

LEXICAL PRIORITY AND THE PROBLEM OF RISK

BY

MICHAEL HUEMER

Abstract: Some theories of practical reasons incorporate a lexical priority structure, according to which some practical reasons have infinitely greater weight than others. This includes absolute deontological theories and axiological theories that take some goods to be categorically superior to others. These theories face problems involving cases in which there is a non-extreme *probability* that a given reason applies. In view of such cases, lexical-priority theories are in danger of becoming irrelevant to decision-making, becoming absurdly demanding, or generating paradoxical cases in which each of a pair of actions is permissible yet the pair is impermissible.

1. The Risk Paradox for Absolutism

Absolute deontological ethical theories (hereafter, 'absolutist theories') hold that certain kinds of action must be performed, or must be avoided, regardless of how good or bad the consequences may be. These action types may be individuated in part by motives or intentions, or by historical or other circumstances (but not by the value of the consequences). For instance, Kant held that one must always keep one's promises, no matter how bad the consequences of doing so may be, or how much good might be brought about by breaking a promise.¹ Elizabeth Anscombe held that it is always wrong to knowingly punish a person for a crime he did not commit.² And Robert Nozick appears to have held that it is always wrong to violate an individual's negative rights against coercion.³

Recent discussions have highlighted an ethical paradox that confronts such views. Suppose that it is always wrong, regardless of the consequences, to perform an action known to be of kind K. What should we say in general about cases in which a contemplated action, A, has a probability p of being of kind K? There seem to be three views worth considering:

- a. We might hold that A is wrong, regardless of the consequences, as long as $p > 0$. Depending on the details of the deontological theory in question, this is likely to result in an implausibly, even absurdly, demanding ethical theory. At

- any given time, nearly any contingent proposition has a nonzero probability on one's evidence, and almost any action has a nonzero probability of falling under almost any interesting kind.⁴ For instance, for any action I might perform, there is a nonzero probability that it will violate a promise I made at some earlier time (but perhaps have forgotten). If one takes an absolutist view about the obligation of promise-keeping, one is then driven to the view that all actions are wrong. Other absolute deontological norms may lead to somewhat less demanding conclusions than the absolute prohibition on promise-breaking, but their consequences are nevertheless likely to be very unpalatable.⁵
- b. We might hold that A is wrong, regardless of the consequences, when $p = 1$, but that the consequences matter when $p < 1$; that is, the absolute prohibition only applies when the probability reaches 100%. This is likely to render the absolutist element in our ethical theory impotent and pointless. Because almost nothing has a probability of 1, the absolute proscription will never, or almost never, apply. This is likely to defeat the purpose of introducing such a proscription.
 - c. We might posit a threshold level of probability, t (call it the 'risk threshold'), such that when $p < t$, A may be rendered permissible if the consequences are sufficiently favorable, but when $p > t$, A must not be performed regardless of the consequences. This theory leads to paradoxical cases in which it is permissible to perform act A (whether or not one performs B), and it is permissible to perform B (whether or not one performs A), and yet it is impermissible to perform both A and B. For suppose there is a probability slightly below t that A is of kind K, and a probability slightly below t that B is of kind K. Suppose further that A and B each would (regardless of whether the other is performed) produce extremely good consequences, so that each is permissible according to the risk threshold theory. If, however, one performs both A and B, there is a probability greater than t that one will perform at least one action of kind K. Depending on the details of the deontological theory, this will typically lead to the conclusion that performing A and B is absolutely prohibited.⁶

To illustrate the problem, suppose one holds, with Anscombe, that it is always wrong to knowingly punish the innocent, regardless of the consequences. If so, what level of certainty of guilt ought we to require before the defendant in a criminal trial may be convicted and punished? If we require absolute certainty, then we have the implausible result that accused criminals should never be punished. If we require something less than certainty, such that criminals may periodically be punished, then it is virtually certain that the system will also punish *some* innocent people. If it is always wrong to knowingly punish an innocent person, then it would seem also to be wrong to institute a system that one knows will punish a number of innocent people. Anscombe's absolutist view of criminal justice thus threatens to generate a prohibition on any meaningful criminal justice

system. Other absolutist proscriptions are likely to lead to similar problems.⁷

2. *The General Problem of Risk*

2.1. THE LEXICAL PRIORITY THESIS

The risk paradox does not only affect absolute deontological ethical theories. It applies to a broad class of theories of practical reasons, namely, those theories that incorporate a lexical priority structure (hereafter, 'lexical priority views'). Lexical priority views hold that some normative considerations have categorically greater weight than others, in the sense that considerations of the latter kind, however multiplied, can never outweigh considerations of the former kind.⁸ These theories hold:

There exists some practical reason R, and some good G, such that R always takes precedence over our reason for promoting G, regardless of the quantity of G involved.

There may be more than one value of R and more than one value of G; that is, there may be multiple practical reasons and goods such that those reasons outweigh our reason to promote any quantity of those goods. In any such case, I shall call the overriding reason a 'higher reason' and the good that is thus overridden a 'lesser good.'

There are at least two different ways of understanding 'practical reasons.' Imagine that Chris is working in his office one night, when he decides to go use the restroom down the hall. Improbably enough, there is a hungry bear roaming the halls at the moment. If he knew of it, Chris would strongly prefer to avoid the bear. Unfortunately, he has no evidence and no suspicion that a bear might be in the hallway. In this situation, is there a *reason* for Chris to avoid the hallway?⁹

Either answer to this question has some linguistic plausibility. Let us say that there is an 'objective reason' for Chris to avoid the hallway, but there is no 'subjective reason' for Chris to avoid the hallway. A subjective reason for S to do A is a fact that suffices, given S's preferences and values, to make it *pro tanto* rational for S to do A. If Chris knew of the bear in the hallway, then the fact *that Chris knew that there was a bear in the hallway* (together, perhaps, with some of Chris' background knowledge concerning bears, hallways, and so on) would constitute a subjective reason for Chris to avoid the hallway. An *objective reason* for S to do A is a fact that is not itself a subjective reason for S to do A but is such that, if S knew it, then the fact that S knew it would be a subjective reason for S to do A. In the example, *that there is a bear in the hallway* (together, perhaps, with certain general facts about the nature of bears, hallways, and so on) is an objective reason for Chris to avoid the hallway. This fact exists as an objective reason whether or not Chris knows the fact.

