

Articles

Do Moral Dilemmas Tell against the Consistency of a Given Moral System?

Jakub Bożydar Wiśniewski
Queen Mary, University of London

1. Introduction and Conceptual Clarifications

In this article I shall look at the issue of moral dilemmas and attempt to analyze its implications for the structure of normative moral systems. My contention is that the phenomenon of moral dilemmas, though real and capable of placing us under two or more jointly unsatisfiable obligations (or duties, moral requirements, etc.), should not be considered a defect of rationality that requires eradication, but, on the contrary, can and should be accepted as an integral part of the application of moral systems, a part oftentimes capable of playing an important corrective role.

In arguing for the above-mentioned conclusion, I shall look at two distinct criteria of consistency¹ that can be applied to the assessment of a given moral system—they are both based on certain standard principles of deontic logic, the second incorporating some significant insights from the field of modal logic. The first criterion assumes that a satisfactory moral system, if properly followed, cannot allow for the occurrence of genuine, irresolvable moral conflicts in the actual world; it can allow for the occurrence of apparent or *prima facie* moral conflicts, but in such cases it must firmly point toward one specific course of action among a pair or a range of apparently conflicting options. The second criterion, on the other hand, counts among consistent moral systems those which allow irresolvable moral conflicts in the actual world, but also holds that there are possible worlds in which no such conflicts occur.

My thesis is that the second criterion should be endorsed as the more plausible of the two, but some of my remarks are intended to suggest that on a

¹ By consistency I mean the absence of logical contradictions in a given theoretical system, which in the case of moral systems translates into the non-existence of situations in which it is logically impossible to meet all of the moral obligations that a given system imposes upon its adherent.

sufficiently strict formulation of what qualifies as genuine adherence to a given moral system, the respective roles that these two criteria assign to moral dilemmas appear quite similar. Having described the overall structure of my work, let me now introduce a few further conceptual clarifications pertinent to the issues at hand.

The category of normative systems that I consider to be relevant to my analysis are rationalist moral systems, by which I mean systems whose principles are to be discovered or constructed by means of the intellect or, more specifically, by means of informal deductive reasoning. Some further, more detailed formal characteristics of rationalist moral systems are specified by Alan Donagan, who writes that

(1) they [the systems in question] rest on a few fundamental principles . . . , which are advanced as true without exception; (2) each of these principles lays down some condition upon all human action as being required by practical reason; (3) those principles do not constitute a set of axioms, from which all the remaining moral precepts of the theory can be deduced; but, rather, (4) the remaining moral precepts are deduced from the fundamental principles by way of additional premises specifying further the conditions those principles lay down as required of all human action.²

As regards the definition of moral dilemma that I shall adopt here, I rely on the summary of the phenomenon in question provided by Ragnar Ohlsson:

In a moral dilemma, the agent acts wrongly whatever he does. Either all available alternatives are forbidden, or two or more actions that cannot conjointly be performed are morally required in the same situation, or one and the same action is both forbidden and absolutely obligatory.³

2. The First Criterion

The first criterion of consistency that I shall look at (hereafter called “the first criterion”) assumes that a consistent moral system, if properly followed, cannot allow for the occurrence of genuine, irresolvable moral conflicts in the actual world. However, as one of my guiding presuppositions is that since moral conflicts are real, it seems to follow from this criterion that although they initially create an impression of logical inconsistency, a well-formed normative system which allows for their occurrence should also allow for

² Alan Donagan, “Consistency in Rationalist Moral Systems” *Journal of Philosophy* 81 (1984), p. 293.

³ Ragnar Ohlsson, “Who Can Accept Moral Dilemmas?” *Journal of Philosophy* 90 (1993), p. 405.

explaining away this impression. Thus, we should ask ourselves: When does a dilemma arise and what rationalist-friendly explanations for its appearance can be adduced?

Two articles written by Bernard Williams in the mid-1960s seem to offer some suggestions on how to answer the first part of the above question. In “Ethical Consistency,”⁴ Williams can be read as saying that a mark of a genuine moral dilemma is the feeling of regret that follows the decision to act on one of two or more conjointly unsatisfiable obligations; in other words, acting on one of two or more apparently conflicting *oughts* does not eliminate the other *ought(s)* from the scene. In “Consistency and Realism,”⁵ on the other hand, he points to situations in which the agent is unable even to reach the stage of feeling regret, since, all things considered, he is in a position where no reasons capable of justifying one rather than another course of action are forthcoming. Williams’s intention is to use these remarks to question the viability of moral realism, but I believe that they could be used in attempts to undermine rationalism as well, since they suggest that there exist difficult moral scenarios in which reason, intellect, or deduction can offer us only little or no help.

