

## **Society for Applied Philosophy Annual Lecture 2017**

### **Three Mistakes about Doing Good (and Bad)**

---

**PHILIP PETTIT**

#### **Introduction**

A great deal of attention is given in ethics to whether the right option in any choice is that which does, or promises to do, the most neutral good. But not enough attention is given to the different sorts of good you can bring about in choosing an option. This article seeks to restore the balance, arguing against some claims that unnecessarily restrict the extent to which your choices, in particular your choices in action, can claim to improve the world.

There are three common claims about the effects of action that put significant restrictions on how much good, or indeed how much bad, you can be said to bring about as a result of what you do. The first, most restrictive claim holds that the goods you bring about must all be causal effects or consequences of action, not consequences of a constitutive or grounded kind, as it is described here. The second holds that they must at least be behavioural effects of action, whether causal or constitutive in nature, not consequences that depend on your having acted out of a benevolent disposition. And the third maintains that whether they are behavioural or disposition-dependent, they must all be consequences of an all-or-nothing or on-off or discrete kind; they do not include the extra effect of pursuing such all-or-nothing consequences with a higher rather than a lower degree of benevolence.

Is the good you can do in action restricted to the good you can do causally, or behaviourally, or via its discrete consequences? There are two ways in which such restrictions might be defended, one based in a theory of the good – hedonic utilitarianism – the other in a theory of action. This article addresses the second defence of the restrictions but it will be useful to comment briefly on the first.

Hedonic utilitarianism supports the three restrictions, because of equating the good that you do with the pleasure you generate.<sup>1</sup> The pleasure you generate will inevitably come about as a causal consequence of what you do, not as an effect that the action constitutes. It will come about as a behavioural consequence, materialising independently of your disposition in acting. And it will come about as an all-or-nothing or on-off or discrete consequence; the degree of pleasure the act itself occasions in me will not apparently vary, depending on how far the act was benevolently directed at producing pleasure.<sup>2</sup>

In this article, I work with the assumption that any theory of the good that restricts it to pleasure, or indeed to anything of the same restrictive kind, is at best unsupported, at worst implausible. The reason is that there are many examples of

presumptive goods whose value does not seem to be sourced in anything of the sort. Examples that will figure in later discussion include the goods involved in people's keeping promises made, satisfying someone's trust, and telling an interlocutor the truth, as well as goods like the respect, honesty and friendship we can enjoy in our relations with one another. Without arguing that any one of those properties has independent value, the article assumes that some properties of the kind certainly do.

Assuming that this is so, the question arises as to whether the sorts of goods illustrated are ones that you can bring about as consequences of your action. The three restrictive claims would imply that they are not, suggesting that the only sort of good that can guide you in action is of a kind with pleasure. But the article argues against those claims, showing how goods of the sort illustrated can be realised as effects of what you do, and defending a model of action – the control model, as it is called here – that makes sense of how they can figure in that role. The upshot is that we should expand our conception of consequences, allowing them to include goods of the kind illustrated as effects of how you act.

Despite its metaphysical character, this thesis has important implications for moral theory. By expanding the beneficial effects that you may bring about in action, the thesis makes it more likely that acting well is significantly related to the neutral goods that your actions bring about or promise to bring about. Thus, it opens up the prospect of reducing the gap between the consequentialist doctrine that acting well is always a function of doing such good and non-consequentialist theories that make it a function of other factors alone or of other factors as well.

The article is in five sections. The first offers a basic theory of action that can serve as a framework for discussion. The three middle sections are devoted respectively to identifying and criticising the three restrictive claims, casting them as the *only-by-causal-consequences* mistake, the *only-by-behavioural-consequences* mistake, and the *only-by-discrete-consequences* mistake. And the final section returns to the theory of action, arguing that the critique of the mistakes is supported by an independently plausible model under which acting involves, not just producing an effect, but controlling for it.<sup>3</sup>

## 1. A Theory of Action

In order to understand the mistakes, we should begin with some basic assumptions about what doing good, or indeed acting to any effect, involves. These provide a framework for thinking about action that all sides to our debates might in principle endorse. The framework involves a number of ideas and distinctions. First, the concept of an option and an act; second, the idea of doing something intentionally; third, the by-relation in virtue of which we say that you can do something intentionally by doing something else intentionally; and, finally, the distinction between a basic option and a derived option.

### *Options and Acts*

When you make a decision, you face a number of mutually exclusive and, by assumption, jointly exhaustive options. These may be: turning left or turning right, posting a letter or sending an email, voting for Clinton, voting for Trump or

declining to vote for either. An option in each of these cases is a possibility allowed by your environment that you can take or not take, realise or not realise, depending on your desire.

Thus, you can turn left or turn right as you wish, given your body is in good order and there are two routes before you. You can post a letter or send an email as you wish, given that there is a postal and internet service available. And you can realise one or other of your voting options, depending on what you wish, given that you are a registered voter and there is an election between the relevant candidates. What happens in any such case is, as we say, up to you: it will be realised if you wish to realise it, at least in the normal run of things, and it won't be realised if you do not wish that.

It will be up to you in this sense, of course, only if you are aware of how things are in the environment: while you may not have a complete degree of belief in the matter, you accept for relevant purposes that the postal and internet service is available, that you are a registered voter, and so on.<sup>4</sup> An option, on the approach taken, is up to you in the sense that it is a possibility that you can realise or not, and consciously realise or not. The possibility lies within your power of realisation and, given how you take the world to be, it lies presumptively within your power of realisation.

This account of an option stands midway between two other possible stories. One would say that a possibility is an option for you only if its realisation depends wholly on you and not at all on the environment.<sup>5</sup> This approach would mean that among possible worldly initiatives the only options you ever face involve trying to do this or trying to do that, never just doing this or doing that. You never confront a choice between doing X and doing Y, where realising either requires the environment to be thus and so; you only face a choice between trying to do X and trying to do Y. And that, I shall assume, is downright implausible.

Another account of options would say that if you desire to achieve some effect – even a chancy effect that, as you accept, is not guaranteed by the environment to materialise as a result of your behaviour – then the fact that you succeed in realising it means that it was an option for you. You line up to kick a penalty in a soccer game and manage to get the ball past the goalkeeper. On this account, scoring the goal was an option for you. But like the first alternative to our proposal, this is counter-intuitive. It means that we can think of the decision you face as one between scoring and not scoring rather than as a decision, more plausibly, between shooting in one or another direction, with the aim of getting the ball into the net.

Rejecting these alternatives, I shall assume that the best way of thinking about an option, as suggested earlier, is as a possibility that you are allowed by your environment, and allowed under your way of taking the environment, to realise or not, depending on what you wish. That it is allowed by your environment means that it does not depend wholly on you, unlike a trying. That it is a possibility that you can realise or not realise, depending what you wish, means that it does not have a chancy character, unlike kicking a goal.

When you choose an option in this sense of option – when you realise the conscious possibility involved – then that normally requires you in the circumstances on hand to make a change in how things are, producing a novel event, although perhaps only the negative event involved in choosing not to do something.<sup>6</sup> There is an interesting exception to this pattern, mentioned in Section 5, but for the moment we can set that aside. The event produced is naturally described as the act you perform. Where

options represent types that may be realised in any of a number of perhaps trivially different ways, acts are those token events that actually realise options.

### *Intentionality*

Your causing the act-event in a case where we ascribe an action to you cannot come about in any old way. If the weight of an object at the other end makes you let go of a rope – you can no longer hold on – then letting go of the rope is not an action that you perform. When we predicate action, so it will be assumed here, we require that you were moved by a certain desire – say, in the rope case the desire to let the weight drop – and that the act you performed in letting go of the rope was prompted by a belief that that would cause the weight to drop, in accordance with your desire. Indeed, we require that this combination of belief and desire would have led you to let go of the rope in more or less any variant on the situation where that remained within your conscious capacity, and was still rationally required by your attitudes: it promised to satisfy your desires according to your beliefs. We would not think that your letting go of the rope was an action, for example, if the onset of the appropriate desire and belief made you nervous and caused you in that deviant fashion to let go.<sup>7</sup>

In the case of an action like this we can say that you intended to let go of the rope and intended to drop the weight at the other end; or, alternatively, you acted with the intention of letting go of the rope and with the intention of dropping the weight. And, to introduce a more general idiom, we can say in that case that you intentionally let go of the rope and intentionally dropped the weight. This way of speaking is more general because you can do something intentionally without intending to do it. Even though you did not intend to damage the floor, for example, we are likely to say that you damaged it intentionally if you recognised that the floor is made of wood, and foresaw that dropping the weight, which you did intend, would be likely to have that effect.<sup>8</sup>

When you intend to X – when you do X intentionally in the stronger sense of that phrase – you must desire that the X-event should materialise; this, in the way in which you desire to let go of the rope and thereby drop the weight. When you do not intend to Y, but still do Y intentionally, you do not desire as such that the Y-event should materialise, only that some event with which the Y-event is consciously packaged should do so; this, in the way in which you desire to drop the weight, recognising that dropping the weight is packaged with damaging the floor. In the second case, you do not drop the weight because it promises to damage the floor, you drop the weight in spite of the fact that it promises to do so.

