Does 'Ought' Imply 'Can'? And Did Kant Think It Does?

ROBERT STERN
University of Sheffield

The aim of this article is twofold. First, it is argued that while the principle of 'ought implies can' is certainly plausible in *some* form, it is tempting to misconstrue it, and that this has happened in the way it has been taken up in some of the current literature. Second, Kant's understanding of the principle is considered. Here it is argued that these problematic conceptions put the principle to work in a way that Kant does not, so that there is an important divergence here which can easily be overlooked.

My aim in this article is twofold. First, I want to consider the plausibility of the principle that 'ought implies can', and in particular to consider how much work this principle can be made to do. I will argue that while the principle is certainly plausible in *some* form, it is tempting to misconstrue it, and that this has happened in the way it has been taken up in some of the current literature. Second, I want to consider Kant's understanding of the principle. Here I will argue that these problematic conceptions put the principle to work in a way that Kant does not, so that there is an important divergence here which can easily be overlooked.

I

The principle 'ought implies can' has been employed in several different debates in ethics and related areas. For example, it has been used to address the issue of free will vs. determinism; of moral dilemmas; of internalism vs. externalism as accounts of moral motivation; of obligation and blame; and of excuses and wrongdoing. None of these ways of using the principle have been entirely free of controversy, in the sense that different sides have disputed the way in which the principle has been employed to argue for one position over another. In these disputes, it is rarely that the principle of 'ought implies can' has been rejected *altogether*; rather, it is usually claimed that while there are clearly some arguments in its favour, these nonetheless establish the principle in a fairly weak form, so that in fact it cannot be used to do what it is being asked to do by one side or other in the dispute. To take one example: In his well-known article 'Obligation and Motivation

© 2004 Cambridge University Press
DOI: 10.1017/S0953820803001055
Utilitas Vol. 16, No. 1, March 2004
Printed in the United Kingdom

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

in Recent Moral Philosophy',[1] William Frankena takes issue with W. D. Falk's attempt to argue from 'ought implies can' to motivational internalism (the view that to have a moral obligation, an agent must have a motivation to act on that obligation). Frankena summarizes Falk's position as follows:

> Falk appeals to the familiar principle that 'I morally ought' implies 'I can,' adding that 'I can' implies 'I want to (in the sense that I have, at least dispositionally, some motivation for doing),' and then [draws] an internalist conclusion.[2]

Frankena then responds to Falk by suggesting that 'ought implies can' may be understood more weakly than Falk's argument requires:

> ['Ought implies can'] may plausibly be understood as saying: (a) moral judgments 'presuppose,' 'contextually imply,' or 'pragmatically imply' that the agent is able to act as proposed or is believed to be, but do not assert or state that he is; or (b) the *point of uttering* moral judgments disappears if the agents involved are not able to act as proposed or at least believed to be; or (c) it would be morally wrong to insist that an agent ought to do a certain action, if he is or is thought to be unable to do it. If Kant's dictum is interpreted in one of these ways the externalist need have no fear, for then it will not serve to refute him.[3]

Frankena thus claims that to use 'ought implies can' to argue for internalism is to misunderstand the principle; in fact, the principle is too weak to establish the desired conclusion.

In general, then, disputes in these areas have arisen because 'ought implies can' has seemed to some to license a particular conclusion that others have disputed, by claiming that in fact the principle is not strong enough to warrant that conclusion. To settle that issue, of course, one must consider the arguments for the principle, which are supposed to support it: do these arguments succeed in establishing it in a (comparatively) weak or a (comparatively) strong sense? Thus, while few would reject the principle altogether, there is disagreement about how exactly it should be understood, and thus about what work it can be made to do, in the light of arguments in its favour.

Now, in this article I want to consider a use for the 'ought implies can' principle which also raises this question, but is different from any so far mentioned. The use is this: It is argued from 'ought implies can' that what is right must be something that we as agents are capable of following or acting upon, so that the principle of 'ought implies can'

[1] William K. Frankena, 'Obligation and Motivation in Recent Moral Philosophy', *Essays in Moral Philosophy*, ed. A. I. Melden (Seattle, 1958), pp. 40–81. Frankena is responding to the following arguments by Falk: 'Morals Without Faith', *Philosophy* 19 (1944), p. 7, and 'Obligation and Rightness', *Philosophy* 20 (1945), p. 139.

[2] Frankena, pp. 59 f.

[3] Ibid., p. 60.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

is said to imply that we should focus on the capacities of agents in moral theorizing and action, and adjust our accounts of what is right and wrong accordingly. It is suggested that the arguments that support 'ought implies can' are sufficient to license a reading of the principle that warrants this conclusion, so that the principle can be employed to help us determine the extent of the normative considered in this way. Against this, I will claim that in fact these arguments are *not* sufficient to license this view of the principle, and that therefore 'ought implies can' is not strong enough to be used in this manner.

II

I will begin by giving some examples of how the principle has been understood in the way that I want to criticize. The examples will come from ethics and epistemology.

The clearest example in ethics is to be found in the work of James Griffin. In a recent book and associated articles, Griffin has argued for what might be called a greater degree of *realism* in ethics, in the sense that we should begin by understanding *ourselves* and our capacities, as a necessary first step to thinking about moral issues. He claims that moral theories have too often neglected facts about human nature and society, and as a result have become distorted and inadequate to our real needs: We have theorized in a vacuum, and so have failed to do so successfully.