The thesis of lexical priority could be stated either in terms of subjective reasons or in terms of objective reasons. For purposes of the following discussion, I will take lexical priority views to apply at least to the subjective reasons that an agent would have in the (perhaps unlikely or impossible) case where the agent has absolutely certain knowledge of some objective reason. That is, I take lexical priority views to be committed to the following, which I label the Lexical Priority Thesis (hereafter, Lexical Priority):

Lexical Priority: There exists some condition C and some good G such that, if one knows for certain that a potential action satisfies C, then one has a subjective reason for (or against) performing the action, which is of greater force than any reason for action one could have in virtue of knowing that an action would promote any quantity of G.

2.2. VARIETIES OF LEXICAL PRIORITY VIEW

Absolutist theories in ethics are one species of lexical priority view. The good of enjoyment is typically taken as a lesser good in these theories. On Kant's view, our reason for avoiding the breaking of promises always takes precedence over our reason for promoting enjoyment, regardless of how much enjoyment might be promoted (or suffering avoided) by the breaking of a promise. Similarly, on Anscombe's view, our reason for avoiding the punishment of the innocent takes precedence over our reason for promoting any amount of enjoyment, and on Nozick's view, our reason for respecting the rights of others takes precedence over our reason for promoting any amount of enjoyment.

Lexical priority views may be held in consequentialist, as well as deontological forms. John Stuart Mill is often credited with a kind of lexical priority view: on the standard reading, Mill thought that high quality, intellectual pleasures are always superior to merely sensory pleasures, regardless of quantity.¹⁰ Thus, our reason for promoting high quality pleasure would take precedence over our reason for promoting any amount of lower quality pleasure.

Derek Parfit offers a contemporary variation on this theme, suggesting that any amount of 'the best things in life' outweighs any amount of lower-quality benefits. He calls this view 'Perfectionism':

[E]ven if some change brings a great net benefit to those who are affected, it is a change for the worse if it involves the loss of one of the best things in life.¹¹

The surrounding context shows that Parfit intends a commitment to Lexical Priority:

[U]nless we can defend this view, any loss of quality could be outweighed by

a sufficient gain in the quantity of lesser goods.¹²

As an example, he suggests that the loss of Mozart's music would be a loss that could not be compensated for by any quantity of sub-Mozart goods.

Many other philosophers have advanced axiological theories with a lexical priority structure. Stuart Rachels has argued that a moment of ecstasy is superior to any duration of mild, 'muzak and potatoes' pleasure. Cardinal Newman thought that the commission of a single, venial sin was worse than any amount of worldly suffering. And some advocates of the 'sanctity of life' argue that life has infinite value, and thus that death is worse than any amount of suffering.¹³

Lexical priority views, then, have been defended in a variety of forms, by thinkers with a wide variety of ethical perspectives. Nor is the Lexical Priority Thesis confined to ethical reasons; Lexical Priority may be asserted for any kind of practical reason. Thus, the aforementioned views of Mill, Parfit, and Rachels presumably apply equally to our prudential reasons for pursuing certain goods for ourselves as they do to our ethical reasons for pursuing those goods for others. Mill, for instance, would presumably have held that one's prudential reason for promoting one's own higher pleasures is lexically prior to one's prudential reason for promoting one's own lower pleasures.

2.3. THE RISK PARADOX FOR LEXICAL PRIORITY VIEWS

Suppose an agent must choose between two actions, A and B. A would promote a large quantity of some lesser good G but would transgress against some higher, objective reason R. B would satisfy R but would fail to promote G. There are no other relevant practical reasons. What should the agent do? If the agent knows for certain that the situation is as described, then Lexical Priority delivers an unequivocal verdict: the agent should choose B.¹⁴

But what if matters are uncertain? Suppose the agent knows that A would definitely promote a large quantity of G, and suppose there is a probability p , on the agent's evidence, that A would violate R.¹⁵ B would certainly not violate R but would also certainly not promote G. What should the agent do in this case? There seem to be three natural answers a lexical priority theorist might give to this question:

- a. The *Zero Risk Tolerance View*: As long as $p > 0$, the agent should choose B, regardless of the quantity of G that would be produced by doing A.
- b. The *Maximal Risk Tolerance View*: As long as $p < 1$, the agent should do A, provided the quantity of G to be produced is sufficiently large.¹⁶
- c. The *Risk Threshold View*: There is some threshold probability t , strictly between 0 and 1, such that if $p > t$, then the agent should do B, regardless of the quantity of G that would be produced by doing A; but if $p < t$, then the agent should do A as long as the quantity of G to be produced is sufficiently

large.

The lexical priority theorist's answer might vary according to the details of the particular case. Perhaps, that is, in some kinds of cases, answer (a) would be appropriate, while in others (b) would be appropriate, and so on. Nevertheless, if we fix all the morally relevant features of a case other than the value of p , then one of the above three views should apply in any given type of case.

As we saw in section 1, all three answers, (a), (b), and (c), are problematic. Zero Risk Tolerance leads to implausibly restrictive practical theories. Maximal Risk Tolerance defeats the point of positing a lexically superior practical reason. And Risk Thresholds lead to violations of one of the following plausible principles:

Two Rights Don't Make a Wrong: If it is appropriate for S to do A whether or not S does B , and it is appropriate for S to do B whether or not S does A , then it is appropriate for S to do A and B .¹⁷

Two Wrongs Don't Make a Right: If it is inappropriate for S to do A whether or not S does B , and it is inappropriate for S to do B whether or not S does A , then it is inappropriate for S to do A and B .

