

16. I am grateful to those who discussed an earlier draft of this paper with me at the 1999 Dartmouth College conference on Gert's Moral Theory. I am especially grateful to Daniel Bush, Bernard Gert, Stephen Nathanson, Geoffrey Sayre-McCord, Walter Sinnott-Armstrong, and Michael Smith.

6

Bernard Gert's Complex Hybrid Conception of Rationality*

Michael Smith

In the early chapters of his comprehensive, original, and startlingly systematic *Morality: Its Nature and Justification*, Bernard Gert argues that our concept of rationality—or, more accurately, our concept of practical rationality—has a hitherto unrecognized “complex hybrid” character.

It is the failure to recognize the hybrid character of rationality that is responsible for the inadequacy of all the previous accounts of that concept. . . . It is only by recognizing the hybrid character of rationality, acknowledging that it has both an egocentric part (irrational actions) and a nonegocentric component (reasons), that an adequate definition can be formulated. (83)¹

Despite Gert's various ingenious arguments for this claim, and the many lessons about the concept of rationality that he teaches us while arguing for it, I remain skeptical. Not only does an alternative theory of rationality, a version of what he calls the “cool moment” desire theory (42-44), seem to me invulnerable to his main objection, a version of that theory also, in my view, provides us with a far more plausible conception of practical rationality than the complex hybrid conception Gert recommends. So, at any rate, I wish to argue here.

The Commonsense Conception

Like all philosophical theories, theories of practical rationality must ultimately answer to the court of commonsense. Let me therefore begin by briefly explaining what I take the commonsense conception to be.

As I understand it, the commonsense conception of practical rationality has the concept of a reason for action at its very core. There are various reasons for action, where reasons for action are thought of as facts about which agents can gain knowledge via reflection. Moreover, some of these reasons are egocentric in character, whereas others are nonegocentric. The reasons we each have to avoid harm to ourselves, to help our friends, and the like, are in the former category—a specification of their content requires an ineliminable reference to one-

self—and the reasons we each have to prevent the suffering of sentient creatures quite generally are in the latter.

Reasons, in turn, whether egocentric or nonegocentric, have different weights. Sometimes egocentric reasons outweigh nonegocentric reasons (the reason to provide a benefit of a certain magnitude to oneself, or a member of one's immediate family, or a friend, is weightier than the reason to provide that same level of benefit to a complete stranger); sometimes nonegocentric reasons outweigh egocentric reasons (the reason to prevent misery to a million strangers is weightier than the reason to prevent that same level of misery to oneself, or a member of one's immediate family, or a friend); and sometimes egocentric and nonegocentric reasons are of much the same weight (perhaps the reason to provide a benefit of a certain magnitude to oneself, or a member of one's immediate family, or a friend, has much the same weight as the reason to provide that same level of benefit to two [or maybe it is three or four. . .] complete strangers).

These various reasons with their associated weights are, in turn, capable of explaining what agents do, at least in those favorable cases in which agents know about their existence and give them the weight that they in fact have. But knowledge of the reasons that there are is sometimes impossible to gain. In determining the rationality of an agent's actions, the commonsense conception of practical rationality therefore insists that we ask not whether agents act as they have reason to, but rather whether they act (or whether it is as if they act) as they believe that they have reason to, and whether these beliefs about the reasons that there are, in turn, elude various ordinary forms of rational criticism to which all beliefs are subject.

When these conditions are met—and when various other familiar conditions on intentional action are met as well—the commonsense conception tells us that the actions in question are rational, and when they aren't, it tells us that we have a case of irrational action. Agents may therefore act rationally without acting as they have reason to, and they may act as they have reason to without acting rationally. What is important to determining the rationality of agents' actions is rather whether they act as they believe that they have reason to, and whether these beliefs are in turn well-grounded.²

My claim in what follows is to be that a version of the cool moment desire theory is preferable to Gert's complex hybrid conception of rationality *inter alia* because that theory better squares with this commonsense conception. Of course, if I am wrong that the conception of practical rationality just described is commonsensical, then my objections to Gert's complex hybrid conception will simply fail. But here is not the place to argue about that. Rather, assuming that the conception just described is commonsensical, I want to focus instead on which theory better squares with it.

Outline of Gert's Complex Hybrid Conception of Rationality

Here is Gert's full-dress definition of rational action.

An act is irrational if and only if it is an intentional action of a person with sufficient knowledge and intelligence to be fully informed about that action and who, if fully informed, (1) would believe that the action involved significantly increased risk of his suffering death, nontrivial pain, loss of ability, loss of freedom, or loss of pleasure and (2) would not have an adequate reason for the action. A reason for acting is a conscious rational belief that one's action will increase the probability of someone's avoiding

any of the harms listed above or gaining greater consciousness, ability, freedom, or pleasure. A reason is adequate if any significant group of moral agents regard the harm or benefit gained as compensating for the harm suffered. Any intentional action that is not irrational is rational. (83-84)

Let me bring out some of the crucial features and implications of this definition.

