

[*Southern Journal of Philosophy*, Summer, 1996, XXXIV, No. 2: 147-168]

### EXPRESSING A GOOD WILL: KANT ON THE MOTIVE OF DUTY<sup>1</sup>

If any action is to be morally good it is not enough that it should *conform* to the moral law--it must also be done *for the sake of the moral law*: where this is not so, the conformity is only too contingent and precarious, since the non-moral ground at work will now and then produce actions which accord with the law, but very often actions which transgress it.<sup>2</sup>

So Kant claimed in the preface to the *Groundwork*, and many more times throughout his moral works.<sup>3</sup> Kant held that only the motive of duty avoids leading by mere accident to the performance of actions which are dutiful. But, he thought, a morally good action should have a motive that does not produce dutiful actions by mere accident. So only actions from duty are morally good. Thus, Kant holds that it is a condition of an action's being morally good that its motive does not lead to dutiful actions by accident.<sup>4</sup> And his well-known view that only dutiful actions from duty have moral worth appears to be based partly on the claim that only the motive of duty meets this condition.

In what follows, I will consider two common criticisms of this position. One accepts the condition that morally good actions must have motives that produce dutiful actions more than merely "now and then", but rejects the claim that only actions from duty meet this condition.<sup>5</sup> Benevolence, love and many other noble motives, according to this criticism, do not produce morally right actions by mere accident. So actions motivated by these motives should not be regarded as lacking moral worth for this reason. The other criticism rejects this condition altogether, arguing that it is overly demanding.<sup>6</sup> For even if a motive would not *always* lead to a moral action it still might be a laudable

motive for persons to have and act on. And so the second criticism also concludes that actions motivated by these motives should not be regarded as lacking moral worth for this reason. Thus, both criticisms aim to show that there is nothing unique in how the motive of duty produces dutiful actions, and that therefore Kant is unjustified in arguing on this basis that actions from other laudable motives have no moral worth.

As I see it, these criticisms are successful only if Kant's position was that duty is *more reliable* than other motives in producing dutiful actions. For it is quite possible that many motives produce right actions far more reliably than the motive of duty. But this, I shall argue, is not Kant's position, nor should it be. Even if the motive of duty were completely unreliable--in the sense of being weak, changeable, or not always available to act on--it would still avoid producing dutiful actions by accident in a different, and more important, sense: When one acts dutifully from the motive of duty, one's motive is, as Barbara Herman has put it, "internally related" to the dutifulness of one's action.<sup>7</sup> My main concern here will be with what it means to say that a motive is internally connected to some feature of the actions it motivates, why it is morally significant that the motive of duty is alone internally connected to the dutifulness of one's actions, and, more generally, what role this claim might play in Kant's overall theory.

In the first section, I will explain how the issue of the connection between motives and actions arises, and briefly set out the two interpretations of Kant's claim that the motive of duty alone leads non-accidentally to dutiful actions. I call these the "internal connection" interpretation and the "reliability" interpretation. In the second section, I explain the two criticisms of Kant's claim and how they are based on the reliability interpretation. These criticisms should lead us in turn to adopt the internal connection

interpretation, but the latter leaves questions about the importance of establishing such a connection. So in the third section, I clarify what this internal connection is and in the fourth explain why it is important. In the fifth section, I explain what I take the role of Kant's claim about the motive of duty to be in the argument of *Groundwork I*.

### *I. THE CONDITION*

Let us summarize this condition under which an action is morally good in the following way:

(C) The performance of a dutiful action is morally good only if it is no accident that its motive led to it.

A familiar sort of example will show how C figures in Kant's account of moral worth. Suppose that Dorit helps Jay, a recently disabled neighbor, by doing his shopping. Jay is a gentle person who has no one else nearby to rely on and would appreciate her help; Dorit is retired and in full health, and has no other pressing demands. Let us suppose, in other words, that Dorit does something that she ought to do--her duty. Now had we occasion to evaluate her action, our attention would naturally be drawn to her motive. But what sort of motive would make us regard her action as morally good?

A common sense answer is that a sympathetic motive would. After all, Jay was in a bad way. Dorit may have been emotionally affected by his ill fortunes, sympathized with him, and, in virtue of that, was motivated to help. Surely we would think highly of this. But Kant famously argues that, even so, her action would have no moral worth.