The intended notion of 'appropriateness' is a broad notion, referring to what is overall consonant with one's practical reasons of a given kind. It is meant to capture both the relation of moral permissibility and the relation of an agent to the action that is the agent's best option. Thus, the Two Rights Don't Make a Wrong principle (hereafter, Two Rights) is intended to imply both (1) that if it is permissible to do A (whether or not one does B) and permissible to do B (whether or not one does A), then it is permissible to do A and B ; and (2) that if it is best to do A (whether or not one does B) and best to do B (whether or not one does A), then it is best to do A and B . Similarly for the Two Wrongs principle.

In section 1, we saw an illustration of how the problem of risk applies to an absolutist ethical theory. Consider now an example aimed at advocates of an *axiological* lexical priority view—specifically, one aimed at Perfectionists. Suppose you have a large sum of money that you intend to donate to a charitable cause. Your only concern is that the money should do the most good. Two causes have attracted your attention. On the one hand, you could donate the money to a poverty relief organization, which you know will result in a large improvement to the welfare of many people who presently have only barely worthwhile lives.¹⁸ On the other hand, you could donate the money to an art school to enable it to expand its operations searching for and supporting unrecognized artistic talent. In the latter event, there is a probability p that, as a result of your support, a new artistic masterpiece would be produced. It seems that, on Parfit's Perfectionist view, if $p = 1$, it would be best to donate to the art school, regardless of how

many people would be aided by the poverty relief organization.¹⁹ For what values of p would this remain true? Three answers naturally suggest themselves:

(a) On the Zero Risk Tolerance View, one should donate to the art school as long as there is a nonzero chance that doing so will result in the production of an artistic masterpiece. There is a nonzero probability of inducing *any* human being to produce an artistic masterpiece through artistic training and, if necessary, paying them to attempt to produce works of art. For most human beings, this probability is very small; for some, vanishingly small. Nevertheless, it is not zero. Since this endeavor is capable of absorbing all available charitable funds, it seems that all charitable donations should be devoted to art—or more precisely, to the endeavor to produce the best things in life, whatever these things may be—and none to raising the quality of life for those whose lives are barely worth living. Likewise, if funding for the arts is within the proper functions of government, perhaps governmental budgets should devote all funds to the encouragement of art and other potential ‘best things in life,’ except for the minimum that must be spent on such other functions as police, roads, and so on, to enable great art to be efficiently produced.²⁰

(b) On the Maximal Risk Tolerance View, one should donate to the anti-poverty organization (assuming that it will do a sufficient amount of good), provided that the art school donation has less than a 100% probability of leading to an artistic masterpiece. Needless to say, no action is ever 100% certain, *ex ante*, to aid in the production of an artistic masterpiece. Nor can any action be *ex ante* certain either to preserve or to destroy an artistic masterpiece. So the alleged infinite value of these best things would never be relevant to practical deliberation.

(c) On the Risk Threshold View, one should donate to the anti-poverty organization if the probability that the art school donation would lead to production of an artistic masterpiece is less than the relevant threshold value, t ; otherwise, one should donate to the art school.

But now consider a slightly different example. You have come into \$20,000, which you plan to use for charitable purposes. There is an international aid organization, which would use any money you give them to greatly improve the lives of many impoverished people, with benefits directly proportional to the size of your gift. On the other hand, there are two struggling artists who each need \$10,000 of support to continue their work; otherwise, they will take jobs at Wal-Mart and give up art. After meeting both artists, you correctly judge that each artist would have a probability of slightly less than t of producing an artistic masterpiece as a result of your support. You are deciding whether to support both artists, to give all the money to the poor, or to support one artist and give the rest of the money to the poor. On the one hand, it seems that you should support both artists, since doing so has a probability greater than t of producing an artistic masterpiece, and this outweighs any quantity of the lesser goods that would be produced by aiding the poor. On the other hand, if we view supporting the artists as making two distinct gifts, one to the first artist and one to the second, then it

seems that these two actions are each misguided, for each gift has a probability less than t of producing an artistic masterpiece, and this potential benefit is outweighed by the large quantity of lesser goods produced by the poverty relief organization. The threshold version of Perfectionism thus violates the Two Wrongs Don't Make a Right principle.

As the criminal justice and charitable donation cases we have considered illustrate, lexical priority theories that incorporate a probability threshold may violate either the Two Rights principle or the Two Wrongs principle: theories that posit a lexically prior reason *against* certain actions will violate the Two Rights principle, whereas theories positing a lexically prior reason *for* certain actions will violate the Two Wrongs principle.

3. Responses to the Problem of Risk

3.1. THE DOCTRINE OF DOUBLE EFFECT

The implausibility of the Zero Risk Tolerance View turns on the fact that almost every logical possibility has a nonzero probability. If, however, a lexical priority theory assigns absolute weight only to certain kinds of features of an action that in fact many actions have *no* chance of possessing, then the Zero Risk Tolerance View might be made plausible. For this reason, an absolutist version of the Doctrine of Double Effect (hereafter, the DDE) might be thought to afford an acceptable treatment of the problem of risk. Suppose one holds that it is always wrong to *intentionally* harm the innocent, regardless of the benefits that could be produced by doing so; but it is not always wrong to harm the innocent as a *foreseen but not intended* side effect of an action aimed at some desirable end.²¹ This principle is meant to explain such things as why it is wrong to intentionally target civilians in war, even when doing so is expected to produce a much greater good, yet it is not wrong in the same way to attack military targets, even when doing so can be expected to cause some civilian casualties as a side effect.²² It might be thought that, even if we cannot know for certain what outcomes will occur in the external world, we can nevertheless know for certain what our own intentions are. Consequently, we can often know for certain that our actions do not violate the Doctrine of Double Effect. For instance, a military pilot may not know whether he is killing civilians, but he may know for certain that, in bombing a particular target, he does not *intend* to kill civilians. The DDE thus seems to enable zero-risk-tolerance absolutists to avoid taking an absurdly demanding ethical stance.