To begin, the definition assumes, rightly I think, that the only people who can act irrationally are those who have the capacity to be fully informed about their actions.³ Those who are incapable of forming beliefs about the chances that their actions will have the features that they have—including that they cause harms such as death, pain, loss of ability, loss of freedom, and loss of pleasure, or that they cause compensating benefits in terms of someone or other's avoiding these harms or gaining increases in consciousness, ability, freedom, or pleasure—cannot act irrationally. But those who are capable of forming such beliefs may well act irrationally even when they act in ignorance of relevant considerations. Specifically, they act irrationally when they could and should have known better. I take it that this is why Gert couches his definition in counterfactual terms: "if fully informed, would. . ."

Second, Gert's definition entails that it is always irrational for an agent to act so as to knowingly harm himself—that is, knowingly to cause himself pain or death or loss of ability or freedom or pleasure—without adequate reason. Thus, according to Gert's definition, people are one and all rationally required not to harm themselves without adequate reason. The concept of rational action is therefore substantive and egocentric: it gives special significance to the effects of an agent's own actions on himself. Here, in Gert's view, we see the grain of truth that lies in egoism.

Third, the definition entails that it may nonetheless be rational to act so as to knowingly harm oneself. For someone who knowingly harms himself in the presence of a conscious rational belief that his action will bring compensating benefits for himself or someone else thereby acts with adequate reason, and so does not act irrationally. Moreover, the strength of the reasons thus provided does not depend on whether the benefits are for the agent himself or for someone else. It is here, according to Gert, that we find the grain of truth that lies in Kantianism. Some reasons for action are nonegocentric, and even in those cases in which there are egocentric reasons, the strength of the egocentric reason is the same as the strength of the corresponding nonegocentric reason. It is thus rationally allowed for people to harm themselves in order to provide benefits either for themselves or for other people. Egoism goes wrong in so far as it denies this. To this extent Gert's definition is both substantive and nonegocentric.

Fourth, and finally, the definition entails that it may, however, nonetheless be rational for someone to act so as to knowingly harm someone else, despite the fact that he has adequate or even stronger reason not to do so. For so long as the action does not harm the agent himself, the definition tells us that the agent does not act irrationally, notwithstanding that he has reasons for acting otherwise. In Gert's view Kantianism goes wrong in so far as it denies this. This is the sense in which Gert's theory is hybrid in character. Contrary to Kantian theories, Gert thus thinks that it is always rationally allowed for people to harm others when doing so will not harm themselves. Here again we see the egocentric character of the concept of rational action. Egoism is once again on the right track.

Critical Discussion of Gert's Conception of Practical Rationality

Gert on Egocentric and Nonegocentric Reasons

According to Gert's theory, practical rationality is a hybrid concept, having "both an egocentric part (irrational actions) and a nonegocentric component (reasons)" (83). Moreover, as I emphasized in drawing out the second and third implications of his theory, he claims that the strength of an egocentric reason is identical to the strength of the corresponding nonegocentric reason. Here, accordingly, we find the first difference between Gert's conception of reasons and the commonsense conception.

As I said at the outset, while the commonsense conception agrees with Gert that there are both egocentric and nonegocentric reasons, it differs from Gert's conception in holding that these reasons may differ in their relative strengths. It holds that though egocentric reasons sometimes outweigh nonegocentric reasons, nonegocentric reasons may sometimes outweigh egocentric reasons, and these reasons may even sometimes be of the same strength. According to Gert, however, this commonsense view rests on a confusion.

The question whether reasons of self-interest are stronger than reasons involving the interests of others arises only if one confuses reasons with motives. If the amount of harm to be avoided and the benefit to be gained is the same, the reasons involving self-interest cannot make rational any acts that reasons involving the interests of others cannot make rational. Any irrational act that would be made rational by a reason of self-interest would also be made rational by a reason of the same strength involving the interests of others. A mere change of person affected does not affect the strength of a reason. (78)

In other words, those who claim that egocentric and nonegocentric reasons may differ in their relative strengths are misled by the fact that people are sometimes *moved* by the fact that they will suffer a certain amount of harm themselves and left *unmoved* by the fact that others suffer that same amount of harm. But, Gert insists, while this is true it doesn't show that the strength of the reason isn't the same in both cases. It is the same, according to Gert, because a "mere change of person affected does not affect the strength of a reason."

The main problem with this diagnosis, however, is not just that it is at odds with commonsense, but that it is also at odds with Gert's own stated criterion for the way in which the weight of the reasons that there are gets fixed. "A reason is adequate," Gert tells us, "if any significant group of moral agents regard the harm or benefit gained as compensating for the harm suffered" (84). But in that case if, as seems likely given the commonsense view, some significant group of moral agents would each give more weight to their making it less likely that they will suffer greater harm themselves than they would give to their making it less likely that just someone will suffer greater harm, then it surely follows that, even by Gert's own criterion, the reasons have different weights. Gert's claim that a "mere change of person affected does not affect the strength of a reason" thus looks to be false even by his own lights.