Philippa Foot presents some arguments aimed at countering Williams’s objections, which might be useful in elucidating the sense in which a dilemma can be seen as something less than an irresolvable moral conflict.⁶ She distinguishes between, as she puts it, type 1 and type 2 ought statements. Type 1 ought statements assert that one ought to do a certain thing in virtue of certain specific reasons. Thus, when an agent is confronted with a dilemma, he is also confronted with diverse reasons for action—but even if one of them ultimately prevails, it does not follow that the other ceases to be a good reason, and the agent’s awareness of that fact might make him feel regret for not actualizing certain positive outcome(s). The regret in question, however, constitutes an emotional remainder, not a logical remainder, so we need not think that failure to act on one of the jointly unsatisfiable type 1 ought statements impugns the power of our moral intellect or indicates an inconsistency in the moral system to which we adhere. By way of further conceptual clarification, in this context I use the phrase “logical remainder” to refer to a reason that, in hindsight, required being acted upon on pain of inconsistency. “Emotional remainder,” on the other hand, refers to a reason that, in hindsight, required being acted upon on pain of sacrificing a certain value (and feeling a subsequent painful sensation of loss), though the

⁴ Bernard Williams, “Ethical Consistency,” *Proceedings of the Aristotelian Society* 39 (1965), pp. 103-24.

⁵ Bernard Williams, “Consistency and Realism,” *Proceedings of the Aristotelian Society* 40 (1966), pp. 187-206.

⁶ Philippa Foot, “Moral Realism and Moral Dilemma,” *Journal of Philosophy* 80 (1983), pp. 384-86.

attainment of this value was not logically required by a given normative system.

Type 2 ought statements, on the other hand, assert not what one ought to do in virtue of any specific reason (or the corresponding duty, obligation, moral principle, etc.), but what one ought to do all things considered. If, having applied the entire strength of one's rational faculties, one still faces a pair of obligatory alternatives, neither of which is less significant than the other, then it seems that no type 2 ought statement is forthcoming. Consequently, the worry here could be that one is following an inconsistent system of moral guidance, which should be revised and perhaps based on something more than pure deductive reasoning, just as an experiment yielding inconsistent results can be thought to be based on a defective procedure.

However, there seems to be in principle no reason to suppose that the structure of the moral domain is the same as the structure of, say, the experimental domain. As Foot points out, it is not an unfamiliar idea that the former may include an element of incommensurability (this is not to say, of course, that considerations of incommensurability were not raised in relation to the latter area as well). Perhaps the aforementioned scenarios, in which there is seemingly nothing conclusive to be said in favor of either of the jointly unsatisfiable obligations, should be recognized as containing no truth of the matter as to which action is *the* proper one, *the* morally correct one. Having said that, though, it is now crucial to investigate whether incommensurability should be treated as a friend of rationalism or as its enemy.

At this point, it might be worthwhile explicitly to unpack the term in question. By incommensurability with regard to a given set of values I mean the impossibility of reducing those values to a common normative denominator, which implies the inability to judge them according to the three standard comparative relations ("better than," "worse than," and "equally good"). The main worry here is that, in cases where the values among which a choice is to be made are genuinely incommensurable, practical reason or intellect could be seen as incapable of offering us any final guidance, that is, incapable of performing an act of comparison against a common standard. It is therefore important to look at the purported examples of such cases and see whether the difficulties about decision-making that they induce can be said to expose any deficiencies of rationalism.

Let us begin with the familiar example known as "Sophie's Choice."⁷ It involves a scenario in which a woman is forced by a concentration camp guard to choose which of her two children is to be spared and which is to be killed. Extra poignancy is added to the situation by the fact that Sophie's refusal to make a choice will result in both children being killed—a factor that provides her with an irrefutable reason to act rather than remain passive. Is

⁷ See William Styron, *Sophie's Choice* (New York: Bantam Books, 1980); Patricia Greenspan, "Moral Dilemmas and Guilt," *Philosophical Studies* 43 (1983), pp. 117-25.

this a *bona fide* case of incommensurability in morality? On a closer look, it seems that rather than being a genuine instance of a choice between incommensurable values, the scenario under discussion contains an instance of a choice between incomparable outcomes, each of which essentially exemplifies an attachment to *the same* set of values—in this case, maternal love, familial devotion, care, etc. Thus, it appears plausible to conclude that Sophie’s predicament is not a moral dilemma, but a practical dilemma—it no longer belongs to an area in which moral reasons play any action-guiding role, so it should not be taken as potentially damaging to the viability of rationalism.