As Jackson and Smith observe, this solution to the problem of risk is not so straightforward. It is doubtful at best that individuals can really be absolutely certain of their own intentions.²³ Perhaps there are such things as unconscious intentions. If there are, then perhaps an agent who takes himself to have no such intention may have an unconscious intention to harm the innocent. One need not accept that unconscious intentions actually exist to feel the force of this objection.

One need only accept that there is a nonzero probability, on most agents' evidence, that unconscious intentions exist, as well as a nonzero probability, given that such things exist, that the agent himself in performing a given act would have an unconscious intention to harm the innocent. This problem might, however, be dealt with by modifying the DDE so that only *conscious* intentions to harm the innocent count as violating the principle.

A second objection to the proposed use of the DDE is more serious. Assume that while it is absolutely impermissible to intentionally harm the innocent, it is permissible, given sufficiently desirable consequences, to intentionally harm those who are not innocent (for instance, those who have wronged others or who are wrongfully trying to harm others). Suppose that one could produce extremely desirable results by harming a person who has a probability p of being innocent. For what values of p would it be permissible to intentionally harm this person?

The proponent of the DDE might give any of three familiar answers: it is permissible only when $p = 0$; it is permissible when $p < 1$; or it is permissible when $p < t$, for some t strictly between 0 and 1. But none of these answers affords an escape from the problem of risk as formulated in previous sections. Is there some other approach open to proponents of the DDE that avoids the problem of risk?

Jackson and Smith consider the following principle as one that might be advanced by a proponent of the DDE:

It is absolutely forbidden to perform an action under the description 'intentionally killing someone innocent.'

Jackson and Smith find difficulty in interpreting this principle, but ultimately they conclude that the principle fails.²⁴ I shall not discuss the merits of this principle, both because its meaning escapes me and because I doubt that partisans of the DDE would advance it. In my view, actions are simply performed or not performed; an action is not *performed under a description*.

An action may, however, be said to *be intentional* under some descriptions, and not intentional under other descriptions.²⁵ This suggests that an advocate of the DDE might advance the following principle:

An action is wrong whenever it is intentional under the description, 'harming someone innocent.'

An action will not violate this deontological proscription merely by being an intentional harming of someone *who in fact* is innocent. Nor will it violate the proscription merely by being an intentional harming of someone *who is believed* to be innocent. Rather, it is the entire phrase, 'harming someone innocent,' that characterizes what the agent who violates the principle is trying to do. Exactly what this means depends on the correct theory of intentional action. We may consider three accounts of intentional action:

- I1 An action is intentional under the description ‘ *ing*’ if and only if the agent at the time intends that his action be a case of *ing*.
- I2 An action is intentional under the description ‘ *ing*’ if and only if the action would have to be a *ing* in order for the agent’s plans to be fully realized.
- I3 An action is intentional under the description ‘ *ing*’ if and only if *that the act would be a case of ing* is part of the agent’s reason for performing it.

Do any of these accounts, combined with the proposed version of the DDE, result in a plausible ethical theory?

To answer this, consider a well-known example:

Terror Bomber: A military pilot plans to bomb a civilian population center, as a means of frightening the enemy into surrendering. If he proceeds, it is virtually certain that many innocent non-combatants will be killed, but also that many more lives will be saved through the shortening of the war.

This case is used in the literature as a paradigm instance in which an agent violates the Doctrine of Double Effect.²⁶ Presumably, then, any reasonable interpretation of the DDE must condemn the terror bombing.

But suppose that the pilot in this case, after proceeding with the bombing, defends himself as follows:

It is true that I intentionally killed many people whom I believed to be innocent and who were in fact innocent. But my action was not intentional under the description “harming the innocent,” for, first, I did not specifically intend that my action be a harming of *innocent* people; rather, I was simply indifferent as to whether the people killed were innocent or not. Second, for my plans to be realized, it was not necessary that any innocent people be harmed. If, improbably enough, all the inhabitants of the city had been enemy combatants, my plans would have been just as fully realized as they in fact were. Third, the fact that my action would harm innocent people was no part of my reason for action; my reason for bombing the city would have applied equally if all the inhabitants had been combatants--either way, the enemy would have been frightened into surrendering.

This defense must fail. Presumably, it misconstrues the DDE. Presumably, proponents of the DDE would reply to the pilot roughly as follows:

To count as violating the DDE, it is not necessary that you intend specifically that the people you harm *be innocents*. It is enough that you intentionally harm certain people, where you have good reason to believe those people to be innocent. Since you intentionally harmed the inhabitants of the city below you, and you had good reason to believe those people to be innocent, you acted contrary to the DDE.

But this reply repudiates the interpretation of the DDE on which actions only violate the DDE when they are intentional under the description 'harming the innocent.' Once we clarify the DDE in this manner, the question arises as to what must be the probability that a particular individual is innocent, for it to be impermissible to intentionally harm that individual regardless of the benefits. And so the problem of risk must again be confronted.

3.2. DENYING ETHICAL INDEPENDENCE

Recall the Two Rights Don't Make a Wrong principle enunciated earlier:

Two Rights: If it is appropriate for S to do A whether or not S does B, and it is appropriate for S to do B whether or not S does A, then it is appropriate for S to do A and B.