The upshot is thus that we can reduce at least some of the distance between Gert's conception of practical rationality and the commonsense conception. On Gert's conception, we should suppose that reasons for action can be both ego-

centric and nonegocentric and that the weight of these reasons can differ. However we cannot reduce all the distance between Gert's conception of practical rationality and the commonsense conception, for according to the commonsense conception, reasons are facts, whereas according to Gert's conception, reasons are conscious rational beliefs. Let's therefore turn our attention to that aspect of his theory.

Gert on Reasons as Beliefs rather than Facts

Gert gives the following argument in favor of the view that reasons are conscious rational beliefs.

If facts were reasons, it would weaken the conceptual connection between reasons, the rationality of actions, and the rationality of persons. Many facts are such that not only is the agent unaware of them, no one could be expected to be aware of them. These facts do not affect the rationality of an action when the rationality of an action counts in determining a person's rationality. (65)

He then gives an example that, in his view, illustrates the difficulty.

Imagine someone who acts in a way that she believes will cause her death—she takes an overdose of pills, say—but without believing that her doing so will be compensated for by the harms she prevents or the benefits she produces. However, suppose further that the pills she takes are the antidote to a poison that she unknowingly ingested at some earlier time, so that her act of trying to kill herself, far from causing her death, in fact saves her life. Gert points out that on the view of reasons as facts, this woman might well have acted as she had reason to. But, he insists, it would be extremely misleading to conclude, on this account, that she acted rationally. On the view of reasons as conscious rational beliefs, by contrast, he tells us that there is no reason to suppose that she acted rationally. Indeed, on his own account of irrational action, since the woman had no conscious rational belief that serves as an adequate reason for committing suicide, he claims that it follows that she acted irrationally. This, he tells us, is the right result.

Let me enter a parenthetical remark here. Though Gert claims that his theory would classify this woman's act as irrational rather than rational, it is not at all plain to me that it would do so. For remember, according to Gert's definition, the irrationality of an act does not depend on whether the person who acts *in fact* believes that she is acting in a way that will harm herself, and, if she does, whether she does not *in fact* have a conscious rational belief that provides her with an adequate reason for so acting. Rather it depends on whether she *would* have this constellation of belief and absence of belief if she were fully informed.⁴ This was the point I emphasized when I spelled out the first implication of Gert's theory.

Suppose, then, that we add the following details to Gert's story. Though the woman he describes is in fact ignorant, both the fact that she has ingested poison and the fact that the pills that she is about to take are an antidote to that poison are suitably available to her. That is, let's suppose that she would, if fully informed, have both of these beliefs. Then, notwithstanding her actual ignorance, and notwithstanding that she actually acts in the belief that she is committing suicide by taking the pills she takes and that her doing so is not in any way compensated for by harms that she thus avoids or benefits that she thus

produces, since, if she were fully informed, she would believe that her action of taking the pills benefits her rather than harms her, it follows that she does not act irrationally. And since acts that are not irrational are rational, so, according to Gert's definition, it would seem to follow that the woman acts rationally, not irrationally. Yet if, as it seems to me, the woman in this more detailed version of Gert's example plainly acts irrationally, then the fact she acts rationally according to Gert's own definition must surely count as an objection to it. Here ends the parenthetical remark.

I cannot see why Gert thinks that his objection to the claim that reasons are facts is so telling. Suppose we agree with him that the only facts that can affect the rationality or irrationality of an action are those of which an agent could have been aware, and suppose that we further agree with him that, since this is so, it follows that reasons cannot be grounded in facts of which agents cannot be aware. I am not positively endorsing the suggestion that we agree with him about this, I am merely suggesting that we imagine ourselves agreeing with him, for the sake of argument. Even with all this agreed it simply doesn't follow that reasons are conscious rational beliefs. All that follows is that, if reasons are facts, then they are facts of which agents could have been aware. But facts of which agents could have been aware are still facts. They are not conscious rational beliefs.

Worse still, it seems to me that Gert saddles those who think that reasons are facts with a view about the rationality of an agent's acts that they simply needn't hold. As I said at the outset, precisely because the commonsense conception holds that reasons are facts of which agents may be *unaware*, it holds that the rationality of an agent's acts is to be determined not by whether she acts in a way that she has reason to act, but rather by whether she acts in the belief that she has an adequate reason for so acting, where that belief eludes rational criticism. Thus, in Gert's example, those who favor the view that reasons are facts can insist that even though the woman did act as she had reason to, since she didn't act in the *belief* that she had that reason she did not, on that account, act rationally. Moreover, they can insist that if the belief about reasons on which she did act was rationally criticizable, then she acted irrationally. As far as I can see, this means that in both of the versions of Gert's example that we have been discussing here, those who hold that reasons are facts will say that the woman acts irrationally. This is the right result.

At one point, Gert seems to suggest that this view about the relationship between reasons and rationality is simply a terminological variation on his own theory of rational action.