A caveat might be in place here: moral reasons do not play any action-guiding role in the situation in question only if we assume that Sophie is able to bring herself to make a choice (provided that she really does not consider one of her children to be more valuable than the other). If, however, we assume her to be unable to do so⁸ (since, e.g., we conceive of her as incapable of living with herself whatever she chooses), then we might consider her rejection of the original choice as the recognition of a moral reason not to play an active part in condemning any of her children to death, a reason sufficient to trump the disvalue of carrying on her life with neither of them at her side. Such a decision would indicate that within the moral system to which Sophie adheres, it is possible to compare the worth of acceptance and the worth of rejection of the original choice with respect to the yardstick of some more comprehensive value, even if that value’s name does not figure in any commonly known moral dictionary.⁹ Such an analysis rules out the possibility of the scenario in question’s involving an instance of value incommensurability.

And yet, one might insist that a dilemma where two jointly unsatisfiable outcomes are generated by a *single* norm or by one’s attachment to a *single* value is still a moral dilemma after all. Such a claim, however, would require specifying the element responsible for making a potential resolution in such cases difficult to find. It cannot be a difficulty associated with value choice, since such a possibility is ruled out by definition. Perhaps it can be a difficulty stemming from the fact that the conflicting outcomes at hand are sufficiently different or directed at sufficiently distinct objects? In Sophie’s case, for instance, such a characterization could apply to her children.

This, I believe, is only an ostensible problem. Its appearance seems to hinge on the fact that the descriptions of such putatively problematic scenarios

⁸ In a strict sense, her inability to choose which of her children to save is a choice in itself. Likewise, if her resultant passivity is to be thought of as more than a result of being incapacitated by fear, then there is clearly a conscious mental activity behind it.

⁹ Ruth Chang, “Putting Together Morality and Well-Being,” in *Practical Conflicts: New Philosophical Essays*, ed. Peter Baumann and Monika Betzler (Cambridge, MA: Cambridge University Press, 2004), pp. 118-58.

are too coarse-grained. Take another case involving what seems to be a pair of incomparable outcomes stemming from adherence to a single value or moral principle—a conflict between fulfilling a promise made to one’s mother and fulfilling a promise made to one’s best friend. My contention is that unless the value in question can be further subdivided into, say, the value of being dependable in family dealings and the value of being dependable in dealings with friends (with one of these values trumping the other), there is no moral reason to prefer one outcome rather than another and either is equally justifiable. I take Sophie’s case to admit of a parallel solution: unless one of her children exemplifies some special moral qualities, she can justifiably choose to save either of them, even though that may not alleviate her subsequent grief and possible (emotional) regret. In both cases the crucial point is to make the description of the situation as fine-grained as one’s moral sensibilities allow. This seems to me sufficient to ensure that none of the above scenarios is capable of harming the viability of rationalism.

There are, however, other familiar dilemmatic stories, whose harmlessness might be more difficult to establish. Cases such as that of Antigone¹⁰ (who has to choose between burying her brother, thus performing duties given by gods, and refraining from doing so, thus respecting the order of her king) or that of Sartre’s student¹¹ (who is torn between going to England to join the Free French and staying with his mother and helping her to live), seem to be more likely candidates for exhibiting pairs of authentically incommensurable choices. The origin of this authenticity is the fact that the sets of values among which the heroes of the aforementioned stories must choose are grounded in obligations (or duties, principles, etc.) stemming from different sources,¹² whose lexical ordering seems implausible.

If, given such predicaments, deductive reasoning offers us no decisive push in either direction, then it appears that genuine incommensurability can threaten to render rationalist moral systems incomplete or otherwise deficient. Similar conclusions can be taken to follow, for example, from Simon Blackburn’s discussion of dilemmas.¹³ According to Blackburn, as soon as an agent finds himself in a stable quandary (that is, in a situation where one does not know how to act and where no “further exercise of thought or imagination

¹⁰ Sophocles, “Antigone,” in *The Tragedies of Sophocles*, trans. Richard C. Jebb (Cambridge, MA: Cambridge University Press, 1917).

¹¹ Jean-Paul Sartre, “Existentialism Is a Humanism,” in *Existentialism from Dostoevsky to Sartre*, ed. Walter Kaufmann, trans. Philip Mairet (New York: Meridian, 1957 [1946]), pp. 287–311.

¹² E. J. Lemmon, “Moral Dilemmas,” *The Philosophical Review* 70 (1962), pp. 139–58.