### *The By-Relation*

These notions make it possible to introduce the by-relation. The relation between letting go of the rope and dropping the weight allows us to say that you dropped the weight by letting go of the rope. And the relation between either letting go of the rope or dropping the weight, on the one side, and damaging the floor on the other allows us to make use of the same locution. You damaged the floor, we can say, by letting go of the rope and, perhaps more informatively, that you damaged the floor by dropping the weight.

In an influential account of the by-relation, Alvin Goldman<sup>9</sup> suggests that we ought to think of letting go of the rope, dropping the weight and damaging the floor as distinct events. But consistently with acknowledging the by-relation, and benefitting from his analysis, we can characterise things differently. We can say that in a case like this there is only one act and one event but that it has many properties, each associated with one of the ways in which the action can be described. We can conceptualise the action coarsely rather than finely, in other words, and avoid what looks like an unnecessary multiplication of entities.<sup>10</sup> In order to emphasise that there is only one event involved, we might say that by making it true that you let go of the rope, you also made it true both that you dropped the weight and that you damaged the floor; by ensuring that the event satisfied the one description you ensured that it satisfied the others too.

### *Basic Options and Acts*

We have introduced the notions of option and act, intentionality and the by-relation. And that makes it possible, finally, to introduce the notions of a basic option and a basic act. A basic option is one that you can intentionally realise other than by intentionally realising anything else: intuitively, anything more basic; and a basic act is the token act that realises a basic option.<sup>11</sup> In the example given, letting go of the rope – that is, releasing your grip on the rope – is a candidate for a basic option and act but dropping the weight and damaging the floor are not. Letting go of the rope intentionally is something that you just know how to do, in a primitive sense of know-how. Dropping the weight or damaging the floor, in the example given, is something that you know how to do by knowing how to do something else: letting go of the rope.

Why is your knowledge of how to let go of the rope primitive? Because you are unlikely to be aware of the neuronal transmission and the muscular contraction involved in letting go of the rope. And, if you are aware, you are unlikely to be able to intentionally initiate that neuronal and muscular process as such. The most basic description under which you will be able to desire and intend the act is simply: ‘I let go of the rope’. You can realise that description intentionally without realising any more basic description intentionally.

That will be so, at any rate, unless you are recovering from an accident or ailment of some kind and are having to learn or re-learn how to let go of something in your grasp. It should be clear that depending on the skills you have mastered, or that have come to you by grace of nature alone, you may be able to treat ever more complex options as basic. The child who learns how to tie its shoelaces masters as a basic option something that might previously have presented as an option to realise by doing this or that with the left hand, this or that with the right hand, and so on in tortuous detail.

### *Acts and Actions*

Given that an act is an event with many different properties, as we have stipulated, a question arises as to which sort of property individuates the act, belonging to it essentially. The natural answer is, the property under which you can realise or enact it as a basic option. On this approach, we cannot think that while remaining one and the

same act, the event might have instantiated a different basic property. But of course, we can imagine that that act might have varied in other properties: for example, that it might have originated in a different source, occurred in a different context, or had different effects.

By the account we will be working with in the following discussion – it will be developed and qualified in the final section – action involves a process of choosing an option and generating an act. The term ‘action’ will be taken to refer to the process involved in the choice and enactment of the option, the term ‘act’ to designate what emerges as a result. Whereas the action is a doing, the act or deed is what is thereby done.<sup>12</sup>

## 2. The Only-by-Causal-Consequences Mistake

### *The Question About Doing Good*

The ideas reviewed up to this point enable us to raise the question of what it is that establishes whether or not you do good in acting. Whatever makes an action good is presumably a function of the properties of the act produced, whether they consist in its having a certain essential nature, its deriving from a certain source, or its generating certain effects. The question, then, is what are the various properties of an act that might figure in ensuring that you do good in taking it. The answer is bound to have implications too for what might figure in determining that you do bad in taking the action but we shall concentrate mainly on the issue with goodness.

In order to understand the question, three qualifications are essential. First, the properties that make an action good will each make it good only in a certain respect, only *pro tanto*. They may not make it good overall, or good *simpliciter*: in effect, they may not make it better than alternatives. Good-making properties are any properties that can contribute to the goodness of an action, making it good as such: good insofar as it instantiates those features.<sup>13</sup>

Second, the properties envisaged here are presumed to make actions neutrally good, not just good from the point of view of the particular agent or a particular beneficiary. They make for goodness in a universal sense of the term, not just for what might seem to be good from the perspective of a particular individual or group.

Third, knowing the properties that make actions good in this sense does not necessarily entail knowing what you ought to do, or ought to have done, in a given choice. Even if those properties are the only factors relevant to what you ought to do or have done, their relevance may vary with the probability that you assigned at the time of choice to their materialising as a result<sup>14</sup> – or perhaps with the probability you ought rationally to have assigned to this – and with how they weigh against one another.<sup>15</sup> And of course those properties might not be taken as the only factors relevant to what you ought to do; this is what non-consequentialists maintain.

### *A Good-Making Essential Property?*

What properties of an action might figure, then, in the good-making role? A first possibility is that the essential property of an act – the property under which the option

involved is enacted – should be itself a good-maker. If it is, then you do good just in virtue of realising a possibility that instantiates that property. The basic option is good as such, good qua bearer of the very property under which you enact it. And since the act realises that property essentially, its goodness makes the action necessarily good.

It is certainly possible that the basic property of an option should be a valued or indeed a disvalued property. But in most cases, it probably doesn't have that character. Basic options in ordinary behaviour involve movements of the limbs that present in a more or less value-free way: moving my hands in this or that pattern or moving my legs in this or that direction. And basic options in speech, which amount just to choices of different words, involve utterances that generally have the same value-free profile.

Perhaps the only exceptions to this comment on the essential properties of acts arise with mental acts. If forming an intention for the future is an option for you, then forming the intention to help me in some way – say, help me to move apartments – is a basic option. And it may be good as such: good in virtue of the very property that makes it into something you can do. That is certainly an open possibility, as it is an open possibility that forming an intention to harm me is bad as such.<sup>16</sup>

In what follows, however, I shall put mental acts aside. They are a special category by all accounts, and a controversial category by some. In any case, they raise complexities that make it impossible to include them in an article of this length. Thus, the focus of the discussion will be on the variety of contingent properties in virtue of which an external act might be good as such, good *pro tanto*.

### *Good-Making Causal Properties*

Assuming that the basic, essential properties of external acts do not make them good as such, whether you do good in any instance is going to depend on the other properties that your acts display. And at this point the causal properties of acts come into prominence. For perhaps the most obvious way in which you can do good in performing an act A, where 'A' picks out the act under its basic property, is by causing something good to ensue: say, to take a utilitarian example, by increasing my happiness without, as we may presume, reducing anyone else's. In that case the act is not only a bearer of the A-property, it is also a bearer of the felicitic property of actually promoting happiness. And in virtue of doing something with that property or relation, you count as doing good.

Causalists, to give a name to those who commit the first mistake, maintain that the only properties of an act in virtue of which the action can have value, or indeed disvalue, are causal properties in this intuitive sense.<sup>17</sup> Their mistake becomes visible once we recognise that what you do may involve doing good, not just in virtue of its causal properties, but also in virtue of other properties.<sup>18</sup> There are two prominent kinds of non-causal properties that offer themselves as candidates. We may describe them as conventional properties, on the one side, circumstantial properties on the other.<sup>19</sup>

### *Good-Making Conventional and Circumstantial Properties*

Suppose that the A-act that has the causal property of making me happy involves returning a treasured book, as you had said you would do when I lent it to you. In virtue of your having said you would return the book, the A-act counts by convention as

a case of keeping a promise. And just in virtue of instantiating the promise-keeping property in the wake of an actual promise, it is likely to count as good by received criteria of value. Thus, you count as doing good in performing the A-act, not just by doing something with the causal property or relation, but by doing something with the conventional property or relation. By realising A you realise that property and in realising that property you do good.

Is keeping a promise a consequence of your doing A? Yes, in a perfectly normal sense of the word, since it is something you bring about by giving me back the book. But keeping the promise, as distinct from making me happy, does not come about by a causal effect: a sort of effect, we may presume, that takes time to materialise. My becoming happy is a distinct event from your returning the book and there is a temporal lapse, however brief, between your performing that act and my experiencing happiness. Keeping your promise is not a distinct event from your returning the book, so that there is no temporal lapse between your realising that act and the promise being kept. The act constitutes or grounds the promise-keeping rather than causing it.