Of course we might distinguish between external and internal reasons. External reasons are facts and are the kind of reason normally referred to when one says "There is a reason." Internal reasons are beliefs and are normally referred to by saying, "One has a reason." If there are external reasons to do something, then if one knows of these facts, one has internal reasons to do it. If one has internal reasons to do something, then one believes that there are external reasons to do it. External reasons determine the external rationality, sometimes misleadingly called the "objective rationality" of an action. Internal reasons determine the internal rationality, sometimes misleadingly called the "subjective rationality" of an action. In the ideal case, which I hope is the normal case, external and internal reasons coincide, so that the external and internal rationality of an action are the same. However, sometimes the external and internal reasons do not coincide. For the ra-

tionality of actions to have the appropriate relationship to the rationality of persons, internal rationality must be taken as the basic sense of rationality. Accepting internal rationality as the basic sense of rationality is regarding basic reasons as beliefs rather than facts. (65-66)

However it seems to me that this involves a serious misunderstanding.

The commonsense view, at least as I have described it—and, indeed, the view that Gert discusses in the passage just quoted—is one according to which reasons are themselves the objects of those of an agent's beliefs that determine his rationality. In other words, the propositional content of the beliefs that determine the rationality of agents is: that there is a reason for them to act in this way rather than that. The beliefs are thus about reasons, in a primary sense of the word "reason". They are not themselves reasons, in this primary sense. Of course, we can, if we wish, go on to define a secondary sense of the word "reason" to refer to beliefs that are about reasons in the primary sense. This is in effect what Gert imagines us doing in the passage just quoted. But, importantly, reasons in this secondary sense, even if we do choose to define such a sense, are beliefs that are about reasons in the primary sense.

This contrasts strikingly with Gert's own official view about reasons. Gert's official view is that reasons are beliefs which, when possessed by agents whose acts would otherwise be irrational, serve to make those actions rational. An example of such a belief, according to Gert, would be an agent's belief that cutting off his arm will save his life. For if we imagine this belief possessed by an agent who cuts off his poisoned arm to stop the poison spreading to the rest of his body, then, in Gert's view, the fact that the agent has this belief serves to make his cutting off his arm, an act which would otherwise be manifestly irrational, rational.

But we can now see the crucial difference between Gert's official view and the view of reasons described in the passage quoted above. The beliefs that Gert thinks reasons are are not the same as the beliefs that count as reasons in the secondary sense, for the beliefs that Gert thinks reasons are not beliefs about reasons, in a primary sense of the word "reason". Rather they are beliefs about the effects that acts have on the weal and woe of agents. Moreover, for this very reason, we can now see that Gert's theory is seriously problematic. For the acts of agents who act in the belief that their acts will have certain effects on the weal and woe of people, but who do not act in the belief that this sort of consideration is reason-giving, are surely no more rational merely for having been performed in the presence of the belief. Imagine, just to make the point vivid, that the agent believed that the consideration provided a reason *against* so acting. Would it be made rational then? I do not think so.

To have any effect at all upon the rationality of an agent's acts, an agent's belief that his act will have certain effects on the weal and woe of people must therefore be accompanied by the belief that that consideration is reason-giving. In other words, the agent must have a reason in the secondary sense described above, a belief about a reason in a primary sense of the word "reason". The idea that the beliefs that Gert thinks reasons are are capable of making otherwise irrational acts rational all by themselves thus seems to me to be in serious tension with our ordinary ways of thinking and talking about reasons. Despite his suggestion to the contrary, it therefore seems to me that there is no way that Gert can turn the commonsense view I have described into a version of his own theory by a feat of redefinition.

As further proof that this is so—in other words, as further proof that the view I have labeled “commonsense” really is commonsensical—imagine the parallel situation with reasons for belief, a parallel that Gert endorses (66–67). Suppose that someone says to me, “I know that you believe that *q*, but tell me what reasons there are for so believing.” If I find myself unable to give an answer to this question then it seems to me that my natural inclination would be to reflect, and perhaps to investigate the world in which we live, in an attempt to discover what reasons there might be for believing that *q*, where this inclination, in turn, would be premised on the idea that reasons are facts, as opposed to beliefs, facts which I can find out about by engaging in acts of reflection and investigation. For it is only if reasons are such facts that I can so much as make sense of the idea that there might be reasons for believing *q* of which I am currently unaware. Moreover and importantly, given that we are comparing our ordinary ways of thinking and talking about reasons with Gert’s, the inclination also seems to be premised on the idea that these facts are suitably independent of the acts of reflection and investigation via which I can come to have beliefs about them. There wouldn’t, after all, be much point in my reflecting or investigating in order to discover what reasons there are for believing that *q* if the acts of reflection and investigation themselves could *change* the facts about the reasons that there are for believing that *q*.