¹³ Simon Blackburn, “Dilemmas: Dithering, Plumping, and Grief,” in *Moral Dilemmas and Moral Theory*, ed. H. E. Mason (New York: Oxford University Press, 1996), pp. 127–39.

can reasonably be expected to alter this”), he has no option but to plump for one alternative. Here is how Blackburn contrasts plumping with choosing (and also with reasoning):

I say “plump” deliberately, because saying that you have to choose carries a bad implicature. Choice is a process that invokes reasons. But the reasoning is all in *before* the case is describable as a stable agent’s quandary.¹⁴ It is because the reasoning leaves no ranking of alternatives, and because this is seen to be irremediable, that there is nothing left to do but plump.¹⁵

So the worry is that in cases of stable quandary, rationalism is unable to propose any alternative to plumping, thus effectively admitting defeat. But is that really so? I believe that the answer is no, but before arguing for it, let me bring into focus another, crucial distinction, drawn from St. Thomas Aquinas, the distinction between moral conflicts *simpliciter* and moral conflicts *secundum quid*. Donagan describes the issue in question in the following informative way:

A moral system allows perplexity (or conflict of duties) *simpliciter* if and only if situations to which it applies are possible, in which somebody would find himself able to obey one of its precepts only if he violated another, even though he had up to then obeyed all of them. For reasons already given, Aquinas held that any moral system that allows perplexity *simpliciter* must be inconsistent. By contrast, a system allows perplexity (or conflict of duties) *secundum quid* if and only if situations to which it applies are possible in which, as a result of violating one or more of its precepts, somebody would find that there is a precept he can obey only if he violates another.¹⁶

The above distinction highlights the importance of retrospective evaluation of our moral lives. This, in turn, suggests a new perspective from which to analyze the apparently irresolvable moral conflicts. So perhaps Blackburn is right in claiming that in stable quandaries reason has little or nothing to offer in prospective terms, but it is plausible that it has much to offer in retrospective terms. In other words, it should enable the agent to

¹⁴ The wording here makes it somewhat unclear as to whether “stable” qualifies the agent or the quandary. However, the broader context of Blackburn’s discussion of the problem at hand makes it clear that “stable” has to qualify the quandary despite the imprecise way he states it.

¹⁵ Blackburn, “Dilemmas: Dithering, Plumping, and Grief,” p. 129.

¹⁶ Donagan, “Consistency in Rationalist Moral Systems,” p. 306.

analyze carefully and minutely the way by which he apparently ended up with two or more conflicting obligations—and having performed the said analysis, perhaps he will find out that he is actually facing a *secundum quid* conflict.

The crucial point here is that the recognition of *secundum quid* conflicts may require quite subtle reasoning. When an irresolvable moral dilemma seemingly arises through no guilt of my own (e.g., I had made several promises to attend various appointments, but my beloved aunt suddenly got sick and I feel that I have a duty to visit and comfort her, and I cannot do all of these things due to time constraints), I need to ask myself whether I have really not made a moral mistake, even though presumably it was a slight one (and hence whether I am not facing a *secundum quid* rather than a *simpliciter* conflict). Perhaps my guilt is very subtle. It might, for instance, consist in violating a general rule of temperance, saying that one should be careful with making promises and avoid burdening oneself with too many of them, knowing that one might not find enough time to fulfill them all.

Such rules, whose violation can result in an irresolvable moral conflict, can be subsumed under a general, second-order, meta-prescriptive principle of the form “One ought to act in such a way that, if one ought to do x and one ought to do y , then one can do both x and y ,”¹⁷ and incorporated as such into a viable and consistent rationalist moral system. The recognition of the importance of observing this principle highlights the retroactive role that reason has to play in an agent’s stable quandaries, where the alternatives and values to be chosen from are genuinely incommensurable.

But can we trace *every* emergence of authentic value incommensurability to our previous infractions of the meta-prescriptive principle mentioned above? It seems that if at least one of the scenarios of the sort adduced by Blackburn turns out to involve a conflict *simpliciter*, then rationalism can be accused of incompleteness after all.

I believe that at this point we have to face a clash of intuitions with regard to what we expect of rationalism and where we see its limits. On the one hand, I do not think it unreasonable to suppose that those who claim that an acceptable rationalist moral system must be capable of providing one with a solution to absolutely every moral conflict could, in principle, find some logically defensible reasons to make a specific, conclusive choice in every scenario involving the apparently incommensurable alternatives, only that such reasons might plausibly seem too far-fetched. In the case of Sartre’s student, for instance, one could claim that the student’s familial duties definitely override his patriotic duties, since his coming into existence was not conditional on the existence of independent France (or any state at all), but it was conditional on the existence of his parents, and hence he owes more to his family (*ceteris paribus*, i.e., assuming that it was not abusive, etc.).