A particular example here, Griffin thinks, is utilitarianism. Utilitarianism has a commitment to impartiality, in the sense that it tells us that the right thing to do is whatever maximizes general utility. But, Griffin says, the reality of human life is that we usually cannot either calculate or act on what this maximization demands, because of our natural partiality to family, our interests and other commitments. Griffin therefore claims that human limitations mean that utilitarianism cannot play a genuine role in our lives, and as a result the moral norms it proposes should be rejected as spurious:

> Moral norms must be tailored to fit the human moral torso. They are nothing but what such tailoring produces. There are no moral norms outside the boundary set by our capacities. These are not some second-best norms – norms made for everyday use by agents limited in intelligence and will – and then, behind them, true or ideal norms – norms without compromises to human frailty. Moral norms regulate human action; a norm that ignores the limited nature of human agents is not an 'ideal' norm, but no norm at all.[4]

[4] James Griffin, 'The Human Good and the Ambitions of Consequentialism', *The Good Life and the Human Good*, ed. E. F. Paul, F. D. Miller and J. Paul (Cambridge, 1992), p. 131. Cf. also James Griffin, *Value Judgement: Improving Our Ethical Beliefs* (Oxford, 1996), p. 105.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

Here, then, we find the 'ought implies can' principle being used in the way I am interested in. From certain facts about human nature and capacities, Griffin sets limits to what standards a moral theory can put forward, thereby using 'ought implies can' as a fundamental determinant of what is right and wrong.

It is worth emphasizing that Griffin does indeed want to use 'ought implies can' in this strong way. This is made clear in his discussion of a possible utilitarian response to his claim that impartiality is impossible for us to achieve, both cognitively and motivationally, the response being that impartially promoting interests is not meant to be a decision procedure (how we should go about deciding how to act on a particular occasion), but a criterion (what in the end makes an act right or wrong). Griffin's reply is that this does not help, as any criterion of what is right and wrong must also be constrained by human capacities, otherwise it will become too remote from human practices, and hence will lose its standing as a criterion: 'What most promotes interests is often permanently beyond our reach. Then a would-be "criterion" like that can play no role, not even that of a criterion.'[5] Thus, for Griffin, 'ought implies can' in a strong sense: No act can be right if it is beyond human capacities to act in this way, or wrong if it is beyond human capacities to avoid acting in this way; therefore (he thinks) utilitarianism is mistaken as a moral theory.

Having set out Griffin's position in ethics, we may now consider a second example of the strong use of 'ought implies can', this time in epistemology. The context here is a form of naturalistic response to scepticism, of the sort proposed by P. F. Strawson. According to this response, one way to answer scepticism is to show that there are certain beliefs which we must hold and cannot give up, for example that there is an external world. This response is not without its ambiguities, but one way of taking it is to use it in conjunction with the strong 'ought implies can' principle, so that the sceptic is defeated by arguing that because we cannot give up the belief in question, there is no violation of a cognitive norm here. Thus, as one proponent of the Strawsonian

[5] Griffin, *Value Judgement*, p. 106. Cf. Owen Flanagan, *Varieties of Moral Personality: Ethics and Psychological Realism* (Cambridge, Mass., 1991), pp. 32–8, where Flanagan claims that act utilitarianism fails his Principle of Minimal Psychological Realism as a decision procedure, because it is psychologically impossible to determine which action promotes the best consequences; but he accepts that this does not rule out act utilitarianism as a criterion of rightness: 'For our purposes the point is best put this way: although utilitarianism qua philosophical theory will tell us that the action is best which produces the best outcome, it need not tell us that agents should always act or be motivated to act to produce the best outcome' (p. 34). Griffin's position is significantly stronger than this, in attacking utilitarianism qua criterion of right action *as well as* qua decision procedure.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

position has put it:

> Showing that we must have such a belief as a condition of experience is not the same as proving that such objects exist. One is stating what we must believe, not how things are; but since the sceptic wishes us to justify the belief, doing so the argument goes is enough to put an end to scepticism.[6]

So, in response to the sceptical challenge that a belief we hold is epistemically illegitimate, the naturalistic strategy is to argue that we must hold this belief, and so cannot be violating any cognitive norm in retaining it, in so far as 'ought implies can' (in a strong sense).

I have chosen to highlight these positions because they are particularly clear instances of the outlook I want to criticize. But I do not think they are the only ones. For example, within political philosophy, the criticism of a viewpoint as 'utopian' would in part seem to reflect the idea that a political philosophy cannot be valid unless the principles it proposes are within the capacities of normal human beings to adopt. Of course, this may simply reflect nothing more than a commitment to politics as 'the art of the possible', in which case such theorizing is dismissed not as invalid, but merely as pointless; but it may also reflect the stronger view, that a theory which argues for principles that are unrealizable by us must be wrong, in which case here 'ought implies can' is once again being used in a strong sense.

In what follows, I will go on to claim that if 'ought implies can' is used in this way, the principle is being used *too* strongly. I will argue that while there are plausible arguments for the principle, these arguments only support a weaker reading of it; on the other hand, I will suggest that arguments that might be used to support the stronger reading are not plausible. I will then go on to consider whether Kant's use of the principle shows him to be committed to the weaker or the stronger reading.

III

Let me begin with what is perhaps the most plausible argument for 'ought implies can', namely what I will call the *argument from blame*. The argument, put simply, is that it is wrong to blame someone for something that they cannot control. Many people find this argument plausible, and although there are complexities (for example, can a

[6] A. C. Grayling, 'Transcendental Arguments', *A Companion to Epistemology*, ed. Jonathan Dancy and Ernest Sosa (Oxford, 1992), p. 508. For further discussion of this approach to scepticism, see Robert Stern, *Transcendental Arguments and Scepticism: Answering the Question of Justification* (Oxford, 2000), pp. 107–12, and 'On Strawson's Naturalistic Turn', *Strawson and Kant*, ed. Hans-Johann Glock (Oxford, 2003), pp. 219–33.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

deliberately incurred incapacity negate blameworthiness?), I propose simply to accept the argument for the sake of this discussion. The issue here, however, is whether accepting the argument is sufficient to establish 'ought implies can' in the strong sense.