Were her motive sympathy, he argues, her action would stand

on the same footing as [actions from] other inclinations--for example, the inclination for honour, which if fortunate enough to hit on something

beneficial and right and consequently honourable, deserves praise and encouragement, but not esteem; for its maxim lacks moral content, namely, the performance of such actions, not from inclination, but from duty.<sup>8</sup>

There are two ideas here. First there is an echo of the passage from the Preface with which I began. A sympathetic motive will lead to a dutiful action only by the grace of good fortune. Hence, Kant thinks that an action motivated by sympathy will fail to meet C. The second idea explains why this is so and why such an action deserves no esteem: The sympathetically motivated person's maxim has no moral content. A "maxim" is the motivating principle, or more simply the motivating rationale, of a person's action.<sup>9</sup> Although what exactly Kant meant by "maxim" is controversial, the basic idea is clear enough.<sup>10</sup> Maxims can be thought of as general principles of behaving in a certain way to gain a certain end given certain circumstances. For instance, my rationale for taking the bus rather than driving to work might be that taking the bus under these circumstances would conserve fuel; or my rationale for going to the gym every day might be that exercising daily will maintain my health. Hence, "to take the bus to work in order to conserve fuel" and "to exercise daily to maintain health" are roughly what Kant has in mind. A person's maxim contains, in a perspicuous way, the rationale which motivates her to act as she does. It thus contains a description of her action, circumstances and end that led her to act. So in saying that acting from sympathy is acting from a maxim with no moral content, Kant is saying that, stated in a perspicuous way, the rationale which brings the sympathetic person to act as she does has no moral content, and so her action has no moral worth. A sympathetic person's rationale typically

does not describe her actions as being required by duty, for that is not a feature of an action that will typically engage her sympathy.

Kant of course held that our duty is to act only on maxims that can at the same time be consistently willed as universal laws, that our duties are defined by the Categorical Imperative (hereafter, "CI"). Quite clearly he would hold that someone whose motive for her action was that it uniquely met the CI would be a motive with a moral content, even if her rationale would not explicitly appeal to the fact that this is what makes something a duty. So the content of someone who acts from duty may well make no mention of the fact that her action is a duty because it mentions something equivalent to this, namely, that it uniquely meets the CI. In this case, there is a necessary connection (on Kant's view) between an action's satisfying the description of being a duty and its satisfying the description of uniquely meeting the CI. So one's motive will have a moral content not only if one's motivating rationale describes one's action as being dutiful, but also if it describes it in terms which (accurately) define actions as duties. (I will return to this point below when I discuss the role of Kant's claims about the connection between the motive of duty and the dutifulness of one's actions.)

Now if a maxim contains the description of the action, circumstances and end that brought a person to act as she did, then if that person was motivated simply by her sympathy, then her rationale would not be that her action was required by duty. This gives us a straightforward way of thinking about "moral content": A person's maxim has moral content only if her motivating rationale describes the action as required by duty. Our maxim must reflect the fact that we conceive of our action under the heading of duty in order for our motive to have moral content.

Kant takes himself, of course, to be only mining “ordinary knowledge of morality”, and so he believes that, were we to reflect just a bit on our own practice, we would agree that this is just what we think of as a morally good action. Clearly he would be wrong if he meant that ordinary people would describe morally good actions in just these terms. I take it that he means that with a little philosophical acuity, one could discern just this position in ordinary moral thinking. He also thinks that this bit of ordinary moral knowledge has a justification, and I think it is clear enough from both the remark in the Preface and the aside in this passage to see that this justification has to do with C. Kant’s thinking in this regard can be reconstructed in the following way. Suppose  $\Phi$  is a dutiful action. Then P’s performance of  $\Phi$  is morally good only if her maxim has a moral content. But P’s maxim has a moral content only if it is no accident that her motive produces  $\Phi$ . And it is no accident that P’s motive produces  $\Phi$  only if her maxim describes  $\Phi$ -ing as dutiful. So P’s performance of  $\Phi$  is morally good only if her maxim describes  $\Phi$ -ing as dutiful. If this reconstruction is correct, then C is a condition for the moral goodness of an action because it is a condition for its maxim’s having a moral content. And this in turn is why Kant holds that only the motive of duty will meet this condition.