On Gert’s conception of reasons as conscious rational beliefs, however, none of this is true. For, on that conception, since I have no conscious rational belief that makes the belief that *q* rational, the proper answer to the question as to the reasons that there are for believing that *q* is that there aren’t any such reasons. Reflection and investigation are thus pointless. Worse still, on that conception it seems that reflection and investigation would simply change the facts about the reasons that there are for believing that *q*. For, on that conception, though there will be reasons to believe that *q* after I reflect and investigate—let’s assume that, as a result of reflection and investigation, I would come to believe both that *p* and that *p* entails *q*—the reasons that there will be for believing that *q* will be reasons that the acts of reflection and investigation themselves brought into existence. Engaging in reflection and investigation thus doesn’t enable me to detect the preexisting reasons that there are for believing that *q*, it simply enables me to create some reasons for believing that *q*. It is thus, in a more or less quite literal sense, up to me whether or not there are any reasons for believing that *q*.

None of this seems to me to square with our ordinary ways of thinking and talking about reasons. Our ordinary thought and talk about reasons is simply shot through with the assumption that reasons are facts, facts which preexist our acts of reflection and investigation but about which we can form beliefs about by engaging in acts of reflection and investigation. The only conclusion to draw would therefore seem to me to be that we must reject Gert’s conception of reasons as conscious rational beliefs and suppose, instead, that reasons are facts.

Gert’s Objection to “Cool Moment” Desire Theories

If reasons are facts, then we desperately need to say what sorts of facts they are. My own view is that reasons are facts whose status as reasons is conferred upon them by their relations to idealized psychological facts.⁵ Since this is similar to an idea that Gert considers and rejects—the idea that we can define rational action as that which satisfies the desires we would have in a cool moment—I want

to consider briefly whether my own view is vulnerable to Gert’s objection to cool moment desire theories.

Gert’s objection to cool moment desire theories of rationality is, in essence, that they deliver the wrong results. Consider, for example, people who have an untreated mental illness (43). They might well desire, in a cool moment, that they kill themselves, or cause themselves pain, or to lose abilities, or to lose freedom, or to lose pleasure, in circumstances in which their so doing will not be compensated for by the avoidance of greater harms or the production of benefits.⁶ But since, according to Gert, it could never be rational to do such a thing, it follows that cool moment desire theories classify as rational acts that are plainly irrational.⁷ People are rationally required not to kill themselves when doing so will not be compensated for by the avoidance of greater harms or the production of benefits, no matter whether they happen to desire to do so in a cool moment or not.

An initial response to Gert’s objection might be that it shows just that cool moment desire theories require augmentation. Those who accept such theories should say not that it is rational for an agent to do what he would desire himself to do, if he formed his desires in a cool moment, but rather that it is rational for him to do what he would desire himself to do, if he formed his desires in a cool moment, providing that when he forms those desires he is not mentally ill. According to Gert, however, this sort of amendment to a cool moment desire theory is tantamount to giving up the theory altogether. For what defines someone as having a mental illness is, at least in some cases, simply their possession of certain desires, desires which are irrational in a much more fundamental sense, a sense that eludes any illumination at all. Gert makes this point—that the irrationality of certain desires is basic, or fundamental—most forcefully in the following passage.

I regard this list of five basic irrational desires (the desires for death, pain, disability, loss of freedom, and loss of pleasure) as complete. I cannot prove that it is complete, indeed I cannot prove that any of the desires that I have included in this list are basic irrational desires. I have not derived this list of irrational desires from any more basic account of irrationality. . . . I cannot and would not try to provide arguments for any item on the list. (51)

In Gert’s view, then, no further account can therefore be given of the irrationality of desiring death, or pain, or disability, or loss of freedom, or loss of pleasure, in circumstances in which there is no compensation to be gained from the avoidance of greater harms or the production of benefits. The irrationality of these desires is the starting point from which all further conversation and argument about rationality begins. Their irrationality is a given.

Let me now spell out my own view about the nature of reasons in order to see whether it is vulnerable to a similar line of objection. My own view, to repeat, is that reasons are facts whose status as reasons is conferred upon them by their relations to idealized psychological facts. To be a little more precise, in my view the fact that an agent *A* can perform an action of a certain kind *K* in certain circumstances *C* by performing an act of kind *K** in those circumstances constitutes a reason for him to perform an act of kind *K** in *C* if and only if everyone, *A* included, would want that they themselves perform an act of kind *K* in *C* by performing an act of kind *K** in those circumstances if they were fully rational. If we call the possible world in which *A* has the inclinations and beliefs that he

has in circumstances *C* the "evaluated" world, and the possible world in which he is fully rational the "evaluating" world, then, the idea is, the facts that constitute *A*'s reasons for acting in the evaluated world have their status as reasons conferred upon them not by what *A*, in the *evaluated* world, wants himself to do in the evaluated world, but rather by what he, in the *evaluating* world, wants himself to do in the evaluated world.