I do not believe it is, because there is a distinction that can be drawn between *agent* evaluation and *act* evaluation. That is, I can say that you are not to be criticized for doing or believing *A* because you were unable to do otherwise, while still holding that what you did or believed was wrong. So, the fact that an agent cannot be blamed for doing *A* does not show that no wrong was committed, and no norm violated.[7] Thus, the argument from blame shows merely that 'blame implies can', not that 'right implies can'; the principle it establishes therefore cannot be used to argue against a moral theory that says that some acts are right that nonetheless are unachievable by human beings. My claim, therefore, is that while the argument from blame is indeed plausible, it is not sufficient to support a strong reading of the 'ought implies can' principle.

If this is accepted, it therefore appears that a proponent of the strong reading needs to find another argument to support his position. I will consider five such arguments: the argument from obligation; the argument from motivation; the argument from anti-utopianism; the argument from agents; and the argument from naturalism. I will claim that while these arguments might be sufficient to support the strong reading of the principle, none is plausible, so that only a weak version of the principle remains defensible, on the basis of the argument from blame.

The *argument from obligation* is an attempt to reinforce the argument from blame in order to deliver a stronger conclusion, where the problem with that argument was that it appeared that an action might be right, even if an agent could not be blamed for not performing it because of their inability to do so. But, it might be asked, is it coherent to take an action to be right, if no agent *whatsoever* could be blamed for not doing it, on the grounds that *no* agent is capable of performing it? This is not coherent, it could be argued, for it would suggest that an action is right, while no agent is under any obligation to do it, because no agent can perform it. But surely, to be right, the action must be a duty for *some* conceivable agent? Now, I think this point has some force, but it is still not sufficient to support the position of those who

[7] Cf. Walter Sinnott-Armstrong, '"Ought" Conversationally Implies "Can"', *The Philosophical Review* 93 (1984), p. 250: 'Another common argument is that we do not blame agents for failing to do acts which they could not do, so it is not true that the agents ought to have done the acts. No such conclusion follows. The premise is about *agents*, but the conclusion is about *acts*. It is possible that an act ought to be done even though the agent would not be blameworthy for failing to do it.'

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

understand 'ought implies can' in a strong sense. For their claim is that for an act to be right, *we* must be under an obligation to perform it, qua human agents, with our cognitive and motivational limitations; but then it is harder to see why, to be right, an act must be a duty *for us*, something *we* are obliged to perform, any more than it must be a duty for a dog or a monkey. (Of course, this might follow if it was claimed that we were morally exemplary in some way, for then it could be said that unless we (qua moral agents) had *A* as our duty, then *A* would not be good (because if *A* were good, then as exemplary moral agents we would have it as our duty); but this is not part of the position I am criticizing.) So, provided the moral theorist can show how the act they take to be right would be obligatory to *some* conceivable agent (where that agent is without our various limitations), their position would appear to meet the requirements of the argument from obligation; the theorist does not have to show it can be made obligatory *to us*, so that once again our capacities cannot be used to set limits on claims about the right.

The next argument I will discuss, the *argument from motivation*, is this: An act cannot fall under a moral rule unless an agent is capable of obeying that rule, otherwise there would be moral rules that do not engage with our motivational set. What is the basis for this argument? A first suggestion might be that it is somehow unfair that there are rules which determine what is right or wrong, when we are constitutionally incapable of obeying them. But if this is the idea behind the argument, then it really takes us back to the argument from blame, and can be handled by allowing (once again) that we would not be *blameworthy* for failing to act rightly, as this *would* be unfair, while maintaining that there is no unfairness *beyond that*, in the fact that something is right which we cannot do. A second suggestion might be that moral rules have the status of *commands*, and that it makes no sense to issue a command that cannot be obeyed. Now, this view of moral rules is not uncommon; but it is important to distinguish two ways in which it can be taken. One way is (so to speak) *phenomenologically*: we *feel* moral rules to be imperatives or orders, telling us what do to in such a way that we take ourselves to be commanded to do it, and so obliged to act in a certain manner. Taken in this way, it is then plausible to infer that we would not feel *commanded* to act unless we thought we could so act. However, this just shows that our *experience* of morality suggests that we can do what we are morally required to do, in so far as we feel moral rules to be imperatives that apply to us. But of course, this does not show that acts are right only if we are obliged to perform them in this way, and so does not show that they cannot be right unless the commands of morality are addressed to us. It therefore appears that the proponent of this arguments needs a stronger claim,

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

namely, that moral rules are not just experienced by us as commands, but *are* commands, in the sense that they are issued as orders for us to follow, and would not be issued otherwise. How might this position be supported? One option, of course, would be to move to a theistic conception of morality, and to conceive of moral acts as commanded by God, who would not so command us unless he thought we could obey him. But the familiar response to such a conception is with a version of the Euthyphro dilemma: Are such acts commanded because they are right, or are they right because they are commanded? Assuming that the first horn of the dilemma is the more attractive (otherwise God's commands become somehow arbitrary), then the rightness of an act remains prior to its status as a command of God, and the normativity of a moral rule is not to be identified with its status as a command. Another, non-theistic, option brings us to the third attempt to defend the argument from motivation, namely that moral rules, like commands, imperatives and so forth, must be action guiding, otherwise they would be *pointless*. This then takes us to the third argument for the strong version of the 'ought implies can' principle, which is the argument from anti-utopianism.