¹⁷ Ruth Barcan Marcus, “Moral Dilemmas and Consistency,” *Journal of Philosophy* 77 (1980), p. 135.

On the other hand, as I mentioned above, such an extension of rationalism could appear unconvincing and redundant. I personally tend to agree with such reservations and think that a viable form of a rationalist moral system should acknowledge that in the situations where reason favors neither of the available alternatives, nor exposes a *secundum quid* conflict, a disjunctive solution is an acceptable option. In other words, in such situations an agent should be allowed to plump. And yet, I do not believe that the system in question should be denigrated as a “watered-down” form of rationalism. On the contrary, I consider it to be a more comprehensive, multi-level rationalism, conscious of the distinction between first-order reasons for choosing a specific outcome among those available, and higher-order reasons for choosing either arbitrarily or on the basis of non-moral evaluations (e.g., aesthetic, pragmatic, etc.) where there are no overriding first-order reasons to prefer any particular outcome.

In this connection, let me spell out more fully what I mean by “disjunctive solution.” What I have in mind is a situation in which: (1) the competing courses of action, that is, the available disjuncts, can be brought under the common denominator of a shared area of concern (e.g., morality, aesthetics, efficiency, etc.)¹⁸; and (2) neither of the disjuncts under consideration can be said to promote the value(s) characteristic of the area in question more or less than the other. What follows from the above is that there is no reason to prefer either course of action, but there is also reason not to try to shun making a choice among them, since that would result in actualizing neither of the valuable outcomes attainable in the circumstances at hand.

I think this also addresses the possible criticism that appealing to “higher-order” principles indicates that rationalism cannot handle value conflicts on its own. Such higher-order principles are, in fact, part of any well-formed rationalist moral system. All that their “higher-orderedness” means is the recognition that as soon as it is found that reason (morally) favors neither of the alternatives at one’s disposal and does not reveal a *secundum quid* perplexity, one can justifiably have recourse to extra-moral reasons in making one’s final decision. In this sense, rationalist moral systems are no different from, say, rationalist economic systems, where, as soon as it is found that there is no reason to prefer either of the alternatives at one’s disposal on the grounds of economic efficiency, it becomes justifiable to tip the balance of one’s choice by appealing to extra-economic grounds.

The above contentions should become even more justified if we compare cases involving value incommensurability with (presumably) more familiar cases of option equality. It does not seem that finding oneself in a situation

¹⁸ If this criterion cannot be met, it indicates that a given pair of ostensibly competing choices cannot be said to figure in any meaningful choice situation (or, to put it proverbially, that the choices in question are as different as chalk and cheese, i.e., too different to be compared in any sensible way). Hence, it is a mistake to think that there is any relevant choice to be made in such situations.

where reason is indifferent to the range of options available indicates any deficiency of rationality on the part of the agent (at least in a given context). After all, such indifference is part of what defines equality, and it seems too much to claim that the notion of genuinely effective reason is incompatible with the recognition of equality among certain options.¹⁹ Consequently, it appears that a consistent moral rationalist can admit that there can be situations where he is confronted with a set of values with regard to which reason remains indifferent, but whose joint actualization he still (subjectively) desires.

To sum up: the criterion of consistency that we looked at assumes that a consistent moral system, if properly followed, cannot allow for the occurrence of genuine, irresolvable moral conflicts in the actual world. Having analyzed it in some detail, can we say that it does justice to the possibilities of any viable form of rationalism? Can it accommodate cases of *bona fide* value incommensurability? In one sense it certainly can, namely, by resorting to the disjunctive solution; it is, however, very important not to equate this step with suggesting the relativist's "anything goes" solution. The rationalist, unlike the proponent of the "anything goes" approach, engages in the process of moral reasoning, aimed at confirming or disconfirming whether the situation he is faced with is one of authentic incommensurability, rather than an instance of a *secundum quid* conflict or a scenario similar to that from *Sophie's Choice*. Thus, the eventual application of the disjunctive solution requires ruling out many other, alternative diagnoses of a given predicament, and that is where reason plays an indispensable role. However, having announced at the outset that there is another criterion of consistency that I find more defensible with respect to the assessment of a given moral system, let me now conclude the present discussion and turn to its analysis.