Thus, to illustrate, as I see things the fact that, in certain circumstances *C*, an agent's cutting off his arm will save his life constitutes a reason for him to cut off his arm in those circumstances if and only if everyone, *A* included, in possible worlds in which they are fully rational, want that, in possible worlds in which they themselves are in circumstances *C*, they save their own lives by cutting off their own arms. The fact that constitutes the agent's reason to cut off his arm—the fact that he can save his life by so doing—thus has its status as a reason is conferred upon it by its relation to idealized psychological facts, facts about the desires of fully rational creatures.

Note that this account of reasons provides us with an intuitive and compelling account of the nature of reasons for and against acting in a certain way. This is because, with regard to a particular set of circumstances, a fully rational agent could have several conflicting desires about what is to be done. Thus, for example, an agent might have a reason to cut off his arm in certain circumstances *C* because, if he were fully rational, he would want himself to do what is required to save his life, and cutting off his arm is what is required. And he might also have a reason not to cut off his arm in those same circumstances because he will cause himself pain by doing so, and, if he were fully rational, he would want himself to do what is required to prevent himself from feeling pain. The agent thus has a reason for and against cutting off his arm, and what he has reason to do all things considered is therefore fixed by the relative strengths of such desires: that is, by facts about what his fully rational self would overall want, or most want, himself to do in the relevant circumstances.

The attraction of this sort of view, much as with the attraction of cool moment desire theories quite generally, should be plain. It is, after all, extremely plausible to suppose that what an agent has a reason to do is whatever someone perfectly placed to give the agent advice would advise him to do. But the person who is best placed to give an agent advice is simply himself, minus all of his imperfections and idiosyncrasies. It is the agent himself, in the possible world in which he is fully rational. In effect, the account of reasons just described simply makes this idea more precise. For the advice a fully rational agent would give is simply an expression of his overall preference among the various options that are available in the circumstances of action in the context of which he is giving advice.

At this point, Gert's objection might seem to loom large. For, he might say, the plausibility of the account of reasons just given is entirely dependent on the answer to a so far unanswered question. What makes an agent *fully rational*? It appears that we face a dilemma. On the first horn, we give an account of what it is to be fully rational that does not simply reduce to saying what Gert says. In other words, we give an account that does not amount to saying that to be fully rational is, *inter alia*, to lack desires for death, pain, disability, loss of freedom, and loss of pleasure in circumstances in which no compensation is to be gained from the avoidance of harms or the production of benefits. On this horn, Gert might say, the account of reasons is vulnerable to the same objection he made to cool moment desire theories. For such a theory is consistent with something that is manifestly false, namely, that agents can have reason to kill themselves

in circumstances in which no compensation is to be gained from the avoidance of harms or the production of benefits. On the other horn of the dilemma, we say exactly what Gert says. In other words, we say that to be fully rational an agent must, *inter alia*, lack desires for death, pain, disability, loss of freedom, and loss of pleasure in circumstances in which no compensation is to be gained from the avoidance of harms or the production of benefits. But once we say this we abandon the project of giving an account of reasons of the sort just described in favor of adopting a view like Gert's. The fundamental idea lying behind the concept of a reason is conveyed by Gert's list of five basic irrational desires.

My own view is that we should grasp the first horn of this dilemma. However, in order to see why the first horn isn't as problematic as Gert supposes, we must first think a little more carefully about Gert's own positive view, his view that irrational desires are defined by a *list* of the five desires already mentioned. What really drives Gert in the direction of this positive view? The answer, it seems to me, is that Gert is driven to the view that irrational desires are defined by a list because of the high degree of confidence he supposes us all to have that the desires for death, pain, disability, loss of freedom, and loss of pleasure are indeed irrational. Indeed, as the argument given on the first horn makes plain, it is as though Gert thinks that if we posit a feature that unifies desires under the labels "rational" and "irrational," then that very fact will somehow call into question, or have the potential to undermine, the confidence we have that the five desires on his list really are irrational. But, of course, that is just a mistake. The unifying feature we posit might underwrite our confidence, rather than undermine it.

Moreover, and more importantly, no matter how high our confidence that the desires for death, pain, disability, loss of freedom, and loss of pleasure are irrational, provided that our confidence isn't 100 percent, the bottom line is that we *do* thereby grant that it is at least possible, even if only barely, that that confidence could be undermined by further reflection. But in that case, the very considerations which strike us as potentially relevant in undermining our confidence look like they will themselves contain the seeds of an explanation of what unifies desires under the labels "rational" and "irrational." In other words, they will provide us with a conception of what it is to be fully rational that *isn't* just equivalent to having no desires on Gert's list of five irrational desires.

So, let's ask the crucial question. What would undermine our confidence that the desires for death, pain, disability, loss of freedom, and loss of pleasure are irrational? Our confidence would be undermined, it seems to me, if we thought that, starting from uncontroversial premises, we could provide a compelling argument that the desires for death, pain, disability, loss of freedom, or loss of pleasure were part of a systematically justified set of desires. If the premises and the steps in the argument, taken together, are something about which we could be more confident than we are that the desires for death, pain, disability, loss of freedom, and loss of pleasure are irrational, then it is our confidence that these desires are irrational that would be undermined, not the former, and rightly so.