The *argument from anti-utopianism* is this: If there were moral rules that we could not act on, then these rules would be pointless, and the normative realm would be utopian, full of high ideals that are unrealizable. But this cannot be the case, so these rules must be ones we can obey. Griffin has put this argument as follows:

> Why choose a standard for moral action so remotely connected to what one can do? Of course, 'strange' does not imply 'wrong'. But 'ought' implies 'can'. Action-guiding principles must fit human capacities, or they become strange in a damaging way: pointless.[8]

Now, Griffin recognizes that this argument assumes that a moral rule is supposed to be action-guiding for us, whereas it might just be taken as a criterion of right and wrong. But Griffin also thinks that a criterion that tells us what is right and wrong must respect our capacities, as otherwise it too is pointless:

> Although criterion and decision procedure can diverge, they should not, I think, get too far apart from one another. Our decision procedures must take account of our capacities, but any criterion for a human practice cannot become too remote from our capacities without losing its point even as a criterion. Health is a reasonable criterion for medical practice because doctors can, directly or indirectly, act to bring it about. In contrast with that, a very demanding moral

[8] Griffin, 'The Human Good and the Ambitions of Consequentialism', p. 123. Cf. Griffin, *Value Judgement*, pp. 163 f.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

criterion (say, Jesus's 'Be ye therefore perfect') may go too far even to be a moral criterion.[9]

Now, Griffin's position here rests on two assumptions. First, if a moral rule says that what is right is something we cannot do, it is pointless; and second, if a moral rule is pointless, it cannot really be a moral rule. Against the first assumption, it might be argued that a moral rule that goes beyond our capacities is not *ipso facto* pointless: for example, it may serve as a source of inspiration, or awe. Many of us admire certain figures or acts which we know we could not follow or even try to follow because of our own incapacities, where nonetheless this admiration gives these exemplars a kind of point.[10] And against the second assumption, it could be argued that nothing has been said to support it: Even if a moral rule has no practical point, why should this bear on the normative question?[11]

One possible response to this is an *argument from morality as 'indexed' to agency*: It could be argued that it is mistaken to speak of an action being right or wrong *simpliciter*, as actions are only right or wrong for particular *agents*, where the capacities of the agent then have a direct bearing on the rightness or wrongness of the action. Thus, it could be argued, we cannot just say that 'Act *A* is wrong', but only that '*A* is the wrong thing for person *S* to do', where this judgement requires us to take into account what *S* is capable of doing. To take an analogy: Someone might argue that we cannot say that a particular pastime is worthwhile *simpliciter*, but only whether particular agents should follow it, given their capacities. So, it does not make sense to say of a child that the right thing for him to do is to read the works of Shakespeare, because whether an activity is worthwhile is 'indexed' to the capacities of the agent. Thus, on this view, someone might argue that while it would be morally right for agents capable of greater impartiality than us to act as the utilitarian says, it is not right for *us* so to act, given our capacities – so that utilitarianism cannot form *our* morality (and so must fit the 'human moral torso' in this sense).[12]

[9] Griffin, 'The Human Good and the Ambitions of Consequentialism', pp. 123 f.

[10] Cf. Nicholas Rescher, *Ethical Idealism: An Inquiry into the Nature and Function of Ideals* (Berkeley, 1987), esp. chs. 1 and 6.

[11] Cf. Sinnott-Armstrong, '"Ought" Conversationally Implies "Can"', p. 251: 'Finally, saying that agents ought to do what they cannot do is often claimed to be pointless and therefore not true. This argument is not valid. The premise concerns the *point* or *purpose* of saying something, but the conclusion concerns the *truth* of what is said. What is said might be true even when saying so could not serve any purpose.'

[12] I think this position and argument is one that Griffin himself would not accept, as it still leaves open the possibility of a morality different from ours, which would apply to agents with different capacities, whereas Griffin emphasizes that morality is really only an institution that applies *to us*: 'Moral norms are shaped for us, with all our limitations.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

The difficulty here, however, is first of all that we still need an argument to show that right or wrong requires 'indexing' to the capacities of agents, rather than simply being properties of actions. This is plausible in the case of pastimes, of course, because the value of pastimes lies largely in the benefits they bring to the person pursuing them; so, if a child would get little or nothing from reading Shakespeare, then it is indeed the case that it would not be right for them to do so. But moral actions rarely get their value in this way. Moreover, unless one embraces relativism, there must be some way of assessing the value of an action that is independent of the capacities of the agent, as when we judge that it would be better if the child *could* read Shakespeare rather than just Harry Potter books, or that we *could* act more impartially rather than less so. But surely this requires us to judge the value of the act as right on its own merits, regardless of the capacities of particular agents?

It is likely, however, that these replies will lead to the fifth argument I want to consider, namely the *argument from naturalism*. For, some may feel that these replies show what is fundamentally right about the 'ought implies can' principle: that it prevents the separation of the moral from what is possible for us as human beings, and so stops morality becoming profoundly *unhealthy*, by being conceived of in a way that fails to take our natures into account. This is the difficulty with my positive replies, it will be argued. Talk of 'exemplars', 'unrealizable ideals', 'moral perfection', and so on leads to a radical separation between how we are and what we value that can only be damaging to our self-conception. Against this, the kind of naturalism associated with Nietzsche, Dewey and others would suggest that we should take care to construct a moral system that begins by taking into account human capacities, in order to avoid the life-denying otherworldliness of an abstract realms of 'oughts'.

Now, this clearly raises large issues that cannot be fully dealt with here.[13] A central question is how far the naturalist's concern about the damage done by a morality of this sort is plausible. For, it could be argued that while undoubtedly *blame* and *guilt* can have this corrosive effect, we have already allowed (in discussing the argument from

There are no moral norms outside the boundary set by our capacities' (*Value Judgement*, p. 100).

[13] As Frankena observed of the related dispute between motivational internalism and externalism, the internalist holds that the externalist fails to recognize 'morality's task of guiding human conduct autonomously', while the externalist accuses the internalist of 'having to trim obligation to the size of internal motives': 'The battle, if war there be, cannot be contained; its field is the whole human world, and a grand strategy with a total commitment of forces is demanded of each of its participants. What else could a philosopher expect?' (Frankena, 'Obligation and Motivation', pp. 80 f.).