The A-act may also count as good in virtue of one or another circumstantial property, not just in virtue of its causal and conventional properties. Suppose that there was no independent assurance that you would return the book but that nonetheless I trusted you to do so. This means that in returning it, not only do you keep a promise and make me happy, you also satisfy my trust in you, and bring about what we may suppose is a good: you act as I had relied on you to act.

Satisfying my trust in you, like keeping the promise, is a consequence of your returning the book; it is something that you bring about by returning the book. But as in the promise case, it is not a causal consequence of the act, A. Returning the book constitutes satisfaction of my trust rather than causing it. The satisfaction is not a distinct event from returning the book that only materialises after some temporal lapse. My trust is satisfied as soon as you return the book, although it may take some time for me to recognise this.

There is nothing mysterious about the idea that a consequence might be a constitutive consequence of an act – a consequence realised by virtue of convention or circumstance – rather than a causal one. Taking an example from beyond the realm of action,<sup>20</sup> consider the relationship between the antibodies in your blood stream and your consequent immunity to a certain disease. Your immunity is a consequence of the presence of the antibodies because a counterpart of the by-relation obtains between them. It is by developing the antibodies that you become immune, and not vice versa. But the relationship between antibodies and immunity is not causal, for once the antibodies are present, the immunity is present; there is no temporal lapse between the two. The antibodies constitute or ground the immunity rather than causing it.<sup>21</sup>

These observations should be enough to show that it is a mistake to think that you can do good only by causing something good to come about; that what you do may be good in virtue of conventional or circumstantial properties, not just in virtue of causal ones; and that there is nothing particularly mysterious involved in taking the realisation of such a property to be a constitutive as distinct from a causal consequence of your basic act. Thus, to go beyond the examples mentioned, the constitutive consequences of acts in virtue of which you do good may include, on the one side, the conventional consequence of saying ‘I do’ that you marry the person you love, or of signing a check that you pay off your creditors; or, on the other, the

circumstantial consequence of making a report that you tell the truth, or of shouting 'Fire' that you signal a danger.

#### *How to Explain the First Mistake?*

Why would anyone ever fall prey to this first mistake? Why would they ignore the constitutive or grounding manner in which you can generate good effects, acknowledging only the causal way of doing so? Classical utilitarians may commit the mistake because of thinking that the only good is pleasure, since this is inevitably a causal consequence of what you do. But why any others?

Walter Sinnott-Armstrong<sup>22</sup> gestures at a possible answer. As consequentialism is generally conceived, he says, it refers to 'any descendant of classic utilitarianism that remains close enough to its ancestor in the important respects'. Thus, he suggests, the consequences that are taken to be relevant by many defenders and opponents of consequentialism are 'opposed to the circumstances or the intrinsic nature of the act or anything that happens before the act'. This mistaken conception of consequences, as I read it, would deny that your keeping a promise or satisfying my trust could be a consequence of what you do in returning my book. It would restrict the relevant consequences, hedonic or non-hedonic, to the causal effects of your action.

### **3. The Only-by-Behavioural-Consequences Mistake**

#### *Behavioural Goods*

The causal, conventional and circumstantial properties that may ensure that you do good in choosing an act all have one feature in common. They attach to the act or behaviour quite independently of the disposition out of which it is performed. They depend for their realisation only on the nature of the *act-in-situ*, i.e. the behaviour itself and the context in which it occurs; context fixes the linkages that mediate the causal effects of the behaviour and the conventions and circumstances that mediate the constitutive. Not depending on the dispositions out of which you act, any goods that you bring about by virtue of realising such properties may be described as behavioural in character.

Thus, if you causally produce a certain pleasure in others by how you act, it does not matter that you did so out of this or that disposition. I may derive pleasure from believing or perceiving that you acted out of a particular disposition, of course – say, friendship for me – but that pleasure is not an effect of the behaviour itself, only an effect of my belief that the action springs from friendship. As the causal consequence of behaviour is disposition-independent, so too are consequences that are grounded in convention or circumstance. If you keep a promise or tell the truth, it does not matter that you act out of this or that disposition; whether you act out of virtue or attachment or just sheer opportunism, you still keep the promise or tell the truth.

The second of our three mistakes consists in thinking that the only goods you can bring about by what you do are behavioural goods. They derive either from the causal effects of your behaviour or from effects of the constitutive kind illustrated by the

goods it generates in the presence of certain conventions or circumstances. In particular, they do not depend in any way on the disposition out of which you act.

### *Robustly Demanding Goods*

Those who make the second of our mistakes – behaviourists, for short – hold that there are no other benefits you can bring about apart from goods of this kind. But that thesis runs up against the fact that many of the presumptive goods in human life are robustly demanding, as we may put it, and such that you can only bring them about in virtue of acting out of suitable dispositions.<sup>23</sup> Some examples will illustrate the sorts of goods in question and, with those in our sights, we will explore the way that they depend for their realisation, not just on how you behave, but on the disposition at the origin of the behaviour.

Goods like those associated with producing pleasure or keeping a promise or satisfying trust are only contingently demanding. They require that you generate the pleasure, keep the promise, or satisfy my trust in the actual scenario that obtains. But they do not demand that you would have acted in any particular manner in counterfactual variations; they do not require that you would have acted to a similar effect if, for example, doing so had been more difficult or costly or inconvenient.

Where my enjoyment of these benefits requires only that things be thus and so in the scenario that contingently obtains, my enjoyment of robustly demanding goods requires that things would also be thus and so in various counterfactual situations. In order to provide me with a robustly demanding good, it must be the case that you actually confer a certain benefit on me but it must also be the case that you would have conferred the same benefit robustly over certain counterfactual variations on the actual circumstances. In other words, the realisation of the good imposes modal as well as actual demands.

Think of the good you do me when you give me respect, or give me the gift of your honesty, or prove yourself a friend. In each of these cases there is a benefit in the background that is merely contingently demanding: the favour of a friend, the truth-telling that honesty requires, or the restraint associated with respect. But in order to give me respect in some interaction, it is not enough that you happen to show restraint on that occasion; in order to let me enjoy your honesty it is not sufficient that you happen to tell me the truth; and in order to act as a friend it does not suffice that, as chance would have it, you give me a certain favour. If you give me the restraint or truth-telling or favour just because it happens to suit you in the relevant situation to do so – if you would not have given me that benefit had it not been so convenient or had I not been so charming – then you do not give me the good associated with respect or honesty or friendship.

What these observations indicate is that goods like respect and honesty and friendship, which are surely among the most important, presumptive goods in human life, are robustly demanding. In order for you to provide me with such a boon it must be the case, not only that you actually deliver the associated benefit – restraint, truth-telling or favour – in response to suitable prompts or reasons but that you would also deliver it over a certain range of counterfactual scenarios.<sup>24</sup>

What range of scenarios? Idealising what may be expected of imperfect agents, we might go for a formula like the following. It is essential that you would deliver the

benefit robustly over variations on the actual situation where the prompting considerations that make a case for respect or honesty or friendship remain in place; where there is nothing like an excuse to deprive you of the capacity to respond to those reasons; and where there are no red lights to indicate that the considerations are trumped by reasons of a manifestly weightier kind: it is not the case, for example, that telling the truth would put someone's life in danger.

In this sense of delivering the benefit robustly, you would have to deliver it robustly over any variations of our circumstances that kept those features in place: say, over any variations that just made it more burdensome or less convenient for you to deliver the benefit. When we speak in what follows of producing a benefit robustly, that is short for speaking of producing it actually and over those sorts of variations on the actual circumstances.<sup>25</sup>

### *The Role of Dispositions*

Assuming that there are robustly demanding goods of this kind available, how are you to do good in one of these ways; how are you to bring about such appealing consequences? The answer is not that providing the associated benefit is going to cause me to enjoy your respect or honesty or friendship; it may cause me to think of you – rightly or wrongly – as respectful or honest or a friend but, given the possibility that you are just opportunistically beneficent, the act can hardly cause the respect or honesty or friendship itself. Nor is the answer that performing that act is going to give me respect or honesty or friendship insofar as it accords with a suitable convention or occurs in a suitable circumstance; nothing of that kind would equip the act to deliver the result sought.