3. The Second Criterion

The second criterion of consistency that I shall focus on (hereafter called "the second criterion") derives from the meta-prescriptive principle introduced in the previous section in order to help recognize and avoid *secundum quid* conflicts. In contrast with this principle, however, it is extended to account not only for the quandaries brought about by previous violation of some moral precept, but also for those brought about by contingent factors outside of the agent's control. This extension is achieved by shifting the viability of the rule "ought implies can" from the actual world to at least one possible world. Ruth Barcan Marcus provides the following formulation of the criterion in question:

¹⁹ For more on the relation between equality and incommensurability, see, for instance, Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), pp. 430-31; James Griffin, *Well-Being: Its Meaning, Measurement, and Moral Importance* (Oxford: Clarendon Press, 1986), pp. 80-81; John Broome, "Is Incommensurability Vagueness?" in *Incommensurability, Incomparability, and Practical Reasoning*, ed. Ruth Chang (Cambridge, MA: Harvard University Press, 1997), pp. 67-89.

One ought to act in such a way that, if one ought to do x and one ought to do y , then one can do both x and y . But the second-order principle is regulative. This second-order 'ought' does *not* imply 'can'. There is no reason to suppose, this being the actual world, that we can, individually or collectively, however holy our wills or rational our strategies, succeed in wholly avoiding . . . conflict [of obligations]. It is not merely failure of will, or failure of reason [that produces such conflict]. It is the contingencies of this world.²⁰

Further caveats with regard to my understanding of the principle "ought implies can" may be in place here.²¹ Certainly, I do not believe that refocusing this principle on possible worlds is appropriate when it comes to non-conflict cases. For instance, it is not the case that I ought to save a drowning person if I cannot swim (as trying to do so would only result in my drowning and nobody being saved), even though there is a possible world in which I can. But in non-conflict cases the question of the consistency of a given moral system does not arise at all, so there is no need to invoke the second-order "ought need not imply can" principle mentioned above. My contention is that it should be reserved for dealing with dilemmatic scenarios. In sum, I think that next to the standard, first-order precept "ought implies can," which applies to individual moral injunctions in the actual world, we should recognize the second-order, regulative precept which holds that with regard to any consistent moral system, "ought" needs to distribute over a conjunction in at least one possible world. This last point might signify a departure from Marcus's view, who seems to hold that ought does not distribute over a conjunction in any possible world. However, I believe that it is a plausible departure, since it appears reasonable to me to claim that if one ought to do A, and one ought to do B, and A and B do not conflict with one another, then one ought to do both A and B.

Another criticism one could raise here is that Marcus's point is purely logical and has nothing to do with what we can reasonably expect of agents in the actual world. I agree with the first part of the preceding statement, but not with the second. It is true that Marcus's point is theoretical in nature, but since it is offered as a recipe for tackling the issue of moral dilemmas, I do not believe that we can think of it as divorced from practical considerations. On the one hand, it provides relief, but on the other hand, it warns against resting on one's laurels. That is, it firmly denies that morality requires the near-impossible, but it also affirms that it requires as much as possible. To expound and reiterate, the second criterion says that even if a strict adherent of a given moral system finds himself in an irresolvable moral dilemma, he should not judge the system in question as inconsistent. This is because the actual world

²⁰ Marcus, "Moral Dilemmas and Consistency," p. 135.

²¹ As was pointed out to me by an anonymous referee.

is full of extraneous influences, which are capable of landing even the most careful agent in situations of inextricable conflict. Bearing that in mind, one should be content with the thought that the moral system one follows is able to resolve every moral dilemma in at least one possible world. However, in order to validate the claim of one's authentic adherence to a given set of principles, one should also strive to make the actual world as close as possible to the ideal possible world mentioned above.

Let us consider one concrete example that might be helpful in elucidating the difference between the diagnoses offered by the meta-prescriptive principle incorporated into the first criterion and its extended version, developed into the second criterion. I start a business with a partner, leaving half of the managerial duties in his hands and taking the rest into my own. Later, I make some commitments to my family or friends, but then I find out that my partner has made a significant blunder and I need to rush to repair it if our business is to survive. Again, the quandary of having to decide between the values of my professional and personal life *might* turn out to be a *secundum quid* conflict, since I might have violated a rule which warns against entrusting morally consequential decisions to those whom we may expect to lack the necessary competences (without at least ensuring ourselves against the potential risk beforehand).

But the same dilemma might have arisen even if I knew that my partner is the most competent man in the world. The contingencies of reality are simply too numerous and pervasive even for the most prudent and industrious person to overcome on every occasion. Thus, even for the most prudent and industrious person "ought" need not always imply "can," although his guiding principle should be to live in such a way as to preserve the above implication as often as possible.

Further illuminating observations on the rule "ought implies can" are presented by Roger Trigg.²² He finds it difficult to agree with the claim made by authors, such as R. M. Hare, according to which an important element of our moral development consists in turning informal rules of thumb into concrete, precise precepts "with their exceptions definitely laid down."²³ Such a view threatens with the assumption that every moral principle must be regimented with a *ceteris paribus* clause. As Trigg writes:

However precise the rule, it always seems possible to be able to invent a situation, however unlikely, where it looks as if we ought to break it. There are very few actions which could not be justified if the fate of the world depended on what we did. If Hare is right, this means that we should not in such situations think of ourselves as

²² Roger Trigg, "Moral Conflict," *Mind* 80 (1971), pp. 41-55.