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

blame) that these moral sentiments are inappropriate in these cases. So, if the moral standard is viewed as exemplary, but in a way that is free from blame and guilt, would this be damaging and harmful? In a well-known passage, Iris Murdoch has suggested a more positive picture:

> Let us consider the case of conduct. What of the command 'Be ye therefore perfect?' Would it not be more sensible to say 'Be ye therefore slightly improved?' Some psychologists warn us that if our standards are too high we shall become neurotic. It seems to me that the idea of love arises necessarily in this context. The idea of perfection moves, and possibly changes, us (as artist, worker, agent) because it inspires love in the part of us that is most worthy. One cannot feel unmixed love for a mediocre moral standard any more than one can for the work of a mediocre artist.[14]

Aside from this big issue, we can also make a narrower point in this context, namely, that if the proponents of the strong version of the 'ought implies can' principle adopt the argument from naturalism, then they are implicitly conceding that they are taking a revisionist position, and so are accepting that this is not how we currently use the principle. For, the argument from naturalism is a critique of our *current* view of morality, from the perspective of a more 'healthy' outlook, where 'can' will determine 'ought'. But if it is a critique of our current view of morality from this perspective, then it is admitted that we currently *do not* conceive of 'ought implies can' in this strong sense. So, one cost of using the argument from naturalism is that the proponent of the strong 'ought implies can' principle must admit that what he says goes against our current understanding of the principle; he is therefore implicitly admitting that he cannot straightforwardly appeal to the principle in attacking some moral system (as Griffin does), for our present understanding of the principle licenses no such criticism. In this sense, then, the naturalist's strategy undermines the dialectical force of the principle when used as a critique of some moral theory. Of course, it is open to the naturalist to *challenge* our present (weak) understanding of the principle and revise it in the light of some argument: but then, as I have tried to suggest in the rest of this section, none of the arguments put forward so far have been successful in this respect.

IV

It seems, then, that the 'ought implies can' principle has been understood by some of its current proponents in a way that is too strong,

[14] Cf. Iris Murdoch, *The Sovereignty of Good* (London, 1970), p. 62.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

because in this strong form it cannot be defended. The question I now wish to raise is a more historical one, namely: Did Kant wish to use this principle in the way it has come to be used recently, and if so how far do my arguments against the principle in this strong form apply to Kant himself?

It is certainly the case that the principle of 'ought implies can' is usually thought of as a Kantian principle, and he is widely seen as bringing it to prominence within modern philosophy.[15] But there is little analysis of what exactly he meant by the principle, or what he supposed it to entail.[16] Moreover, given the common association between Kant and the principle, it is perhaps surprising that there are rather few passages in which Kant actually uses it, and none where he provides any discussion of or argument for it. I will consider those passages which are usually cited as cases where Kant adopts the principle, and I will claim that they show him using it in a weak sense, which suggests that he did not have the strong understanding of the principle that has been discussed and criticized in previous sections.

I will consider the following passages:[17]

Passage A: *Critique of Pure Reason*, A807/B835:

> Pure reason, then, contains, not indeed in its speculative employment, but in that practical employment which is also moral, principles of the *possibility of experience*, namely, of such actions as, in accordance with moral precepts, *might* be met with in the *history* of mankind. For since reason commands that such actions should take place, it must be possible for them to take place.[18]

Passage B: *The Metaphysics of Morals*, 6: 380:

> Impulses of nature, accordingly, involve *obstacles* within the human being's mind to his fulfilment of duty and (sometimes powerful) forces opposing it, which he must judge that he is capable of resisting and conquering by reason not at some time in the future but at once (the

[15] Cf. Frankena's reference to it as 'Kant's dictum' in the passage cited above.

[16] But for a very helpful discussion, which distinguishes different ways in which Kant related 'ought' and 'can' that are overlooked in the contemporary uses of the principle, see Jens Timmermann, 'Sollen und Können: "Du kannst, denn du sollst" und "Sollen impliziert Können" im Vergleich', *Philosophiegeschichte und logische Analyse* 6 (2003), pp. 113–22.

[17] References are given first to the Akademie-Ausgabe, with the exception of the *Critique of Pure Reason*, where reference is given to the A (1781) and B (1787) editions. References are also given to standard translations.

[18] Immanuel Kant, *Critique of Pure Reason*, trans. Norman Kemp Smith (London, 1933), p. 637.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

moment he thinks of duty): he must judge that he *can* do what the law tells him unconditionally that he *ought* to do.[19]

Passage C: *Religion Within the Boundaries of Mere Reason*, 6: 47:
But if a human being is corrupt in the very ground of his maxims, how can he possibly bring about this revolution of his own forces and become a good human being on his own? Yet duty commands that he be good, and duty commands nothing but what we can do.[20]

Passage D: *Religion Within the Boundaries of Mere Reason*, 6: 50:
For if the moral law commands that we *ought* to be better human beings now, it inescapably follows that we must be *capable* of being better human beings.[21]

Passage E: *Critique of Practical Reason*, 5: 143, footnote:
[T]he moral law ... necessarily binds every rational being and therefore justifies him a priori in presupposing in nature the conditions befitting it and makes the latter inseparable from the complete practical use of reason. It is a duty to realize the highest good to the utmost of our capacity; therefore it must be possible; hence it is also unavoidable for every rational being in the world to assume what is necessary for its objective possibility. The assumption is as necessary as the moral law, in relation to which alone it is valid.[22]

Passage F: *Critique of Practical Reason*, 5: 142:
[A] need *of pure practical* reason is based on a *duty*, that of making something (the highest good) the object of my will so as to promote it with all my powers; and thus I must suppose its possibility and so too the conditions for this, namely God, freedom, and immortality, because I cannot prove these by my speculative reason, although I can also not refute them.[23]