The clauses 'whether or not S does B' and 'whether or not S does A' are included to avoid such counterexamples as the following:

Bomb Squad: Susan works for the bomb squad. She has just been called in to defuse a bomb. She knows that if she cuts only the red wire, the bomb will be defused. Also, if she cuts only the blue wire, the bomb will be defused. But if she cuts both wires, the bomb will explode, killing several innocent people, including Susan.

Here, it is appropriate, both ethically and prudentially, to cut the red wire, and it is equally appropriate to cut the blue wire, but it is inappropriate to cut the red wire and the blue wire. The problem is that there is a special relationship between the two actions, such that whether one performs the first action makes a difference to the effects of performing the other. This example refutes the simple claim that if it is appropriate to do A, and it is appropriate to do B, then it is appropriate to do both A and B. But the example does not refute the Two Rights principle stated above, for in Bomb Squad, it is not true that it is appropriate to cut the red wire *whether or not* one cuts the blue wire. Rather, if Susan cuts the blue wire, she must not cut the red wire. The 'whether or not' clauses in Two Rights serve to ensure that actions A and B are suitably independent of each other. Given the sort of independence stipulated by Two Rights, it should be acceptable to combine two appropriate actions.

An advocate of Lexical Priority might seek to reconcile a risk threshold view with the Two Rights principle by denying the ethical independence of actions that jointly contribute to creating unacceptable risks. The risk threshold t might be taken as restricting not only what simple, individual actions one may perform, but also what sequences of acts one may perform: one may not perform a sequence of actions that creates a risk above the threshold. Thus, suppose that A and B

each create a risk slightly below threshold t of doing something which one is absolutely prohibited from doing knowingly. One might argue that B is permissible as long as one does not do A, but that B is impermissible given that one does A, because given that one does A, doing B will push one over the acceptable level of risk.²⁷

This approach may seem plausible in cases involving *pairs* of actions, each of which creates a risk slightly below t . But whatever plausibility the approach may have in such cases evaporates once we turn our attention to cases involving policies that subsume large numbers of actions. Assume that there is an absolute prohibition on punishing the innocent, with a risk threshold of t . And consider a policy of punishing defendants in criminal trials whenever the probability of the defendant's being innocent is less than $.001t$. This would seem a more than sufficiently cautious policy. But if the policy is applied in millions of criminal trials, then the sequence of trials will almost certainly far exceed the acceptable risk of punishing someone innocent. On the view in question, then, the implementation of this policy is absolutely prohibited. In general, the view will prohibit almost any criminal justice system for a large society—unless either the standard of proof for individual defendants or the risk threshold is set absurdly high.

The view may also lead to implausible results when we consider series of actions involving more than one agent. The core idea of the approach is that we must avoid an action that, given some set of other actions that have occurred (or that one knows will occur?), pushes one over the risk threshold. Must these other actions be performed by the same agent, or should we also include actions performed by other agents? Suppose that we include actions performed by other agents. Then I might be prohibited from doing A because someone else performed a risky action yesterday. Turning again to the case of criminal justice, the appropriate standard of proof for criminal trials would increase dramatically as more trials were conducted—to decide whether to convict a given defendant, one would have to take into account the total risk of punishing the innocent created by all other juries. To stay under the acceptable risk threshold, we would have to rapidly increase the standard of proof toward absolute certainty.

But suppose we exclude other agents—in other words, suppose that a given agent need only take into account the total risk created by his own actions. In that case, the appropriate standard of proof in a criminal trial might depend upon who is on the jury. One jury might, for instance, need only a 95% probability of guilt, but another jury—because it convicted someone else with a 4% chance of innocence in another trial—might need a 99% probability of guilt for this trial. Different members of the same jury might even have different appropriate standards, depending on how many juries they had previously served on and what transpired in those trials. These sorts of consequences illustrate why the approach of denying ethical independence is not a promising one.

3.3. INDIVIDUAL RIGHTS

Consider another version of Risk-Threshold Absolutism. On this view, there is a threshold probability, t , such that it is absolutely impermissible to impose on any one individual a risk greater than t of suffering a rights-violation; however, when the probability of violating a given individual's rights is less than t , then consequentialist reasons may sometimes justify the risk. I assume that, among other things, individuals have a right not to be punished for crimes they did not commit. So one may not perform an action that imposes on any particular person a probability greater than t of being punished for a crime he did not commit. One may sometimes perform an action or series of actions that creates a probability greater than t that *some person or other* will suffer a rights-violation, as long as there is no specific person on whom a risk greater than t is imposed.²⁸

This position seems to reconcile a threshold version of absolutism with the Two Rights Don't Make a Wrong principle. Where A and B each create a risk slightly below t of violating someone's rights (but not the same person for the two actions), and each has sufficiently good consequences, A and B are each permissible. The conjunctive action of performing both A and B will also be permissible, just as Two Rights requires, because although there will be a probability greater than t that at least one person's rights will be violated, there will be no particular person with a probability greater than t of suffering a rights violation.

The reconciliation here is imperfect. For suppose that act A creates a risk slightly below t of violating someone's rights, while act B creates a risk slightly below t of violating the same person's rights, where the two rights-violations are probabilistically independent of each other. Then each of A and B might be permissible, provided the consequences are sufficiently desirable, while the combination of A and B would be absolutely prohibited, since in this case there is a single person who would have a risk greater than t of suffering a rights-violation. Despite this difficulty, one might feel that at least some progress has been made. This sort of failure of agglomeration may seem less troubling, because cases in which we have to consider multiple actions imposing risks on the same person are rarer than cases in which we must consider series of actions imposing risks on different people. The criminal justice system, for instance, has a near 100% probability of punishing some innocents every year, but only a small probability of imposing undeserved punishment on any given individual.

But the individual-rights-based reconciliation faces further difficulties. The most natural understanding of the position is as follows:

IR It is absolutely impermissible to perform A if, for some individual x , there is a probability greater than t that A will violate x 's rights. However, if there is no x about whom this is true, then A may be justified by sufficiently desirable consequences.