²³ R. M. Hare, *The Language of Morals* (Oxford: Oxford University Press, 1952), p. 51.

breaking the rule. We rather modify it, so that it does not apply to that particular type of combination of circumstances.²⁴

Why should the incorporation of *ceteris paribus* clauses into particular moral precepts be considered a threat (and a threat to what)? It may be true that their introduction can preserve the classical or “actualist” interpretation of the “ought implies can” rule,²⁵ but it also undermines the very need for having clearly formulated moral principles. The point is that reality often confronts us with situations where others things are not equal, and that, given the views of people like Hare, invites the practice of inventing exceptions in every more or less troublesome context, starting with the former and gradually including more and more of the latter.

Consequently, a moral principle reduces to a summary of all the decisions made in the past in specific circumstances, which could help one in being “consistent in the future if exactly similar situations arise, but as situations very often are not exactly similar even in morally relevant ways we still have to make up our minds without any rule to guide us.”²⁶ Relying on such summaries can be acceptable for an act-utilitarian, who subjects all of his actions to the overarching principle of maximizing the impartially and quantitatively understood utility (together with all of the numerous problems surrounding its formulation and application, which we need not discuss), but not to someone who, like Hare, sincerely employs rule-based moral language, thereby committing himself to genuinely rule-based moral reasoning.

Let us recall that the crucial insight offered by the second criterion is that an occurrence of a conflict of rules should not immediately be interpreted as an indication that the normative system to which they belong is logically inconsistent. This interpretation might turn out to be correct if the rules in question prescribe doing two mutually incompatible things for theoretical reasons, that is, reasons having to do with their logical structure—for instance, if one of them is an absolute injunction to tell the truth and another is an absolute injunction not to denounce innocent people. If, however, they prescribe doing two mutually incompatible things due, broadly speaking, to problems of practical implementation, then the resulting inconsistency is most probably not a logical defect of the normative system to which the agent adheres, but an inevitable effect of the contingencies of the world. Again, according to the second criterion, for a given set of principles to be consistent, there needs to be at least one possible world in which the application of that set of principles does not yield any dilemmas. A natural further extension of

²⁴ Trigg, “Moral Conflict,” p. 41.

²⁵ By the classical or “actualist” interpretation of the “ought implies can” rule, I mean the one according to which “ought” always implies “can” in the actual world.

²⁶ Trigg, “Moral Conflict,” p. 43.

this line of reasoning is that one should aim at making the actual world as close as possible to that ideal possible world. Hence, the rules comprising the moral system of one's choice should be treated seriously, that is, kept fixed and not modifiable in the face of contingently (though irresolvably) dilemmatic situations.

Drawing once more on Foot's distinction introduced in the previous section, we can say that even though one needs to resort to a type 2 ought statement (and thus break one of the conflicting precepts) in order to overcome a dilemma, the two jointly unsatisfiable type 1 ought statements (each of which corresponds to one of the precepts in question) do not lose their force. What is the significance of their retaining this force and how does this phenomenon manifest itself? It seems that the main issue here is the importance of keeping the moral institutions constituted by the observance of the relevant rules strong and stable. Suppose, for instance, that one admits that, for the sake of getting out of an otherwise irresolvable quandary, he broke the rule of, say, promise-keeping. If such an individual realizes that to the extent that something bad happened and some sort of compensation or restitution has to be made, he can be said to respect, understand, and recognize the binding force of the institution of promise-keeping. This attitude is what crucially distinguishes the agent in question from someone who nonchalantly breaks his promises due to being totally unconcerned with acting morally.

The compensatory behavior referred to above might be thought of as prompted by feelings such as hesitation, reluctance, or disgust at the moment of breaking a given principle and regret, remorse, or guilt after breaking it, but such a picture appears to be laden with emotional overtones not at all consonant with rationalism. It is better to interpret the need to make redress for the wrongs done as indicative of rule internalization, that is, the process of embedding and solidifying one's own attitudes, beliefs, and values, which results in making them genuinely integral to one's moral behavior. This, in turn, serves to ensure that they are given serious attention whenever the agent engages in the process of moral reasoning and that they play a prominent, action-guiding role in his ultimate decisions.