Passage G: 'On the Common Saying: That may be Correct in Theory, but it is of no Use in Practice', 8: 276–7:
But in a theory that is based on the *concept of duty*, concern about the empty ideality of this concept quite disappears. For it would not

[19] Immanuel Kant, *The Metaphysics of Morals*, trans. Mary J. Gregor, in Immanuel Kant, *Practical Philosophy*, trans. and ed. Mary Gregor (Cambridge, 1996), p. 513.
[20] Immanuel Kant, *Religion Within the Boundaries of Mere Reason*, trans. George di Giovanni, in Immanuel Kant, *Religion and Rational Theology*, trans. and ed. Allen W. Wood and George di Giovanni (Cambridge, 1996), p. 92.
[21] Kant, *Religion Within the Boundaries of Mere Reason*, p. 94.
[22] Immanuel Kant, *Critique of Practical Reason*, trans. Mary J. Gregor, in Immanuel Kant, *Practical Philosophy*, trans. and ed. Mary J. Gregor (Cambridge, 1996), p. 255.
[23] Kant, *Critique of Practical Reason*, p. 254.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

be a duty to aim at a certain effect of our will if this effect were not also possible in experience.[24]

Passage H: 'On the Common Saying: That may be Correct in Theory, but it is of no Use in Practice', 8: 278–9:
I explained morals provisionally as the introduction to a science that teaches, not how we are to become happy, but how we are to become worthy of happiness. In doing so, I did not fail to remark that the human being is not thereby required to *renounce* his natural end, happiness, when it is a matter of complying with his duty; for that he cannot do, just as no finite rational being whatever can.[25]

Passage I: *Religion Within the Boundaries of Mere Reason*, 6: 62:
From the practical point of view this idea [of a moral exemplar] has complete reality within itself. For it resides in our morally-legislative reason. We *ought* to conform to it, and therefore we must also *be able* to.[26]

Passage J: *Religion Within the Boundaries of Mere Reason*, 6: 64:
For let the nature of this human being well-pleasing to God . . . be thought as superhuman . . . inasmuch as his unchanging purity of will, not gained through effort but innate, would render any transgression on his part absolutely impossible. The consequent distance from the natural human being would then again become so infinitely great that the divine human being could no longer be held forth to the natural human being as *example* . . . [T]he idea of a conduct in accordance with so perfect a rule of morality could no doubt also be valid for us, as a precept to be followed. Yet he himself could *not* also be presented to us *as an example to be emulated*, hence also not as proof that so pure and exalted a moral goodness can be practised and attained *by us*.[27]

Passage K: *Critique of Pure Reason*, A548/B576:
The action to which the '*ought*' applies must indeed be possible under natural conditions.[28]

Taken together, these represent the main examples where Kant uses the principle of 'ought implies can', and so can best help us gauge his

[24] Immanuel Kant, 'On the Common Saying: That may be Correct in Theory, but it is of no Use in Practice', trans. Mary J. Gregor, in Immanuel Kant, *Practical Philosophy*, trans. and ed. Mary J. Gregor (Cambridge, 1996), p. 280.
[25] Kant, 'On the Common Saying', pp. 281 f.
[26] Kant, *Religion Within the Boundaries of Mere Reason*, p. 105.
[27] Ibid., pp. 106 f.
[28] Kant, *Critique of Pure Reason*, p. 473.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

views. The question, then, is not whether Kant adopted the principle, but which *version* of the principle: How strongly did he understand it? In order to consider this, I will discuss the passages in groups, beginning with Passages A, B, C and D.

On the face of it, these passages may seem to show that Kant is indeed using the principle in a way that suggests he understands it in a strong sense, where, for example, Kant claims that 'duty commands nothing but what we can do' (Passage C). This surely is to argue that the moral law is fixed by our capacities, which is what Griffin and others would also claim. However, I would suggest that in fact, these passages are not quite so straightforward. For, what Kant is focused on here is not the moral law *as such* (so to speak), but how the moral law relates to us, as something that commands *us* (Passages A, C and D), that tells *us* what to do unconditionally (passage B), that has a certain *authority* over us.[29] Kant holds that the moral law is addressed to us as agents, as something we are called upon to carry out, so that (as Henry Allison has put it) for Kant 'the moral law confronts us not merely as a lofty and admirable ideal but also as a source of an unconditional, inescapable demand upon the self';[30] and the moral law could not relate to us in this way unless we were capable of acting upon it, otherwise it would appear as no more than a 'lofty and admirable ideal' rather than as an 'inescapable demand'. So, the way we experience the 'right', as something we are in fact obliged to follow, shows that we think of ourselves as capable of acting as the moral law prescribes, and that we must accept the conditions that (Kant believes) explain this possibility, such as free will and the existence of God (cf. Passage E). But nothing here shows that Kant is committed to the view that the moral law *as such* must be constrained by the capacity of agents to obey it, as is suggested by the strong version of the 'ought implies can' principle. Nor does his argument for the postulates (of God, freedom and immortality) require this. The fact that we are obliged to act in accordance with the moral law is sufficient to make these postulates rational, as explaining how we can come to be obliged in this way. Kant does not need the stronger claim, that these postulates make it possible for us to follow the moral law, and unless we could follow it, it would lose all normative content. So, I would argue, a closer inspection of these passages shows

[29] Cf. also Immanuel Kant, *Toward Perpetual Peace*, 8: 370, trans. Mary J. Gregor, in Immanuel Kant, *Practical Philosophy*, trans. and ed. Mary J. Gregor (Cambridge, 1996), p. 338: 'Morals is of itself practical in the objective sense, as the sum of laws commanding unconditionally, in accordance with which we *ought* to act, and it is patently absurd, having granted this concept of duty its authority, to say that one nevertheless *cannot* do it. For in that case this concept would of itself drop out of morals (*ultra posse nemo obligatur*) [no one is obligated beyond what he can do].'