If the action of giving me the relevant benefit – restraint or truth-telling or favour – is to provide me with the robustly demanding good involved, then what is required is rather that you perform the act out of a suitable disposition.<sup>26</sup> Absent excusing obstacles and red lights, to invoke the provisos mentioned, you act out of responsiveness to the reasons that argue for giving respect, for being honest or for acting as a friend. And acting out of that disposition, it is not only the case that you provide the restraint or truth-telling or favour in actual circumstances; you would provide it also under a range of suitably varying, counterfactual scenarios. You provide me with the robustly demanding good as a consequence of acting out of the disposition.<sup>27</sup>

When you act out of a suitable disposition and give me the robust good associated with it, I may or may not recognise that you are doing that. And when you fail to act out of the necessary disposition, I may still think that it is in place and that I am the beneficiary of your respect or honesty or friendship. But regardless of my epistemic success or failure on that front, the fact is that just insofar as you deal with me out of the appropriate disposition, you confer the robustly demanding good as well as the merely contingently demanding benefit that it involves. You show me respect as well as restraint; give me the benefit of your honesty, not just your truth-telling; and bestow friendship as well as favour.

These observations direct us to a way in which you can do good that is distinct from the ways of doing good reviewed in the previous section. While this way of doing good has not been sufficiently recognised in the literature, it is similar in one regard to the constitutive way you do good in keeping a promise or telling the truth.<sup>28</sup> When the

behaviour is prompted by a suitable disposition, showing restraint in dealing with me just is a case of giving me respect; in the presence of that guiding disposition, the act of restraint constitutes an act of respect. And in the presence of suitable guiding dispositions, telling me the truth constitutes a display of honesty and giving me a certain sort of favour constitutes an act of friendship. As a good like promise-keeping or truth-telling is constitutively dependent on convention or circumstance, the richer good in each of these cases, whether of respect, honesty or friendship, is constitutively disposition-dependent.

### *Good, Bad and Robustness*

In illustrating the ways in which causation, convention and circumstance may ensure that you do good, we noted that they may equally ensure that you do bad. Just as you may do good by making me happy, keeping a promise or satisfying my trust, so you may do bad by making me unhappy, breaking a promise or disappointing me. Is there a similar balance between doing good and doing bad in the domain where the relevant consequences are disposition-dependent rather than dependent on the causal effect of the act, the conventions that bear on the act, or the circumstances under which it is performed? While this is not the place to explore the issue properly, it is worth observing that there is an asymmetry in this domain of a kind that does not appear in the others.<sup>29</sup>

You must robustly provide restraint in order to show me respect, robustly provide me with truth-telling in order to give me the benefit of honesty, and robustly provide me with favour in order to give me friendship. But now consider the corresponding evils of disrespect, dishonesty and breaches of friendship. In order to disrespect me you do not have to deny me the benefit of restraint on a robust basis, you just have to deny me restraint contingently. In order to be dishonest in dealing with me, you do not have to tell me lies robustly, you just have to tell me a lie contingently. And in order to prove yourself just a fair-weather friend, you do not have to refuse me favour robustly, you just have to refuse it contingently.

Hobbes<sup>30</sup> notes this asymmetry between good and bad in discussing the difference between what is required for being someone who does justice to others, and what is required for being unjust. Whereas a just man, he says, is ‘he that taketh all the care he can that his actions may all be just, an unjust man is he that neglecteth it’.<sup>31</sup> Most of the actions we indict as instances of injustice, Hobbes suggests, involve contingently breaching the standards of justice, not robustly flouting them. But all the actions that we treat as instances of justice involve robustly conforming to those standards: in that sense, justice is a robustly demanding good.

Although the asymmetry between good and bad in these domains is surprising, it is readily explicable. Hobbes gives the clue to the solution when he comments that the unjust man is not moved by the prospect of doing injustice to others but rather ‘by the apparent benefit of what he is to do’: that is, by the self-interested benefit to himself.

Assume that self-interest is the problem that prompts us to establish standards like those associated with justice, or indeed with any good like respect, honesty and friendship. That means that agents will count as doing good by resisting self-interest robustly: that is, by conforming robustly to the relevant standards. And it means

equally that agents will count as doing bad by giving in to self-interest and breaching one of the standards: and this, even if the breach is contingent rather than robust. Where beneficence in such areas is robustly demanding, maleficence generally is not.<sup>32</sup>

Some people may prove to be robustly maleficent, of course, seeking even at a cost to their own interests to impose this or that sort of harm on others. That pattern may be associated with extreme envy or vengeance or racism, for example. It is exemplified by Aaron's remark near the end of Shakespeare's *Titus Andronicus*: 'If one good deed in all my life I did, I do repent it from my very soul'. And it is summed up in the mantra of Milton's Satan: 'Evil be thou my good'. But however fascinating it may be, especially in literary representation, the bad that most of us do is of a decidedly more banal character.<sup>33</sup>

#### *How to Explain the Second Mistake?*

Why would someone be led to embrace behaviourism as a general theory of the good you may do in action? Why would they neglect the importance of goods that require, not just the performance of certain acts, but their performance out of appropriate dispositions? What could prompt such neglect?

The answer, I suspect, is a tendency, common among recent moral theorists, to distinguish between the assessment of actions and the assessment of agents in a way that puts dispositions entirely on the side of agents. The idea is that in determining the moral value of an action you perform – in determining the value of what you do – we should look to its disposition-independent properties alone. We should look to your dispositions only in determining what sort of person you are: how respectful or honest as an agent, how faithful as a friend.

This tendency is particularly prominent in Jonathan Bennett's influential book *The Act Itself*.<sup>34</sup> Drawing on earlier work by Alan Donagan,<sup>35</sup> he argues that while 'most moral theorists' take the moral value of an action to depend on its relational properties, those properties exclude any 'relation to a motivating state of mind': that is, to the dispositions manifested in the action.<sup>36</sup> Thus, he allows only the consideration of disposition-independent properties – for example, causal, conventional or circumstantial properties – in determining what you bring about in action and its moral value. He restricts a consideration of dispositions to the evaluation of agents, something that he describes as an exercise of second-order rather than first-order morality. In the case of a given action, this looks at whether the 'person deserves credit or discredit – perhaps including praise or blame'.<sup>37</sup>

Many figures in moral philosophy agree with Bennett in focusing mainly on disposition-independent properties of actions, although they generally moderate the position by conceding that the intentions of agents, if not their more general dispositions, are at least sometimes relevant in action assessment. T.M. Scanlon<sup>38</sup> thinks that intentions are relevant in those cases – a minority, he suggests – where they help to determine the significance of an action, as he puts it. And Derek Parfit<sup>39</sup> suggests, although only in passing, that they may be relevant in all, being presupposed in the description of an action that fixes its moral value. But both writers treat the presence of most dispositions as relevant only in the assessment of agents, not in the assessment of actions.

Scanlon<sup>40</sup> follows Bennett particularly faithfully, when he treats blame as an attitude prompted by the character of the agent, not by the character of the action.<sup>41</sup>

Whether or not an exception is made in the case of intentions, the hostility to dispositions as factors relevant to determining what you bring about in action may explain the appeal of behaviourism. But why the hostility to dispositions in thinkers like Scanlon, Parfit and others? The reason, plausibly, is that in the moral appraisal of action, they focus on finding a code to determine what you may do and may not do: what is permissible and what not.<sup>42</sup> In Scanlon's<sup>43</sup> case, this is a contractualist code; in Parfit's<sup>44</sup> a code that contractualists, rule-consequentialists and certain sorts of Kantians can endorse in common.

In constructing a legal code, it is a natural desideratum that actions that are deemed legal or illegal should be readily identifiable by those who live under the law, so that those agents can be effectively guided by the laws. That means that so far as possible, it will be useful to identify permitted and prohibited actions independently of the dispositions out of which they are produced. Let the dispositions be relevant to determining the range of permitted or prohibited actions and it will be more difficult for subjects to be sure of how to keep on the right side of the law.<sup>45</sup>

A similar consideration applies with any moral code. In constructing a code of this sort, it is also going to be attractive to identify the actions that are permitted or prohibited, so far as that is possible, without taking account of the dispositions out of which they are performed. But morality is not just about constructing a moral code and determining what it permits. It may be useful to have general rules for purposes of cooperation and coordination but the ethics of action encompasses other ends as well. The aim, presumably, is to provide a moral compass for shaping and assessing action, not just a moral code. And so, a desideratum that applies only in relation to a code should not be allowed to restrict the morally relevant properties of actions to their disposition-independent features.

#### **4. The Only-by-Discrete-Consequences Mistake**

##### *The Motivating Idea*

According to the argument so far, the good that you are capable of doing or bringing about in action may include the good associated with non-causal as well as causal consequences and with robustly demanding as well as contingently demanding consequences. For all this shows, however, it may still be the case that the good you can do is restricted to the good you bring about by what we may describe as discrete or all-or-nothing or on-off consequences. And that is certainly a matter of common assumption.

But however common, this assumption too is a mistake. You can do more good, not just in virtue of the purely contingent consequences that you realise, whether these be causal or non-causal, but also in virtue of the degree of robustness with which you realise them. And this can be so, even when you do not bring them about with a sufficiently high degree of robustness to generate goods like friendship or honesty or respect.