Despite its merits, however, some authors propose alternatives to the above story, presumably in an attempt to save the principle "ought implies can." Hillel Steiner, for instance, argues in a rejoinder to Trigg that in the apparently problematic scenarios described by the latter, the agent should be thought of neither as breaking nor as modifying a given moral rule. Instead, he should be thought of as creating a new rule, since a dilemmatic situation confronts him with a new circumstance, and neither of the putatively conflicting rules available beforehand was formulated with the intention of dealing with such a circumstance. As Steiner writes:

[A] situation in which factual conditions corresponding to both C and C2 exist, is one in which our individual actually confronts a circumstance—namely, C + C2—which is different from either C or

C2. . . . Therefore, moral rules respectively covering what ought to be done when C or when C2, do not apply to this situation. What applies to this situation is a rule covering what ought to be done when C + C2. And this is a different moral rule, enjoying the same logical status as the other two.²⁷

But this proposal seems to parallel Hare's view and share all of its objectionable features. Harean prescriptivism conceives of moral development as a process whereby commonplace action-guiding practices are turned into formalized principles, each with its scope of application precisely spelled out. Hence, for instance, it enjoins the observance of a given rule X, but only if none of the exceptional circumstances C1, C2, C3 . . . CN obtains. Steiner's suggestion, on the other hand, requires one to include in the description of a given rule all specific, morally relevant circumstances constitutive of the range of situations that the rule in question is supposed to cover. (Consequently, one is supposed to end up with a set of injunctions of the form "apply the rule X only if all the circumstances C1, C2, C3 . . . CN obtain.") It is not difficult to notice that the two views under discussion are mirror images of one another.

Just as Hare's position threatens to impose on us an insurmountable task of finding and listing infinitely many possible circumstances which override the applicability of a given principle, Steiner's position requires us to accomplish an equally insurmountable task of specifying and listing infinitely many possible factors that might bear some moral relevance to a given principle. Potential combinations of such factors are endless, and Steiner seems to acknowledge this fact, but instead of recognizing it as a difficulty for his proposal, he enjoins us to grasp the infinite and extract the finite out of it (presuming, of course, that the rules of a normative system one adheres to should have finite descriptions):

Indeed, any one factual statement may be partially descriptive of a wide range of different circumstances, covered by a correspondingly diverse range of moral rules. Consequently, in order to know what sort of rule applies in a particular situation, it is vital to ascertain *all* the morally relevant facts about that situation.²⁸

I take it that normative rules are supposed to work prospectively, that is, the point of having them is to know how to act in future situations. They cannot fulfill this role, however, if we are to invent a new rule for every new set of circumstances that we encounter. It appears implausible to suppose that such sets of circumstances, in exactly the same form, will occur very often. If,

²⁷ Hillel Steiner, "Moral Conflict and Prescriptivism," *Mind* 82 (1973), p. 587.

²⁸ *Ibid.*, p. 588.

on the other hand, we wish to avert the above problem by specifying all of the rules of a given normative system in advance, while at the same time respecting the requirements of Steiner's proposal, then we end up with a practical impossibility. As I have already mentioned, the number of possible combinations of circumstances that one can consider is infinite, and it is impossible to adhere to a system containing an infinity of rules (or a system whose rules have infinite descriptions).

In sum, the attempts of authors such as Hare and Steiner to save the universal applicability of the principle "ought implies can" appear to lead to counterintuitive and overambitious conclusions with regard to the nature of rational moral thinking. The second criterion, on the contrary, makes very good sense of the claim that "ought" need not always imply "can" (though we should strive to ensure that this implication holds as often as possible) and that rules can sometimes be broken. It need not indicate that the system to which they belong is logically inconsistent, but it always points toward the sensible conclusion that the contingencies of this world may overcome even the most sophisticated moral logic.

4. Conclusion

I have attempted to show that the occurrence of moral dilemmas in the course of one's life need not imply that the moral system one follows is logically inconsistent, regardless of whether assessed against the first or the second criterion. In relation to both of these criteria, the emergence of dilemmas may play the important role of suggesting what corrective measures should be taken in order to avoid similar predicaments in the future.

In the case of adopting the first criterion, reflecting on an apparently dilemmatic situation can reveal a *secundum quid* conflict, an instance of misidentifying a practical conflict as a moral conflict or an instance of genuine value incommensurability. Adopting the second criterion, on the other hand, in principle enables one to make each of the above discoveries as well, but it also allows one to trace the origins of some of the irresolvable moral conflicts to the unconquerable contingencies of the actual world. This is, in my opinion, the more plausible approach to take. None of these potential findings implies an inconsistency in the set of moral principles one adheres to, but each of them offers some clues as to how to restructure one's relations with reality so that the risk of running into dilemmas is minimized.