What are the characteristics of a desire set that is systematically justified? The answer, it seems to me, is that a desire set that is systematically justified comprises elements that are immune to revision on grounds of ignorance or error about matters of fact; the elements of the desire set cohere well with each other; the desire set, as a whole, exhibits a kind of unity; and so on and so forth. The idea, in other words, is that a desire set is systematically justified to the extent that it exhibits analogues of the standard *epistemic* virtues. One standard epis-

temic virtue, of course, is empirical adequacy. A desire set that is systematically justified must therefore have an analogue of that as well.

This is where Gert's confidence that the desires for death, pain, disability, loss of freedom, and loss of pleasure are irrational gets assigned its proper role. After all, we do not begin the task of constructing a systematically justified desire set utterly devoid of substantive views about which desires will, and which desires will not, be elements of a desire set that is systematically justified. Thus, for example, to the extent that we think that it would make no sense whatsoever for someone to desire death, pain, disability, loss of freedom, or loss of pleasure in circumstances in which there is no compensation to be gained from the avoidance of harms or the production of benefits, we start out with the assumption that no desire set containing these desires *could* be systematically justified. But the important point is that this is a *defeasible assumption*. It can be defeated by further reflection.

Suppose that, starting out from desires which we provisionally admit are systematically justified, we could provide a compelling argument for the conclusion that the set of desires that we would have if we were to add and subtract desires in the light of the information and considerations of coherence and unity *does* contain one or another of the desires for death, pain, disability, loss of freedom, or loss of pleasure, in circumstances in which there is no compensation to be gained from the avoidance of harms or the production of benefits. If such an argument could be constructed, then it seems to me that we might well happily concede that these desires *can* be elements of a systematically justified set after all. Even though we started out thinking that no fully rational agent could have such desires, we would have convinced ourselves that we were wrong, that they could.⁹

Here, then, we have a way of negotiating the first horn of the dilemma. Notwithstanding the fact we characterize a fully rational agent *without* reference to Gert's list—in other words, notwithstanding the fact that we characterize a fully rational agent as someone whose desire set is systematically justified, where a desire set is systematically justified if and only it comprises all and only those desires that survive after desires which we provisionally assume to be justified are formed into a maximally informed and coherent and unified whole—we still allow our *confidence* that certain desires are irrational to play a significant role in fixing what reasons there are. The role they play, however, is essentially that of providing us with a *defeasible constraint* on what reasons there are.

Gert on the Category of Rationally Allowed Acts

One of the reasons Gert gives in favor of his own theory of rational action is that it has the advantage over alternative theories of making room for acts that are rationally allowed, though not rationally required. So long as agents don't act so as to harm themselves, Gert claims, whatever they do is rationally allowed, notwithstanding the reasons that they might have for acting in one way rather than another. This was the point I emphasized when I spelled out the third and fourth implications of Gert's definition of rational action.

For example, suppose an agent faces a choice between buying himself a holiday and giving his money to a charity that alleviates the suffering of others, and suppose further that parting with his money does cause him some harm, at least insofar as it reduces his freedom by cutting down the number of options that he will have in any future choice situation. In such a case, Gert claims that

the man is rationally allowed to give the money to charity, because the harms to others that are thus avoided more than adequately compensate for the harm that he thus causes himself. (It was in this sort of case, you will remember, that Gert claimed to find the grain of the truth that lies in Kantian theories.) But, Gert insists, the man described is not rationally required to give the money to the charity. For he is also rationally allowed to buy himself a holiday, as the benefits he gains by doing that would also more than adequately compensate for the harm he causes himself. (It was in this sort of case, you will remember, that Gert claimed to find the residual grain of truth that lies in egoism.) So long as an agent's action does not harm himself overall, he is therefore rationally allowed to do whatever he likes.

I think that Gert is right to emphasize that there is a category of rationally allowed acts alongside those that are rationally required, and I think that he is also right to insist that some of the so-called conflicts between morality and self-interest are best thought of in this sort of way. Agents have a free choice to decide in which way they will act, at least within certain limits. They are not rationally required to act in the one way or in the other. Moreover, I think he is right that a theory of rational action that suggested otherwise would be flawed in a quite decisive way. But I think he is wrong that his hybrid conception of rationality is unique in making room for the category of acts that are rationally allowed, as opposed to rationally required. The view of rational action I described above makes room for the class of rationally allowed acts as well. It does so by allowing that certain reasons are disjunctive in form.