The idea of robustly realising a consequence applies, of course, only with contingently demanding effects: benefits like favour and truth-telling and restraint. These are

benefits that you can generate in actual conditions without its being the case that you would generate them under any counterfactual variations on those circumstances. While they must be realised with a particularly demanding range of robustness in order to generate friendship, honesty and respect, they may also be realised at lower levels of robustness over the relevant variations.

The third of our mistakes consists in thinking that while producing such a benefit with enough robustness to generate the demanding good may do good, producing the benefit at any lower level of robustness does no more good than producing it only contingently would have done. You give me a good in bringing about a benefit like restraint. And you may give me a good in bringing it about robustly enough to provide me with respect. But, so the idea goes, there is no distinct sort of good that you bring about by virtue of giving me the restraint with any lesser degree of robustness.

### *The Mistake in the Idea*

If the arguments of the last section make a case against the second behaviourist mistake, as they intuitively do, they combine with other considerations to make a case against this mistake too. Let there be robustly demanding goods of the kind that they illustrate, and there are at least two considerations, one of indeterminacy, the other of grounding, that argue against the third mistake. They each suggest that you will do more good in producing the benefit in question, the more robustly you realise it over any scenarios in the relevant range of possibilities. They argue that as the good you do in conferring a robustly demanding good is disposition-dependent, so the good you do more generally is disposition-sensitive.

The first consideration is that the level of robustness at which the provision of a benefit generates a robustly demanding good – the level, for example, at which the provision of restraint constitutes an act of respect – is bound to be somewhat vague or indeterminate; to that extent, describing it as an all-or-nothing or on-off consequence may be somewhat inappropriate. Idealising the requirement for robustly demanding goods, we said that in order to give me respect it must be the case that you actually provide the restraint – this, presumptively, in response to reasons of respect – and, in addition, it must be the case that you would provide it across all possible variations on actual circumstances that preserve those reasons, do not deprive you of the capacity to respond to the reasons, and do not put on the red lights by revealing important considerations that trump them in importance.

We can hardly require this fully idealised performance whenever we take you to give me respect in actual life or indeed to give me the benefit of your honesty or friendship. We will be happy, presumably, to attribute a suitably robust beneficence so long as you come close to the idealised performance. We must require, of course, that you actually display restraint. But beyond that, we may only require that you would also display it in most of the relevant counterfactual scenarios or, if that is distinct, that you would be more or less likely to display it in all.<sup>46</sup> The requirement may be relatively indeterminate or vague.<sup>47</sup>

The indeterminacy of the threshold of robust provision at which restraint constitutes respect, truth-telling honesty, and favour friendship – the fact that there is a threshold range, itself with indeterminate boundaries, not a threshold point – argues against the idea that you do better than you would do in contingently providing the benefit only

when you reach the threshold. If being in an appropriate, indeterminately marked vicinity is what is good, then it seems natural to think that steps that take you towards that vicinity must be good. And that is to say in the examples on hand that providing restraint, truth-telling or favour with increasing robustness means doing increasing good, even if it falls short of taking you into the region of respect, honesty or friendship.

This argument is strongly supported by a second, more general consideration. Suppose that the threshold at which the robust provision of restraint constitutes respect were absolutely determinate or exact. Even in that case, it would be strange to think that short of the threshold limit you do not do any better by approximating it than you would do by displaying merely contingent restraint. If goodness is associated with attaining the threshold, that is presumably grounded in some property of the threshold. If it were not grounded in this way, after all, then the goodness of reaching that particular threshold would be primitive and unexplained. But what can ground the goodness of the threshold, if not the fact that it involves a particularly high degree of robustness in the provision of the associated benefit? And if that is what ensures the appeal of reaching the threshold, it must ensure that any degree of robustness in providing the benefit is appealing. Thus, the more and more robustly you provide a benefit, the more good you do; and this, even if you do not get to provide the benefit at the threshold level.

This grounding consideration may argue for a conclusion that bears on any contingently demanding benefits, it should be noticed, not just on those benefits, like restraint, truth-telling and favour, that are associated with familiar robustly demanding goods. It suggests that the more robustly you provide for any contingently demanding benefit, the more good you do, and indeed that the more robustly you provide for any contingently demanding cost, the more bad you do. The benefit or cost remains constant as the robustness of provision rises but I enjoy a greater and greater degree of benevolence or malevolence at your hands. And as more benevolent restraint will ground respect at a certain limit, so greater benevolence or malevolence may be taken to ground a kind of robustly demanding good or bad: sympathy or antipathy, as we might call them.

### *The More-Robustness Principles*

There is no plausible way of measuring the graded or range good that you bring about as you increase the robustness with which you generate such a benefit; in that sense, there is no discrete consequence associated with your action. But it is plausible that, other things being equal, the more robustly you produce the contingently demanding benefit, the more good you do. By producing the benefit more robustly, you provide me with more robust access to enjoying it. And you provide me with such enhanced access in virtue of the disposition out of which you act, albeit that disposition is not sufficiently dependable to support the provision of any corresponding, robustly demanding good. You may not provide me with friendship but you do provide me with a relatively robust degree of favour. You may not grant me the good of honesty but you do give me a relatively robust degree of truth-telling. You may not provide me with respect but you do provide me with a relatively robust degree of restraint. And so on, in other cases.

If this principle governing good is plausible, so too is a corresponding principle governing bad. Thus, it is intuitively going to be worse, other things being equal, to impose a cost on me robustly over any cases where it is mildly to your advantage than to impose it only over cases where your welfare is seriously at risk. And it is going to be even worse, other things being equal, to impose that cost over any case where it hurts me, even cases where doing so actually goes against your own independent interests. This principle explains why the extremes of maleficence associated with deep envy or vengeance or racism – in a word, antipathy – are so horrid and fascinating.

Apart from being intuitively plausible, the more-robustness principles are useful in explaining why various traditional doctrines are *prima facie* attractive. Take the doctrine that doing harm is worse than allowing harm, doing good better than allowing good. In most cases allowing a harm or a good to materialise – failing to prevent it – means controlling less robustly for its realisation than bringing it about yourself: merely allowing the effect to materialise will not necessarily lead to its realisation in scenarios when others can prevent it instead. This consideration may explain why the doctrine is attractive to many, although it does not vindicate it in the problematic cases where it is often applied: say, in Peter Singer's<sup>48</sup> case of allowing a child to drown when you are the only one who can prevent it. Their utility in this regard should help make the case for the plausibility of the more-robustness principles.<sup>49</sup>

#### *How to Explain the Third Mistake?*

As already suggested, the explanation for why people fall prey to this mistake may involve one of two failures. First, a failure to recognise robustly demanding goods and to see that behaviourism is false. Or second, a failure to see that if robustly demanding goods are recognised, then that argues, in view of the indeterminacy and grounding considerations, that you must do more good the more robustly you realise any contingently demanding benefit.

Neither of these failures will look like a failure, however, under a model of action – the production model, we may call it – that appeals in many quarters. In order to defend the positions taken, then, we need to return to the general theory of action.

## **5. Action, Production and Control**

### *The Production Model of Action*

Our arguments about the second and third mistakes turn crucially on an acceptance of the examples given of robustly demanding goods. If those goods are admitted as even candidate effects of action, that reveals the error in the second mistake; and, together with the considerations about indeterminacy and grounding, it also reveals the error in the third. But notwithstanding the claims made by the examples, the mistaken doctrines may still retain an attraction. This is because they may seem to be unavoidable under the production model of what action involves. The model fits with the framework theory of action sketched in the first section but is not entailed by it; that framework is equally consistent with the control model to be introduced here.

According to the production model of action, the disposing or motivating state that prompts you to act in a certain way is just a contingent cause of the act, linked to it

by a contingent causal law, so that cause and act are distinct existences, in Hume's phrase. To perform an action is to instantiate and be prompted to realise the required act by such a productive state. This means that whatever the effects of the act, it would still have generated them, had it been sourced in a distinct cause: a distinct disposing or motivating state. The marks left by the cause of the act can only get transmitted to the act's effects, after all, via the ahistorical marks it leaves on the act itself; they cannot leapfrog the intervening act, because causation operates, by standard accounts, on the basis of local connections.

It follows, on this way of conceiving action, that all the effects of an act must be due to disposition-independent properties, not due to how it is motivated. And that means that those effects cannot include the realisation of any robustly demanding good, or the realisation of any benefit at one or another degree of robustness, since such effects materialise only when the act manifests a suitable disposition. The second and third mistakes are not mistakes, then, by this account: they register straightforward implications of the production model.

The line for which we have been arguing is not destabilised by this challenge from the production model, however, since our approach directs us to an alternative picture in which control rather than production is the primary element. And this model, it turns out, is independently appealing.