For purposes of illustration, suppose that t , the maximum allowable risk of violating individual rights, is 10%. Assume also that, in the conduct of a just war, the killing of non-combatants is a rights-violation, but the killing of enemy combatants is not a rights-violation. We may then apply IR to the following scenario.

War Options: Joe is a soldier fighting a just war. He knows that there are three tactics he can take, any one of which would end the war and save many times more lives than it cost.

Tactic 1: Joe may shoot the brother of the enemy country's leader. This will frighten the leader into surrendering. The brother has a 10.1% probability of being an innocent non-combatant.

Tactic 2: Joe may bomb a certain building that he knows to contain 500 people. There is a 10% probability that all of these people are innocent non-combatants, and a 90% probability that all are enemy combatants.

Tactic 3: There are ten large cities in the enemy country, each containing one million innocent non-combatants. Joe may program a computer to randomly select one of the ten cities and launch a nuclear missile that will destroy the city.

To repeat, any of these tactics would end the war, saving many times as many lives as it destroyed; the disadvantages are that the first tactic creates a 10.1% probability of killing one innocent person, the second creates a 10% probability of killing 500 innocent people; and the third creates a 100% probability of killing one million innocent people, but which people will be killed will be a matter of chance.

What is the rational preference ranking among these options? Intuitively, Tactic 1 is least objectionable, followed by Tactic 2, followed by Tactic 3. However, IR implies that Tactic 2 is preferable to Tactic 1, since Tactic 1 creates a probability greater than 10% of violating a specific individual's rights, whereas Tactic 2, while threatening 500 times as many people, does not exceed the allowable threshold probability of 10%. Even more absurdly, Tactic 3 would be judged preferable to Tactic 1, because any one city would have only a 10% probability of being targeted by the computer; thus, Tactic 3 imposes only a 10% risk of death on any given innocent person. So if Joe had only Tactics 1 and 3 available, he would have to select Tactic 3.

The scenario could easily be modified to take account of different possible values of t . I have assumed, as seems plausible, that the killing of non-combatants violates their rights. But analogous counter-examples could be devised for most other views of what rights individuals have. For instance, suppose an advocate of IR holds that torturing a person is a rights-violation. We may then imagine a scenario in which one chooses between an action that has a 10.1% probability of torturing someone and an action that randomly selects one out of ten groups of one million people to torture; in either case, stipulate that extremely large benefits

would result. In general, the problem is that IR entails that it can be preferable to perform an action that will definitely violate many people's rights, rather than an action that might violate a single person's rights in the same manner, provided that the former action selects its victims randomly from many potential groups of victims.

4. Conclusion

Risk poses a philosophical problem for a wide class of theories of practical reasons. Lexical Priority theories hold that there is some non-trivial condition C such that, when an available course of action is known to satisfy C , then the agent has a subjective reason for (or against) the action that outweighs any amount of lesser goods. Any such theory invites questions about the subjective reason provided by the existence of a non-extreme probability that an action satisfies C . The defender of a lexical priority view may answer, (i) that there is a subjective reason of infinite weight relative to the value of some quantity of lesser goods, only when a potential action is *sure* to satisfy C ; (ii) that a reason of infinite weight exists when a potential action has any nonzero chance of satisfying C ; or (iii) that a reason of infinite weight exists when a potential action passes some threshold probability, t , of satisfying C .

Each of these alternatives is problematic. For most values of C , the first alternative renders the theory of little interest, as very little is ever absolutely certain. The second alternative renders the theory overly restrictive, or overly anti-consequentialist, as almost every logical possibility has a nonzero probability. Finally, the third alternative creates the possibility of paradoxical situations in which each of a set of actions is justified (unjustified), irrespective of whether the other actions are performed, yet the combination of actions is unjustified (justified). This is because a set of actions, each of which individually has a probability slightly less than t of satisfying some condition, will typically have a probability greater than t of containing at least one member satisfying that condition.

Some lexical priority theorists might seek to avoid the problem of risk by arguing that, because of the special nature of a particular condition C , we often *are* absolutely certain that an action does or does not satisfy C . This might enable the theory to avoid the twin threats of uninterestingness and excessive restrictiveness. However, it is not clear whether this strategy succeeds for any actual, widely-held theory.

The most straightforward way of avoiding the problem of risk is to adopt a non-lexical (Archimedean) theory of practical reasons. This is a theory on which any reason can be outweighed by a sufficient quantity of any kind of good. This includes most consequentialist theories as well as moderate deontological theories. It includes, for example, Bentham's utilitarianism, as well as Ross' pluralistic deontology.²⁹ As long as we avoid the claim of infinite weight for some

practical reasons, there is a natural way of dealing with risk: if an action's certainly satisfying C gives one a practical reason of some finite weight W, then the existence of a probability p of an action's satisfying C gives one a practical reason of weight $p \times W$ (where the weight of a practical reason may be measured by the quantity of some simple, aggregative good required to outweigh that practical reason).