Thus, in the case just described, we can imagine that if the man had the set of desires that would result if we take those of his desires that we provisionally agree to be justified, and then add and subtract desires in order to find a set that is maximally informed and coherent and unified, then he would desire himself, in the circumstances of action he faces, either to buy himself a holiday or to give his money to charity. Since his desire has this disjunctive content it follows, on the theory of reasons for action that I described, that the reason for action he has is disjunctive in form as well: he has a reason either to buy himself a holiday or to give his money to charity. If no other reasons are in play, apart from this disjunctive reason, and if the man believes that this is so, and if his belief eludes rational criticism, then, according to the view of rational action described above, it follows that the man will be rationally allowed either to buy himself a holiday or to give his money to charity. Whatever he does will be rational notwithstanding the fact that he had adequate reasons for acting in the alternative way.

The upshot is thus that Gert is wrong to claim that his theory is alone in making room for the category of acts that are rationally allowed, as opposed to being rationally required. Indeed, it seems that any theory can make room for that possibility simply by acknowledging the existence of disjunctive reasons.

Conclusion

Despite Gert's various ingenious arguments for his complex hybrid conception of rationality, it seems to me that, in the end, we should reject it. This is not to say that there isn't much Gert says with which we can and should agree, there most certainly is: the idea that a theory of rational action must combine egocentric and nonegocentric elements; the idea that certain desires wear their irrational nature more or less on their sleeve; the idea that rational actions come in two forms, acts that are rationally required on the one hand and acts that are ration-

ally allowed on the other; and so on. My point is simply that we do not need to move beyond a theory which is, for all intents and purposes, much like the cool moment desire theory insofar as we agree with Gert on these crucial issues.

Notes

* An earlier version of this paper was read at *A Conference on Gert's Moral Theory* held at Dartmouth College, May 1999. I would like to thank Robert Audi, Ruth Chang, David Copp, David Cumiskey, Bernard Gert, Joshua Gert, Shelly Kagan, Geoffrey Sayre-McCord, Stephen Nathanson, Walter Sinnott-Armstrong, and Susan Wolf for their many helpful criticisms and suggestions.

1. All otherwise unexplained references are to Bernard Gert, *Morality: Its Nature and Justification* (New York: Oxford University Press, 1998).

2. Walter Sinnott-Armstrong has suggested that we can usefully make a distinction between the claim that *there is a reason for an agent to act* in a certain way and the claim that *the agent has a reason to act* in that way. In his view, claims of the latter sort (claims about *the reasons that agents have*) are best analyzed as claims about the beliefs that agents have about claims of the former sort (in other words, they are claims about *the reasons that agents believe that there are*). In these terms, the commonsense conception of practical rationality insists that we ask not what reasons there are for agents to act in certain ways, but rather what reasons they have for acting in those ways. As we will see, Gert himself says that he embraces something like this distinction (65-66). Despite the good company I would keep, I am reluctant to accept this distinction, as ordinary usage seems to me not to support it. "Though John has a reason to act in a certain way, he doesn't believe that he has a reason to do so" doesn't seem to be equivalent to the claim that though John believes that there is a reason to act in a certain way, he doesn't believe that he believes there is a reason to act in that way. It seems, instead, to be equivalent to the claim "Though there is a reason for John to act in a certain way, he doesn't believe that there is a reason for him to act in that way." Claims about the reasons there are for agents to act in certain ways thus seem to me to be equivalent to claims about the reasons they have for acting in those ways. When an agent asks whether there is a reason for him to act in a certain way, that is just a way of asking whether he has any reason to act in that way.

3. Note that when Gert says that someone is "fully informed," he means not that she is omniscient, but rather that she has all the information that is suitably available to her (184, note 11).

4. Though remember what Gert means by "fully informed" (see note 3 above).

5. Michael Smith, *The Moral Problem* (Oxford, Eng.: Blackwell, 1994).

6. Remember that, for Gert, the greater harms referred to here are death, pain, loss of ability, loss of freedom, and loss of pleasure, and the benefits referred to are greater consciousness, ability, freedom, or pleasure. This will be important subsequently (see note 9 below).

7. Why could it never be rational? According to Gert, it is a constraint on any adequate theory of rational action that it not classify as rational acts that "every fully informed rational person would advocate that all persons for whom they are concerned, including themselves, never [perform]" (31), and an agent's knowingly harming himself when his doing so will not be compensated for by the production of benefits or the avoidance of harms is such an act.

8. Michael Smith, "The Incoherence Argument: Reply to Schafer-Landau," forthcoming in *Analysis*.

9. How likely is it that a fully rational agent would have one or another of the desires for death, pain, disability, loss of freedom, or loss of pleasure, in circumstances in which there is no compensation to be gained from the avoidance of harms or the production of benefits? Well, remember that the harms referred to here are death, pain, loss of ability, loss of freedom, and loss of pleasure, and that the benefits he refers to are greater consciousness, ability, freedom, or pleasure (see note 6). The question is therefore: how likely is it that a fully rational agent would desire death, pain, disability, loss of freedom, or loss of pleasure in circumstances in which something other than these harms and benefits provides the compensation? For example, how likely is it that a fully rational agent would desire death, pain, disability, loss of freedom, or loss of pleasure in circumstances in which the loss enables the preservation of a work of art or area of rain forest? Certainly it doesn't seem impossible.