### *The Control Model of Action*

When you give me a robustly demanding good like respect, the action involved is not adequately analysed into a sequence involving the presence of a suitable disposing state and the production of the act of restraint as a contingent, causal result. The production model misses the fact that to give me respect is not just to be the site or locus where a mental state contingently produces behavioural restraint. It is to be a system in which, for more or less any instance where respect requires you to exercise restraint, a disposing state will materialise – maybe this, maybe that – to generate the required restraint. It is to be a system that controls for respectful restraint: it produces it, not just under actual circumstances, but under any of a range of circumstances where it is required. These are circumstances, like those that actually obtain, where reasons of respect are in place, there are no excusing obstacles to action and there are no red lights to suggest the presence of trumping considerations.

There is nothing mysterious in the notion of your controlling in the manner envisaged for an effect like respectful restraint. The temperature system in a building controls in a parallel way for the range of temperature inside. It not only produces a degree in the appropriate range under actual circumstances; it would also produce a degree in that range under more or less all variations in weather conditions. In other words, it produces an appropriate degree of temperature, not just contingently on things being as they actually are, but robustly over suitable variations in the weather. You control for my enjoyment of a suitable form of restraint in roughly an analogous manner. You produce that restraint, not just contingently on things being as they are here and now, but robustly over variations of circumstance that leave the reasons of respect in place and do not introduce excusing factors or red lights.

The fact that the temperature system controls for the range of temperature means, not just that the system is in a state that produces the required adjustment in response

to actual circumstances, but that it would enter a state designed to produce the required adjustment – this, at least, when all goes well – in response to any in a corresponding range of possible circumstances. The fact that you control for the respectful restraint that you display in dealing with me and others means, analogously, that if restraint is required in any interaction with one of us then, all going well, you will enter a state that produces restraint in that instance. You will enter such a state, and produce restraint, if reasons of respect apply, and excuses and red lights are absent. Or you will do so, at any rate, if this is so according to your beliefs; this is a complication we may put aside in the present context.<sup>50</sup>

This picture of the control you exercise in giving me respect, displaying restraint as it is required, fits nicely with the idea that the restraint derives from a suitably respectful disposition. To exercise the disposition for respectful restraint in any situation is to instantiate the higher-order property of instantiating a lower-order property – the disposing state – that produces the restraint in that situation. And to be generally disposed to act in that way is to instantiate the higher-order property of being such as to enter a suitably productive, disposing state in any situation where respectful restraint is required.<sup>51</sup>

The difference between merely producing an effect and controlling for an effect raises a straightforward question in the theory of action. Does acting consist in just producing an act or, more broadly, in controlling for the act? The control model is appealing, both on grounds local to our concerns here and for more general reasons.

#### *The Local Appeal of the Control Model*

The control model obviously has attractions from the point of view of this article. If we equate acting with controlling for an act, rather than just producing it, there is no reason to deny that the effect whereby I enjoy respect at your hands is one that you bring about via your action. Respecting me is presumably an option for you – it is a possibility that you can realise or not, depending on what you wish – and you can clearly control for its realisation; you will do this insofar as you control for the display of restraint in the presence of suitable reasons, and in the absence of excuses or red lights. And as you can control for respect, so you can control for respect under a proviso: say, the proviso that your self-interest is not deeply compromised. Thus, there is no reason to deny that the robustness you display in pursuing restraint, whether or not this is sufficient to give me respect, is something you bring about in what you do; the robustness is a direct reflection of your control.<sup>52</sup>

These observations explain why the control model of action enables us to expand the set of consequences in virtue of which you may do good or bad to include those, as maintained here, that are non-causal, non-behavioural and non-discrete.<sup>53</sup> But apart from having that local appeal, the control model should also appeal on the more general ground that it helps to explain what makes the production of an act intentional.

#### *The General Appeal of the Control Model*

When you act intentionally there is always a desired end – say, the end of producing a benefit – that you pursue under the guidance of your beliefs: say, your beliefs about the opportunities and means at hand and about the obstacles that may get in the way.

As we noted earlier, however, it is not enough that this belief-desire combination produces a suitable act contingently on some particular feature of the situation: say, that its onset makes you nervous, as in the example with the rope. It must produce that act, assuming it is within your capacity, as a result of the fact that your attitudes rationally require it. This means that implementing the belief-desire set in action, you must produce the act actually and over variations in the actual circumstances where – as we may take your beliefs to register – the option remains within your ability and suitably appealing. You must control for the realisation of the end you pursue.

This observation shows that acting with the intention of achieving a certain effect – say, that of dropping the weight, to return to our early example – is an instance of controlling for that effect. But not only is an intentional action like dropping the weight bound to involve control, so too is the action that is intentional in the weaker sense distinguished earlier: say, the action of damaging the floor, where that effect is not desired as such but is foreseeably packaged with the desired option of dropping the weight. Despite the fact that it is unwelcome, you control also for that that side-effect. You control for it because, while you might prefer that the floors were not made of wood and were not exposed to damage, you still control for dropping the weight in a range of scenarios in all of which, as you recognise, the floor is of that kind.

The upshot of these points is that acting intentionally, whether in the stronger or weaker senses, means controlling for the effects of the act. But, as we saw in Section 1, every action is necessarily intentional in the stronger sense, and possibly intentional in the weaker. And the fact that this is explicable under the control model, and only under the control model, strongly argues in favour of that view.

The overall picture should now be clear. You produce certain effects in the course of normal actions but you will produce them intentionally – and the actions will count properly as actions – only to the extent that you control for them. Apart from the effects produced, then, your actions may have independent effects associated with the degree of control you exercise over the production. Control with a suitable degree of robustness for producing favour, truth-telling or restraint in dealing with me, and this control will have the extra effect of giving me friendship, honesty or respect. Increase the robustness with which you control for producing any benefit whatsoever, and this in itself will increase the good that you do me: at least, presumptively, the sympathy you bestow.

This picture requires one final amendment, introducing a minor qualification to the framework theory of action sketched in Section 1. The amendment needed, which argues for the primacy of control over production, is that you may control in action for the robust production of a result, where you do not actually produce that result yourself; you rely on an independent productive factor.

Suppose you control, like the cowboy in the classic Western, for getting a herd of cattle to the railroad. The best way of doing this may involve doing nothing when the cattle go spontaneously in the required direction – when an independent factor ensures the desired result – and only intervening when that is required to put one or another animal back on track. This shows that you may control for a given effect, and perform a corresponding action, without actually producing any act that is meant to generate that effect, whether causally or otherwise. The control you exercise may be wholly virtual, being determined, not by anything you actually do, but by what you would do if needed.<sup>54</sup>

Moving away from the case of action, we should note in concluding this discussion, that the control model may be buttressed more generally by the fact that, on many accounts, control is relevant in representing any causal relationship, at least when the representation is meant to be explanatory. According to those accounts, we should take C to cause E – or should invoke C in causal explanation of E – only if it is the case, not just that C produces E in actual circumstances, but that C would have produced E under a certain range of counterfactual variations.<sup>55</sup> But however supportive and congenial those accounts, the argument provided here is meant to stand independently of them.

## 6. Conclusion

The three mistaken claims reviewed in this article are intimately connected with one another. If we make the first mistake – if we think that you can do good only by the causal consequences of your acts – then we will also make the second: limiting the good effects of an action to its causal consequences means limiting them to its behavioural consequences. And if we make the second mistake – if we think that you can do good only by the behavioural consequences of your acts – we must also make the third: limiting the good effects of an action to its disposition-independent consequences means denying that they can include the distinct effect of bringing about those consequences more or less robustly.

The three claims are nonetheless distinct. We may reject the first mistake – we may accept that actions have constitutive as well as causal effects – without rejecting the second: viz., without holding that actions also have robustly demanding effects. And we may reject the second mistake, accepting that actions have robustly demanding effects, without rejecting the third. We might discount the indeterminacy and grounding considerations, reviewed in the fourth section, that link the rejection of the second mistake with the rejection of the third.<sup>56</sup>



**Figure 1:** *The mistakes related in the manner of three concentric circles*

These observations mean that the three mistakes are related to one another in the manner of three concentric but not coincident circles; see Figure 1. In the innermost circle are those positions guilty of the first mistake, in the middle circle are those guilty of the second, and in the outermost circle are those guilty of the third. Thus, we cannot make the first mistake without making the second and the third. But we might make the third mistake without making the second, and we might make the second without making the first.