[30] Henry E. Allison, *Kant's Theory of Freedom* (Cambridge, 1990), p. 68.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

that Kant is using the principle of 'ought implies can' in a weak sense, by arguing that the moral law only has its status of being obligatory for us because we are able to act upon it, and that we can thus only explain this obligatoriness by accepting certain claims about our capacities and their conditions ('we ought implies we can'). But this is distinct from the claim that no act can be right (rather than just obligatory for us) unless we are able to perform it, which is how the principle is understood in the strong sense.

It might be argued, however, that this way of taking Kant's position is highly problematic, as it would appear to open up a gap between what I have called 'the moral law as such', and our relation to it as moral agents. Surely Kant would have denied that the moral law could obtain *without* our being commanded by it, in which case isn't the latter really constitutive of the former? My response to this objection is to dispute this strong claim as a reading of Kant. Of course, Kant certainly held that *given what we are*, the moral law is a command to us: but that is (so to speak) a fact about us, rather than a fact about the moral law, that it must be such that it can be commanded to human agents.[31] So, in so far as Kant thinks that nothing can prevent the moral law commanding us, he does so because his conception of our agency is such that he holds us to be essentially capable of acting as right requires, not because our capacities as human agents naturalistically conceived puts limits on what the moral law can comprise, so that no such gap can arise based on his conception of us qua exemplary agents, and not merely on his conception of 'ought implies can'.[32]

[31] Cf. Kant, *The Metaphysics of Morals*, 6: 399 (trans. Mary J. Gregor, p. 528), where Kant says that every human being has the predispositions of '*moral feeling*, *conscience*, *love* of one's neighbor, and *respect* for oneself (*self-esteem*) . . . and it is by virtue of them that he can be put under obligation. Consciousness of them is not of empirical origin; it can, instead, only follow from consciousness of a moral law, as the effect this has on the mind.'

[32] Cf. Immanuel Kant, *Groundwork of the Metaphysics of Morals*, 4: 389, trans. Mary J. Gregor, in Immanuel Kant, *Practical Philosophy*, trans. and ed. Mary J. Gregor (Cambridge, 1996), p. 44, where he suggests that he has worked out 'for once a pure moral philosophy, completely cleansed of everything that may be only empirical and that belongs to anthropology'. Cf. also *The Metaphysics of Morals*, 6: 405 f. (trans. Mary J. Gregor, p. 533): 'Ethical duties must not be determined in accordance with the capacity to fulfil the law that is ascribed to human beings; on the contrary, their moral capacity must be estimated by the law, which commands categorically, and so in accordance with our rational knowledge of what they ought to be in keeping with the idea of humanity, not in accordance with the empirical knowledge we have of them as they are.' It is hard to see he could make these claims unless he understood 'ought implies can' in the weak sense suggested here; for of course, it is precisely by understanding it in the strong sense that theorists like Griffin hope to bring such empirical and anthropological issues back in. Kant makes clear his hostility to any naturalistic arguments for doing so in his *Lectures on Ethics*, 27: 294 f. and 301 (ed. Peter Heath and J. B. Schneewind, trans. Peter Heath (Cambridge, 1997), p. 86 and pp. 91 f.).

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

Let me now consider Passages E and F. Once again, there is some reason to take these passages as supporting the view that Kant had a strong reading of the 'ought implies can' principle. However, I also think appearances here are misleading. For, what Kant is focusing on here is the *content* of the moral law as it is addressed to us, namely that it tells us that we must 'realize the highest good to the utmost of our capacity' (Passage E) and that I must 'promote [the highest good] with all my powers' (Passage F). Thus, Kant clearly thought the moral law was such as to demand of us that we use our abilities in such a way as to further the highest good, and that this demand would be incoherent unless we could so act. But again, this is a fairly weak use of the 'ought implies can' principle: it takes it that there would not be a moral rule telling *us* how we ought to use *our* capacities unless we could so use those capacities, which seems very plausible, because unless we had those capacities (in ourselves, or internally) and were able to exercise them (in the world, or externally), how could our use of them be prescribed by the moral law? Thus, on this view, the moral law engages with us (tells us what to do) because it fits with our capacities; but this is distinct from the stronger view, that *unless* it engages with our capacities, and so states what is right *for us* to do, then it cannot tell us what is right *at all*.[33] To take an example: It certainly seems correct to say that the moral law could only say that 'people should try as hard as possible to act benevolently' if people are capable (both internally and externally) of acting benevolently, as otherwise the injunction to try would be meaningless. But this is a moral law that concerns our capacities, and it does not follow that if we lack that capacity, we can no longer say that benevolent acts are right. Thus, Kant would seem to be arguing here merely that an agent can only be under an obligation to try to bring about an end (the realization of the highest good) if he has the ability to do it. But this does not commit Kant to denying that the end (the realization of the highest good) is right, even if no human agent is under any such obligation to use his abilities to bring it about, because he lacks those abilities.

Similar remarks apply to Passage G and H. These occur in Kant's essay 'On the Common Saying: That may be Correct in Theory, but it is of no Use in Practice', which is in part a reply to Christian Garve's criticism that for Kant, we have a moral duty to give up our desire to be happy, but that this is contrary to our nature. In Passage G, Kant

[33] As I have observed above, this might follow with the additional premiss that we are morally exemplary in some way, so that what is morally right is then necessarily something we are capable of acting upon, and so necessarily engages with us in this way; but this is an additional step, rather than something that follows from 'ought implies can' on its own.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

accepts the seriousness of this kind of worry, but in Passage H he says that Garve is wrong to claim that he saw our moral duty in this way. Now, again, these passages may appear to support a strong reading of 'ought implies can'. But, once more, I believe this is mistaken. For, it is clear on closer inspection that Kant is talking about what the moral law commands *us* to do, what *our* obligations consist in, where Kant accepts that something cannot be an obligation for *us* unless we can bring it about; but, as I have argued, this is a weak conception of 'ought implies can'. And, in Passage H, Kant says it cannot be someone's duty to do something he is incapable of doing, which again fits with the weak conception. What he does *not* say here is anything that implies the strong reading: namely, that nothing can be right that we are incapable of achieving; rather, all he seems to be accepting is that we cannot *be obliged to do what is right* unless we are capable of acting in that way, which is a weaker claim than the one made by Griffin et al.

