certain perfectionist ideas about marriage and think that your marriage does not live up to the standards. However, if you stay unmarried, you will accept less exacting requirements and prefer being married to being unmarried. Now, if you get married, then you ought not to get married, since you will then prefer not being married, whereas if you do not get married, you ought to get married, since this is what you will then prefer.

In contrast, a theory can be consequence-sensitive or incorporate an iff can and still be satisfiable.²⁶ It is true that such a theory will generate normative shifts in situations where the agent's power over the future is limited, or his ability to perform an action is at stake. But these shifts will never saddle the agent with a bind in which each option would be wrong if performed. There is always an option that would be right if performed.

In sum, what is wrong with the indebted procreator theory and the world-relative theories is not the mere fact that they violate NI. What is wrong with them is that they violate NI in ways that make them especially poor guides to action. First, they short-circuit rational deliberation by asking us to decide what we will do before we decide what we ought to do. Second, they create dilemmas where there is no way out: whatever we did we would act wrongly. Since these implications hold no matter how ideal we assume we are, there is reason to think that this is a failing in the theories rather than in us.

26 More exactly, they are satisfiable given that they employ a sound axiology, an axiology that does not allow that the value of an outcome depends on whether it obtains, or that the value is cyclical. If the value of an outcome might depend on whether it obtains, then obviously a value-maximizer might find himself in a situation in which whatever he were to do he would fail to do what would maximize value. Perhaps outcome A is worse than outcome B if A obtains, and B is worse than A if B obtains. Similarly, if value can be cyclical so that outcome A is better than outcome B, B is better than outcome C, and C is better than A, then a value-maximizer will not be able to maximize value – all options are wrong.

5. Conclusions

I have argued that a normative theory is not disqualified just because it allows that the normative status of an action depends on whether it is performed. A theory is disqualified if it allows a shift in normative status that makes it impossible for even ideal agents to be guided by the theory. The indebted procreator theory, the world-relative theories, and the option-promiscuous theories all imply normative shifts that will prevent ideal agents from engaging in rational deliberation in some situations. In order to apply these theories, you have to decide what they will do before you decide what you ought to do. But you cannot rationally deliberate about what you ought to do when you have already made up your mind about what you will do. The indebted procreator theory and the world-relative theories have a further defect. They imply normative shifts that prevent the agent from acting in compliance with the theories: in some situations, no matter what we did, we would do something that would be wrong.

In contrast, a quasi-maximalist theory (or a theory that incorporates an iffy can) will always be satisfiable and action-guiding for agents that are ideal in respect to what they know, want to do, and can do. This, of course, does not show that this theory is acceptable all things considered. But it does show that, despite the fact that it violates NI, it is still sufficiently action-guiding to be a good candidate for an acceptable normative theory. NI is therefore not a sacred principle.

References

- BRICKER, P., 1980, 'Prudence', *Journal of Philosophy*, Vol LXXVII, No 7, pp. 381–401.
- BROOME, J., 1999, *Ethics Out of Economics*, Cambridge University Press.
- BUCHANAN, A, BROCK, D. W., WIKLER, D., 2000, *From Chance to Choice*, CUP.
- BYKVIST, K., 2002, 'Alternative Actions and the Spirit of Consequentialism', *Philosophical Studies*, Vol. 107, No 1.
- CARLSON, E., 1995, *Consequentialism Reconsidered*, Kluwer.
- GINET, C., 1962, 'Can the Will Be Caused?', *The Philosophical Review* 71, 1962, pp. 49–55.
- GOLDMAN, A. I., 1970, *A Theory of Human Action*, Princeton University Press.
- HOLTUG, N., 2001, 'On the Value of Coming Into Existence', *The Journal of Ethics*, pp. 361–384.
- LEVI, I., 1997, 'Consequentialism and Sequential Choice', in his *Covenant of*

- Reason*, Cambridge University Press, pp. 70–101.
- NARVESON, J., 1996, 'Future People and Us', in *Obligations to Future Generations*, Sikora R. I., and Barry, B., eds., The White Horse Press, Cambridge.
- PARFIT, D., 1984, *Reasons and Persons*, Oxford University Press.
- PRICHARD, H. A., 1968, *Moral Obligation. Duty and Interest. Essays and Lectures by H.A. Prichard*, OUP, Oxford.
- SCHICK, F., 1988, 'Self-knowledge, uncertainty, and choice' in *Decision, probability, and utility. Selected readings*, Gärdenfors, P., Sahlin, N-E, (eds.), Cambridge University Press, pp. 270–286.
- SMITH, H., 1986, 'Moral Realism, Moral Conflict, and Compound Acts', *Journal of Philosophy* 83, pp. 341–345.
- SMITH, H., 1988, 'Making Moral Decisions', *Noûs* 22, pp. 89–108.
- SPOHN, W., 1977, 'Where Luce and Krantz Do Really Generalize Savage's Decision Model' *Erkenntnis* 11, pp. 113–134.
- TAYLOR, R., 1966, *Action and Purpose*, Prentice-Hall, Inc.
- TENNANT, N., 1997, *The Taming of The True*, Oxford University Press.
- YABLO, S., 1993, 'Definitions, Consistent and Inconsistent', *Philosophical Studies* 72, pp. 147–175.
- ÅQVIST, L., 1969, 'Improved Formulations of Act-Utilitarianism', *Noûs* 3, pp. 299–323.
- ÖSTERBERG, J., 1986, *Self and Others. A Study of Ethical Egoism*, Synthese Library, Kluwer.