The mistakes are worth exposing for reasons of both metaphysics and morality. Exposing them enables us, metaphysically, to see the problems in the production model of action, and the attractions of the rival, control model. And exposing them prompts us to reopen questions in moral theory about the importance of neutrally good consequences and about how far the recommendations of a neutralist consequentialism may coincide with those of non-consequentialism. Let the neutrally good consequences associated with action, whether actually or as a matter of expectation, include constitutive as well as causal effects, robustly demanding as well as contingently demanding effects, and even effects in the robustness of control that do not show up as discrete consequences. That makes it much more plausible to think that acting well is a significant function of how much your action thereby improves or promises to improve the world.<sup>57 58</sup>

*Philip Pettit, Center for Human Values, Princeton University, 308 Marx Hall, Princeton, NJ, 08544-1006, USA; School of Philosophy, Australian National University, Canberra, ACT, 2601, Australia. ppettit@princeton.edu*

## NOTES

- 1 This is the case, however, only insofar as it focuses on the hedonic consequences associated with your actions, not on those brought about independently by your motives or maxims or whatever (R.M. Adams, 'Motive utilitarianism', *Journal of Philosophy* 73 (1976): 467–481; for critical perspectives see P. Railton, 'How thinking about character and utilitarianism might lead to rethinking the character of utilitarianism.' *Midwest Studies in Philosophy* 13 (1988): 398–4416; P. Pettit & M. Smith, 'Global consequentialism' in B. Hooker, E. Mason & D.E. Miller (eds) *Morality, Rules and Consequences* (Edinburgh: Edinburgh University Press, 2000)). I work here with a relatively simple, unanalysed notion of hedonism. For the standard work in analysis of the doctrine; see F. Feldman, *Pleasure and the Good Life: Concerning the Nature, Varieties, and Plausibility of Hedonism* (New York: Oxford University Press, 2004).
- 2 A preferentialist form of utilitarianism, aiming at the maximal satisfaction of at least certain sorts of preference, may avoid the mistakes, since the preferences it would have you satisfy may be the preferences of agents overall about what you do under relevant conventions or circumstances, about how you are disposed to act, or about how robustly you aim at certain goals.
- 3 The article draws heavily on themes in my book *The Robust Demands of the Good* (Oxford: Oxford University Press, 2015b) but organises them along different lines. I wrote that book in ignorance of Steven Sverdlík's excellent study *Motive and Rightness*, which defends a congenial line: S. Sverdlík, *Motive and Rightness* (Oxford: Oxford University Press, 2011).
- 4 I assume here without argument that there a suitable, on-off notion of acceptance available. For some discussion, see P. Pettit, 'Making up your mind', *European Journal of Philosophy* 24 (2016): 3–26.
- 5 B. Hedden, *Reasons without Persons* (Oxford: Oxford University Press, 2015).
- 6 I do not explore the issues raised by such intentional omissions. See C. Sartorio, 'Omissions and causalism', *Noûs* 43 (2009b): 513–530, for a useful discussion.
- 7 See D. Davidson, *Essays on Actions and Events* (Oxford: Oxford University Press, 1980). The fact that the belief causes you to let go of the rope, as you desire, does not mean that you must be aware of the desire

- as such. See P. Pettit, 'Preference, deliberation and satisfaction', in S. Olsaretti (ed.) *Preferences and Well-being* (Cambridge: Cambridge University Press, 2006), pp. 131–153, for some discussion.
- 8 T.M. Scanlon *Moral Dimensions: Permissibility, Meaning, Blame* (Cambridge, MA: Harvard University Press, 2008), p. 8
  - 9 A. Goldman, *A Theory of Human Action* (Englewood Cliffs, NJ: Prentice-Hall, 1970).
  - 10 G.E.M. Anscombe, *Intention* (Oxford: Blackwell, 1957); Davidson op. cit.
  - 11 J. Hornsby, *Actions* (London: Routledge, 1980).
  - 12 On related complexities, see C. Sandis, *The Things We Do and Why We Do Them* (London: Palgrave Macmillan, 2012).
  - 13 Strictly, the argument need not reject the claim that the individual contributions of properties to the goodness of options may vary somewhat, depending on the other properties present in any choice. It may be inconsistent with particularism, however, which is an extreme version of that holistic claim (J. Dancy, *Ethics without Principles* (Oxford: Oxford University Press, 2004)). But for a general consideration that tells against particularism, see Pettit 2015b op. cit., pp. 54–59 and for a deeper challenge, see F. Jackson, P. Pettit & M. Smith, 'Ethical particularism and patterns' in B. Hooker & M. Little (eds) *Moral Particularism* (Oxford: Oxford University Press, 2000, pp. 79–99); reprinted in F. Jackson, P. Pettit & M. Smith, *Mind, Morality and Explanation* (Oxford: Oxford University Press, 2004).
  - 14 F. Jackson, 'Decision-theoretic consequentialism and the nearest and dearest objection', *Ethics* 95 (1991): 461–482.
  - 15 Associating probabilities with rightness, I put aside for convenience the issue of how far an act may be good in virtue of the probability of its realising a certain property.
  - 16 By my lights, forming an intention to do something in the future, as distinct from acting with such and such an intention now, may be an action and, by its nature, a mental action. See M. Bratman, *Intention, Plans, and Practical Reason* (Cambridge, MA: Harvard University Press, 1987); M. Bratman, *Faces of Intention: Selected Essays on Intention and Agency* (Cambridge: Cambridge University Press, 1999).
  - 17 Causalism in this sense is distinct from the doctrine to which Sartorio (2009b op. cit.) gives that name; causalism in her sense is close to the production model of action mentioned in the final section. As it happens, Sartorio ('Causation and ethics' in H. Beebe, C. Hitchcock & P. Menzies (eds) *The Oxford Handbook of Causation* (Oxford: Oxford University Press, 2009a), pp. 575–591) considers some interesting problems and responses in relation to causalism in our sense. She herself rejects causalism in that sense and raises problems for causalism in her own.
  - 18 D. Sosa, 'The consequences of consequentialism', *Mind* 102 (1993): 101–122.
  - 19 Goldman op. cit.
  - 20 D. Lewis, *Philosophical Papers Vol 2* (Oxford: Oxford University Press, 1986).
  - 21 The notion of grounding employed here and elsewhere in the article answers broadly to the technical concept developed in recent literature; see G. Rosen, 'Metaphysical dependence: Grounding and reduction' in B. Hale & A. Hoffman (eds) *Modality: Metaphysics, Logic, and Epistemology* (New York: Oxford University Press, 2010, pp. 109–135).
  - 22 W. Sinnott-Armstrong, 'Consequentialism' in E.N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*, 22 October 2015.
  - 23 Pettit 2015b op. cit.
  - 24 Why should I concern myself with goods that are distinguished by their implications for counterfactual situations? Part of the answer is surely that I live in constant awareness of non-actual possibilities, as prospectation theory claims (M.E.P. Seligman, P. Railton, R.F. Baumeister & C. Sripada, 'Navigating into the future or driven by the past', *Perspectives on Psychological Science* (2013) 8: 119–141); for a fuller account, incorporating this element, see Pettit 2015b op. cit., ch. 7. The control model of action, introduced in the final section, makes sense of why this should be so when I interact with other human agents.
  - 25 You might be said to produce an effect robustly over any set of variations, not just the sorts of variations at which we gesture here. To say that you produce an effect robustly is always shorthand for saying that you produce it robustly over certain variations in the circumstances but not over others.
  - 26 A disposition to X is always a disposition to X robustly over certain scenarios but it need not be robustly present; it may be highly susceptible to disruption. The disposition may be dependable, as we might say, without being durable. For a use of this distinction in countering situationist critiques of virtue, see Pettit 2015b op. cit., ch. 2.
  - 27 The disposition may be present in a once-for-all-cases way, so that you are now fully disposed to provide the associated benefit in relevant cases; or it may materialise case-by-case, as when you are predisposed