This is compatible with maintaining that justice, respect for individual rights, and other non-utilitarian values are of great importance, outweighed only in extreme circumstances. Though some deontologists would remain unsatisfied with this moderate position, it is unclear what evidence supports an absolutist position, as against the moderate position that avoids the problem of risk.³⁰

Department of Philosophy
University of Colorado

NOTES

1. Kant 1959, p. 17.
2. Anscombe 1981, pp. 39-40.
3. Nozick 1974, pp. 32-3; but see also p. 30n, where he comes close to taking back the absolutism he elsewhere seems to endorse. For further absolutist sentiments, see Anscombe 1981, pp. 64-6; Donagan 1977, pp. 152-7; Nagel 1972; and Gewirth 1984. For absolutist theories in political philosophy, see Dworkin 1984 and Rawls 1999, pp. 37-9, 72.
4. On the elusiveness of certainty, see Unger 1971; Russell 1960, p. 278.
5. Friedman 1989, pp. 167-9; Jackson and Smith 2006, p. 272.
6. This kind of argument appears in Von Magnus 1983, p. 25; and Jackson and Smith 2006, p. 276. See also Simons 1999, p. 63, for a similar point. See Aboodi et al. 2008 and Hawley 2008 for replies.
7. Jackson and Smith use an example involving killing a skier who poses a mortal threat to several other people's lives, when one is not sure whether the skier is creating the threat intentionally. Jackson and Smith assume that the absolutist would hold it wrong to kill the skier if the threat is unintentional; many absolutists, however, would demur (Anscombe 1981, p. 67; Nozick 1974, pp. 34-5; Donagan 1977, pp. 162-3). I focus on the criminal trial example in the text because the prohibition against punishing the innocent is much less controversial.

8. I take the term 'lexical priority' from Rawls (1999, pp. 37-9). Lexical priority views are also called *non-Archimedean* theories, because they deny that the weights of practical reasons have what mathematicians call the Archimedean property.
9. I owe a variant of this example to Chris Heathwood.
10. Mill 1998, p. 56 (sect. 2, para. 5). For the standard reading of this passage, see Brink 1992, pp. 72, 92; Crisp 1997, p. 29; Riley 1993. However, Booher (2007) and Schmidt-Petri (2003) question the standard reading of Mill.
11. Parfit 2004, p. 19.
12. Parfit 2004, p. 20. It is the non-Archimedean character of Perfectionism that Parfit hopes to use to avoid the Repugnant Conclusion.
13. On muzak and potatoes, see Rachels 2001, pp. 215-16; cf. Parfit 2004, pp. 17-18. On venial sins, see Newman 1969, pp. 239-40. On the sanctity of life, see the remarks of Dr. Moshe Tendler, quoted by Brody (1976, p. 66); and Rabbi Jakobovits, quoted by Heenan (1974, p. 7).
14. For expressive convenience, I assume that what an agent has most reason to do in a given situation is identical with what the agent 'should' do.
15. Hereinafter, I interpret probability in an epistemic sense; thus, all probabilities are to be taken as probabilities in the light of the relevant agents' evidence. The same goes for 'risks' and 'chances.'
16. This leaves open the question of how large the quantity of G must be. For deontological absolutists, it would be open to maintain that, even in the case where one's reason to avoid a risk of violating R is overridable, the benefits must still be much larger than the costs, i.e., that consequentialist reasoning is not correct even in such cases. For the Maximal Risk Tolerance View, it is only when p reaches 1 that it becomes impossible for *any* quantity of G to outweigh the force of reason R.
17. The Two Rights principle is comparable to axiom A2 of von Wright's (1956, p. 509) deontic logic. Jackson (1985) poses objections to the related principle in deontic logic that if $O(A)$ and $O(B)$, then $O(A\&B)$. However, Jackson's counterexamples do not apply to my Two Rights principle.
18. Because Parfit (2004, p. 20) leaves open the possibility that relieving great suffering is a goal on a par with preserving the best things in life, I assume that the people helped by the anti-poverty charity are not suffering greatly; rather, they simply have a quality of life barely above zero. Parfit's use of Perfectionism to avoid the Mere Addition argument for the Repugnant Conclusion implies that

helping such people might be of infinitesimal value compared with the value of the best things in life.

19. I assume, following Parfit's Mozart example, that the artistic masterpiece would count as one of the best things in life. I assume also that the production of such an object is of equal value to the preservation of an existing object of this kind.

20. This conclusion does not strictly follow from Perfectionism, since we cannot rule out that some deontological principle prohibits this, nor that there is some other good that is also infinitely more important than assisting the needy. Nevertheless, the claim that any nonzero chance of producing an artistic masterpiece outweighs any increase in any number of people's well-being that does not involve production of one of the best things in life, must surely have profound and revisionary implications for how one ought to behave.

21. See Aquinas 1920, 2a2ae, Q. 64, art. 7; Anscombe 2001; Nagel 1986, pp. 179-85. I here pass over the debates as to the precise formulation of the doctrine; see Quinn 1989; Boyle 1991; Davis 2001; Marquis 2001. The doctrine does not, of course, dictate that actions causing merely foreseen harms fail thereby to be *prima facie* wrong. Rather, such actions are *prima facie* wrong, but they can nevertheless be justified by the production of a sufficiently greater good. In contrast, actions that *intend* harm to innocent others cannot be justified by the production of a greater good.

22. Anscombe 1981, p. 66; Quinn 1989, p. 336.

23. Jackson and Smith 2006, p. 279; Kant 1959, p. 23.

24. Jackson and Smith 2006, pp. 281-2.

25. Davidson 1980, p. 46.

26. Quinn 1989, p. 336; McIntyre 2009. Cf. Anscombe's (1981, pp. 62-71) condemnation of similar real-world actions.

27. This is intended as an approximation of Hawley's (2008) view in his response to Jackson and Smith.

28. This is, as I understand it, the position advanced by Aboodi et al. (2008) in response to Jackson and Smith.

29. Bentham 1996; Ross 1988, ch. 2.

30. I would like to thank Eric Chwang, Jackie Colby, and the reviewers at *Pacific Philosophical Quarterly* for valuable comments on earlier versions of this paper.