Turning now to Passages I and J, these concern the role of a Christ-like figure as a moral example to us. Again, these passages may seem to support a position like Griffin's, which opposes the notion of moral ideals that we cannot follow, and thus the counsel of perfection 'Be ye therefore perfect'. Nonetheless, I would claim that it would be wrong to assimilate Kant to this view, for Kant's focus is on Christ as an actual *example*, as a figure we are told we ought to imitate, and would be to blame if we did not. Now of course, if the example of Christ is treated in this way, then we must be *able* to imitate him. But this does not show that there could not be moral ideals which were still valid, even though we could not imitate them (in these cases the most we could do would be to *try* to imitate them, and so would not be to blame for failing to attain any real likeness to our ideal). Indeed, Kant seems to accept the 'validity' of these ideals himself, as something we might strive to copy, without being able to imitate. I do not believe, therefore, that these passages show that Kant had anything resembling the naturalist's concerns discussed earlier. He is merely suggesting that if we are to be blamed for not following a moral example, then our capacities must resemble those of the example; but as we have seen, this argument from blame can be accepted without being committed to the strong interpretation of the 'ought implies can' principle.

Finally, let me consider Passage K. This passage is in fact the one most often cited in support of the claim that Kant had a strong conception of 'ought implies can'. And certainly, taken on its own, it seems fairly compelling. Nonetheless, it is much less conclusive when set in context, as follows:

> The action to which the '*ought*' applies must indeed be possible under natural conditions. These conditions, however, do not play any part in determining the

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

will itself, but only in determining the effect and its consequences in the [field of] appearance.

I would argue that the context alters the implication of the original sentence. For now it seems clear that Kant merely intended that sentence to say that 'natural conditions' (i.e. the phenomenal world) must play a role in determining moral actions as well as other actions, but that that role is not a determinant of the will. The passage is therefore all about Kant's view of moral action and the will, and not about 'ought implies can' at all, and is therefore of no real concern to us here.

I think the discussion of these passages suggests something more generally about the difference between Kant's outlook and the stronger use of 'ought implies can' that we have been analysing. A feature of this stronger use is that it tries to develop an account of right and wrong by *beginning* with an account of human capacities, to set the parameters of moral theorizing. (As Griffin puts it: 'The limits of "ought" are fixed by, among other things, the limits of "can".'[34]) I think Kant's procedure is in many ways the opposite of this: that is, he first fixes his moral theory, in which what matters is not what we are capable of qua human beings, but what obligations can be shown to apply to rational agents capable of acting rightly; and then, once the moral law is fixed, he uses 'ought implies can' to determine what we are capable of qua human beings, in so far as we fall under this law.[35] Thus, whereas the strong conception argues from what we can do to what we ought to do, Kant's weaker conception of 'ought implies can' argues from what we ought to do to what we can do, and so is used to provide his ethical argument for freedom and the existence of God.[36] It is therefore hardly any surprise

[34] Griffin, *Value Judgement*, p. 96.

[35] Cf. Kant, *Critique of Practical Reason* 5: 36 (trans. Mary J. Gregor, p. 169): 'But the moral law commands compliance from everyone, and indeed the most exact compliance. Appraising what is to be done in accordance with it must, therefore, not be so difficult that the most common and unpracticed understanding should not know how to go about it, even without worldly prudence.'

[36] Griffin therefore emphasizes the 'modesty' of his moral position: 'Ethics, particularly the ethics studied in modern universities, strikes me as often too ambitious. It usually fails to operate with a realistic conception of human agency' (*Value Judgement*, p. 100). While Griffin uses 'ought implies can' as a justification for such modesty, I would argue that Kant had no desire to use the principle in this way, but only to establish what our capacities are, in the light of the demands of morality (which in Kant's case are of course notoriously high). Cf. Kant, *Critique of Practical Reason* 5: 30 (trans. Mary J. Gregor, pp. 163–4), where Kant argues that because a person judges he ought not to give false testimony even on pain of death, he will therefore judge that he can overcome his love of life, when 'ought implies can' is used in this way: 'He judges, therefore, that he can do something because he is aware that he ought to do it and cognizes freedom within him, which, without the moral law, would have remained unknown to him.' As Timmermann argues (in 'Sollen und Können') the Kantian principle in these cases is best rendered not

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press

that, on close inspection, Kant's position diverges from the current one.

V

We have therefore considered how far the principle of 'ought implies can' should be used, and the extent to which Kant (as one of the proponents of this principle) wanted to use it. If my discussion has been right, some current uses of the principle go too far, and in a direction that receives little support from the comments on this principle made by Kant himself.[37]

r.stern@sheffield.ac.uk

as 'ought implies can', but as 'you can because you ought' ('Du kannst, denn du sollst'), in so far as here the 'ought' is presupposed rather than being used to determine whether a duty can properly be demanded of an agent.

[37] I am grateful to Christopher Bennett, Fabian Freyenhagen, David Owens, Walter Sinnott-Armstrong, Philip Stratton-Lake, Jens Timmermann and Leif Wenar for helpful comments on earlier versions of this article.

https://doi.org/10.1017/S0953820803001055 Published online by Cambridge University Press