- for some or all of those cases to form the appropriate disposition only as required; you may do this as a result of conscious effort. See V. McGeer & P. Pettit, 'The hard problem of responsibility' in D. Shoemaker (ed.) *Oxford Studies in Agency and Responsibility* (Oxford: Oxford University Press, 2015, Vol. 3, pp. 160–188; Pettit 2015b op. cit., ch. 1.
- 28 The voluminous literature in virtue theory, as well as the long, mainly Aristotelian tradition on which it draws, does give attention to the importance of dispositions like friendship and honesty and respect, as indeed does much of the Kantian literature. But what is missing in most discussions is the idea that acting out of such dispositions is not just good in itself, or good for the agent – it is not just a case of acting well – it is also good for the person with whom the agent deals.
- 29 Pettit 2015b op. cit., ch. 5; P. Pettit, 'The asymmetry of good and evil' in M. Timmons (ed.) *Oxford Studies in Normative Ethics* (Oxford: Oxford University Press, 2015a).
- 30 T. Hobbes, *Leviathan*, ed. E. Curley (Indianapolis, IN: Hackett, 1994), ch. 15.
- 31 Charles Pigden ('Anscombe on "ought"', *Philosophical Quarterly* 38 (1988): 20–41, at p. 29) ascribes a similar view to Aquinas, using a very similar formulation: 'the unjust man is not one who makes a policy of injustice – merely one who does not make a policy of justice'.
- 32 As argued in Pettit 2015b op. cit., ch. 5 and Pettit 2015a op. cit., pp. 15–37, this asymmetry makes sense of the Knobe effect (J. Knobe 'Intention, intentional action and moral considerations', *Analysis* 64 (2003): 181–187), so called, according to which we more readily ascribe acts of intentionally harming than we do acts of intentionally helping. The asymmetry becomes intelligible, if helping is taken to require robustly keeping relevant standards – say, as in Knobe's original example, environmental standards – and harming is taken just to require contingently breaching them. On that construal, the bar for intentionally helping will naturally be higher than the bar for intentionally harming, and so we will be readier to ascribe acts of intentional harm. For a broadly similar approach, see R. Holton, 'Norms and the Knobe effect', *Analysis* 70 (2010): 1–8.
- 33 H. Arendt, *Eichmann in Jerusalem: A Report on the Banality of Evil* (New York: Viking Press, 1963). For an excellent study that argues in a different way for the relevance of its motive to the value and deontic status of an action, see Sverdlik op. cit.; it focuses, however, on badness or wrongness rather than on goodness or rightness.
- 34 J. Bennett, *The Act Itself* (Oxford: Oxford University Press, 1995).
- 35 A. Donagan, *The Theory of Morality* (Chicago: University of Chicago Press, 1977).
- 36 Bennett op. cit., p. 49.
- 37 Bennett op. cit., p. 46.
- 38 Scanlon 2008 op. cit., ch. 2.
- 39 D. Parfit, *On What Matters*, Vol 1 (Oxford: Oxford University Press, 2011), pp. 208–212.
- 40 Scanlon 2008 op. cit.
- 41 For a congenial critique of Parfit in particular, see B. Herman, 'A mismatch of methods' in S. Scheffler (ed.) *Derek Parfit: On What Matters*, Vol 2 (Oxford: Oxford University Press, 2011), pp. 83–115; and for a more general, equally congenial critique of behaviourism, see C. Korsgaard, *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press, 2009), p. 12. For related discussions see B. Herman, 'On the value of acting from the motive of duty', *Philosophical Review* 66 (1981): 233–250; and M. Baron, 'The alleged moral repugnance of acting from duty', *Journal of Philosophy* 81 (1983): 197–220.
- 42 For other considerations that may be influential, see Pettit 2015b op. cit., ch. 5, and Sverdlik op. cit., ch. 1.
- 43 T.M. Scanlon, *What We Owe To Each Other* (Cambridge, MA: Harvard University Press, 1998).
- 44 Parfit op. cit., p. 25.
- 45 This desideratum on the law, in particular the criminal law, is not a strict constraint: it allows us to require, for example, that agents should have *mens rea*, a culpable mind, in the performance of certain criminal acts, if they are to be punished for them.
- 46 Pettit 2015b op. cit., Appendix II.
- 47 Vague, that is, in a non-epistemic sense: see T. Williamson, *Vagueness* (London: Routledge, 1994). There may be a sorites paradox in the offing here, of course. As one pebble added to a collection of pebbles cannot make it into a heap, so one degree added to a scale in the robust provision of restraint can hardly make it into respect. But even if that sort of paradox arises, it is hard to deny that you get closer and closer to the target – to heaphood or respect – as you add more and more pebbles, more and more degrees of robustness.
- 48 P. Singer, *The Expanding Circle: Ethics and Sociobiology* (Oxford: Oxford University Press, 1981).
- 49 Pettit 2015b op. cit., ch 6 uses the more-robustness principles to explain the appeal of a number of such doctrines, but without justifying them, including the doctrine of double effect. For an overview and

- qualified defence of the doing-allowing distinction, see F. Woollard, *Doing and Allowing Harm* (Oxford: Oxford University Press, 2015).
- 50 Going into that state may not occur automatically or mechanically, of course, in the manner in which the temperature system adjusts. You may do so only agonistically: only as a result of a conscious effort to be respectful. See V. McGeer, 'The moral development of first-person authority', *European Journal of Philosophy* 16 (2008): 81–108.
- 51 See E. Prior, F. Jackson & R. Pargetter, 'Three theses about dispositions', *American Philosophical Quarterly* 19 (1982): 251–257. The disposition, as mentioned in an earlier footnote may be a case-by-case disposition that needs to be maintained by conscious effort, not a once-for-all-cases disposition. See McGeer & Pettit op. cit., Pettit 2015b op. cit., ch. 1.
- 52 T.M. Scanlon (2008 op. cit., p. 24) suggests that there is an inappropriate introspective focus involved in the idea that you might deliberate and act on a basis which makes the appeal of the action turn on whether it displays a disposition like respect (N. Kolodny, 'Scanlon's investigation: The relevance of intent to permissibility', *Analytic Philosophy* 52 (2011): 100–123, at p. 122). But to deliberate on the basis of respect is not to consider the option and ask introspectively whether, if you chose it, you would choose it out of respect. Doing that might serve at best as a useful heuristic. To deliberate on the basis of respect for me is to consider the reasons of respect – say, that I am an adult, able-minded, autonomous individual – and to control for the act that those considerations require. See Pettit 2015b op. cit., pp. 164–67.
- 53 But, it should be noted, they also argue for contracting the set of consequences in virtue of which you may do good or bad: these must be consequences that you control for – or at least are in a position to control for – not just consequences that you accidentally or unforeseeably produce.
- 54 P. Pettit, 'The virtual reality of homo economicus', *Monist* 78 (1995): 308–329. Expanded version in U. Maki (ed.) *The World of Economics* (Cambridge: Cambridge University Press, 2000); reprinted in P. Pettit, *Rules, Reasons, and Norms* (Oxford: Oxford University Press, 2002). Virtually controlling for a result is a way of ensuring that it happens and, like actively controlling for a result, is distinct in the degree of robustness involved, not only from merely allowing the result, but also from enabling it (C. Barry & G. Øverland, *Responding to Global Poverty: Harm, Responsibility, and Agency* (Cambridge: Cambridge University Press, 2016)).
- 55 J. Woodward, *Making Things Happen: A Theory of Causal Explanation* (Oxford: Oxford University Press, 2005); C. List & P. Menzies, 'Non-reductive physicalism and the limits of the exclusion principle', *Journal of Philosophy* 106 (2009): 475–502. See also F. Jackson & P. Pettit, 'Program explanation: A general perspective', *Analysis* 50 (1990): 107–117; reprinted in F. Jackson, P. Pettit & M. Smith, *Mind, Morality and Explanation* (Oxford: Oxford University Press, 2004); F. Jackson & P. Pettit 'In defence of explanatory ecumenism', *Economics and Philosophy* 8 (1992): 1–21; reprinted in F. Jackson, P. Pettit & M. Smith, *Mind, Morality and Explanation* (Oxford: Oxford University Press, 2004). They argue that factors at higher levels can be invoked to causal-explanatory effect only on the assumption that they control or program for the effects they explain: in effect, only on the assumption that no matter how they are realised at lower levels – more or less – a more fundamental realiser event will produce that effect. For a comparison between this programming model and some alternatives, see P. Pettit, 'The program model, difference-makers, and the exclusion problem' in H. Beebe, C. Hitchcock & H. Price (eds) *Making a Difference: Essays on the Philosophy of Causation* (Oxford: Oxford University Press, 2017).
- 56 Something akin to an indictment of the third mistake might still remain available, however; we might argue that robustly demanding goods can be less than fully realised, taking them in that case to constitute partial goods. Equally, for all that has been said, it remains possible to think that good can come about other than by behavioural or discrete consequences – to indict both the second and third mistakes – without holding that there are robustly demanding goods. This would be to hold, for whatever reason, that despite the fact that there are no robustly demanding goods, providing contingently demanding goods more and more robustly involves doing more good.
- 57 P. Pettit, 'A consequentialist perspective on ethics' in M. Baron, M. Slote & P. Pettit (eds) *Three Methods of Ethics: A Debate* (Oxford: Blackwell, 1997); Pettit 2015b op. cit. ch. 7.
- 58 This article is the text for the Annual Lecture of the British Society for Applied Philosophy, delivered on 24 October 2017. I benefitted greatly from discussion of earlier drafts at the University of California at San Diego in Oct 2016 and at the Australian National University in April 2017. I also benefitted enormously from comments from Chris Bennett, Johann Frick, Richard Lindley, Joseph Moore, Brian Hedden, Tori McGeer, Doug Portmore, and Peter Singer. I am particularly indebted to a range of comments from Samuel Preston, which led me to reconceive and reorganise an early version.