REFERENCES

- Aboodi, Ron, Adi Borer, and David Enoch. 2008. 'Deontology, Individualism, and Uncertainty: A Reply to Jackson and Smith,' *Journal of Philosophy* 55: 259-72.
- Anscombe, G. E. M. 1981. *Ethics, Religion, and Politics: The Collected Philosophical Papers of G. E. M. Anscombe*, vol. 3. Minneapolis, Minn.: University of Minnesota Press.
- Anscombe, G. E. M. 2001. 'Medalist's Address: Action, Intention and "Double Effect,"' pp. 50-66 in P. A. Woodward, ed., *The Doctrine of Double Effect: Philosophers Debate a Controversial Moral Principle*. Notre Dame, Ind.: University of Notre Dame Press.
- Aquinas, Thomas. 1920. *The Summa Theologica of St. Thomas Aquinas*, second and revised edition, tr. by Fathers of the English Dominican Province. London.
- Bentham, Jeremy. 1996. *An Introduction to the Principles of Morals and Legislation*, ed. J.H. Burns and H.L.A. Hart. Oxford: Clarendon.
- Booher, Troy. 2007. 'J.S. Mill's Test for Higher Pleasure,' *Studies in the History of Ethics*. <<http://www.historyofethics.org/122007/122007Booher.shtml>>.
- Boyle, Joseph. 1991. 'Who Is Entitled to Double Effect?' *Journal of Medicine and Philosophy* 16: 475-94.
- Brink, David. 1992. 'Mill's Deliberative Utilitarianism,' *Philosophy and Public Affairs* 21: 67-103.
- Brody, Howard. 1976. *Ethical Decisions in Medicine*. Boston, Mass.: Little, Brown and Co.
- Crisp, Roger. 1997. *Mill on Utilitarianism*. London: Routledge.
- Davidson, Donald. 1980. *Essays on Actions and Events*. Oxford: Clarendon.
- Davis, Ann. 2001. 'The Doctrine of Double Effect: Problems of Interpretation,' pp. 119-42 in P. A. Woodward, ed., *The Doctrine of Double Effect: Philosophers Debate a Controversial Moral Principle*. Notre Dame, Ind.: University of Notre Dame Press.
- Donagan, Alan. 1977. *The Theory of Morality*. Chicago: University of Chicago Press.
- Dworkin, Ronald. 1984. 'Rights as Trumps,' pp. 153-67 in Jeremy Waldron, ed., *Theories of Rights*. Oxford: Oxford University Press.
- Friedman, David. 1989. *The Machinery of Freedom*. LaSalle, Ill.: Open Court.
- Gewirth, Alan. 1984. 'Are there any Absolute Rights?' pp. 91-109 in Jeremy Waldron, ed., *Theories of Rights*. Oxford: Oxford University Press.
- Hawley, Patrick. 2008. 'Moral Absolutism Defended,' *Journal of Philosophy* 55: 273-5.
- Heenan, Cardinal John. 1974. 'A Fascinating Story,' pp. 1-7 in *The Hour of Our Death: A Record of the Conference on the Care of the Dying Held in London, 1973*, edited by Sylvia Lack and Richard Lamerton. London: Geoffrey Chapman.
- Jackson, Frank. 1985. 'On the Semantics and Logic of Obligation,' *Mind* 94: 177-

- Jackson, Frank and Michael Smith. 2006. 'Absolutist Moral Theories and Uncertainty,' *Journal of Philosophy* 103: 267-283.
- Kant, Immanuel. 1959. *Foundations of the Metaphysics of Morals*, tr. Lewis White Beck. Indianapolis, Ind.: Bobbs-Merrill.
- Marquis, Donald. 2001. 'Four Versions of Double Effect,' pp. 156-85 in P. A. Woodward, ed., *The Doctrine of Double Effect: Philosophers Debate a Controversial Moral Principle*. Notre Dame, Ind.: University of Notre Dame Press.
- McIntyre, Alison. 2009. "Doctrine of Double Effect," *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), <<http://plato.stanford.edu/archives/fall2009/entries/double-effect/>>.
- Mill, John Stuart. 1998. *Utilitarianism*, ed. Roger Crisp. Oxford: Oxford University Press.
- Nagel, Thomas. 1972. 'War and Massacre,' *Philosophy and Public Affairs* 1: 123-44.
- Nagel, Thomas. 1986. *The View from Nowhere*. Oxford: Oxford University Press.
- Newman, John H. 1969. *Certain Difficulties Felt by Anglicans in Catholic Teaching* vol. 1. Westminster, Md.: Christian Classics. From the lectures of 1850.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.
- Parfit, Derek. 2004. 'Overpopulation and the Quality of Life,' pp. 7-22 in *The Repugnant Conclusion: Essays on Population Ethics*, ed. Jesper Ryberg and Torbjörn Tännsjö. Dordrecht, the Netherlands: Kluwer.
- Quinn, Warren. 1989. 'Actions, Intentions, and Consequences: The Doctrine of Double Effect,' *Philosophy and Public Affairs* 18: 334-51.
- Rachels, Stuart. 2001. 'A Set of Solutions to Parfit's Problems,' *Noûs* 35: 214-38.
- Rawls, John. 1999. *A Theory of Justice*, revised edition. Cambridge, Mass.: Harvard University Press.
- Riley, Jonathan. 1993. 'On Quantities and Qualities of Pleasure,' *Utilitas* 5: 291-300.
- Ross, W.D. 1988. *The Right and the Good*. Indianapolis, Ind.: Hackett.
- Russell, Bertrand. 1960. *An Outline of Philosophy*. Cleveland, Ohio: Meridian.
- Schmidt-Petri, Christoph. 2003. 'Mill on Quality and Quantity,' *Philosophical Quarterly* 53: 102-4.
- Simons, Kenneth. 1999. 'Negligence,' *Social Philosophy and Policy* 16(2): 52-93.
- Unger, Peter. 1971. 'A Defense of Skepticism,' *Philosophical Review* 80: 198-219.
- Von Magnus, Eric. 1983. 'Rights and Risks,' *Journal of Business Ethics* 2: 23-6.
- Von Wright, G. H. 1956. 'A Note on Deontic Logic and Derived Obligation,' *Mind* 65: 507-9.