But what does it mean to say that the motive of duty alone does not lead to dutiful actions merely “now and then”, or by “fortune” or “accident”? What is it that ordinary moral knowledge holds is a condition of morally good action? There are, it seems, two main interpretations of what this might be, the “reliability” interpretation and the “internal connection” interpretation. On the former interpretation, a motive that leads to dutiful actions by accident leads to them *unreliably*.<sup>11</sup> Call a motive “reliable” if it makes the performance of a dutiful action more probable than not. There are many

properties of motives that might make it improbable that they will lead to dutiful actions. Many will not have to do so much with the *content* of a person's rationale for acting, with a motive's *cognitive* aspects, but with its *noncognitive* aspects, such as its invariability, availability or strength. Hence, being changeable, weak, or not always available to act on, for instance, will make a motive unreliable in a noncognitive sense.<sup>12</sup> A failure of reliability in this sense will be when, because of such properties, it will not be probable that if an agent has an unreliable motive, she will succeed in acting as she intends to act, and if what she intends to do is a dutiful action, then it will not be probable that she will act dutifully.

A person's motive might be unreliable for a quite different reason, however. For suppose a person's rationale for acting makes no reference to the fact that her action is dutiful. Then, no matter how strong, invariable and available such a motive is, it will still not be perfectly probable that if it motivates a person, she will perform a dutiful action. For the description of the action contained in such a rationale will not single out all and only dutiful actions. For instance, the description of the action contained in Dorit's motivating rationale might have been something like "helping a neighbor in need", with no mention of the fact that it is dutiful. But an action that satisfies this description need not be a dutiful action; one can easily imagine helping a neighbor in need by doing something one ought not do (lying, for example, might satisfy such a description, since there does not seem to be anything about lying in itself that is unsympathetic). Hence, if this is what motivated Dorit to help, then even if it were strong, invariable, and always available, it could well have led her to do something that is contrary to duty. And if such a rationale could lead her to do something contrary to duty, then it will not be perfectly

reliable, not because of its noncognitive properties, but because of its content. Obviously, it is more probable in this cognitive sense that a motive whose content is a sympathetic rationale will lead to actions which are dutiful than one whose content is a malicious rationale. We can imagine a range of act descriptions that single out a set of actions that are more or less dutiful. But the idea is clear enough: If the description of the action contained in a person's rationale for acting does not single out all and only dutiful actions, then it will be not be perfectly reliable in leading her to act dutifully.

It is this sort of idea that is behind the "internal connection" interpretation, which has been advocated by Barbara Herman. The concern here is not with a motive's overall reliability in leading to dutiful actions.<sup>13</sup> Indeed, even if it turned out that the set of actions that is described in a sympathetic person's rationale is composed of almost all and only dutiful actions, and so is highly cognitively reliable, we still seem to be able to conceive of actions that are immoral but engage a sympathetic motive. Moreover, it is a contingent fact that the act description contained in a sympathetic rationale is aligned with dutiful actions. So even if a motive is highly reliable in both a noncognitive and cognitive sense, it may not be *perfectly* reliable in a cognitive sense.

Now, imagine a motive that *is* perfectly reliable in a cognitive sense. This will be the case if a motive has what Herman calls an "internal connection" between it and the dutifulness of the action.<sup>14</sup> On her view, only if the motive on which the agent acted was internally connected to the dutifulness of her action will it be, in the relevant sense, no accident that it led to the dutiful action. And only if the motive on which the agent acted has a moral content in Kant's sense will there be such a connection. Herman does not spell out what an "internal connection" between a person's motive and the dutifulness of

her action is. For the time being, assume that it means simply this: (i) The agent is disposed to act dutifully, (ii) she believes that her action is dutiful (or has those features that accurately define an action as a duty), and this belief is reflected in the fact that her maxim describes her action as dutiful (or has those features that accurately define an action as a duty), and (iii) her action is in fact dutiful.<sup>15</sup> What she is motivated to do and what she does are internally connected in the sense that she does exactly what she was motivated to do. What she did was dutiful, and had she believed that it was not dutiful, or not been disposed to perform dutiful actions, then *ceteris paribus* she would not have done it. We can see, then, that if a motive has a “moral content” as we set this out above, then it will be internally connected to the dutifulness of one’s dutiful actions.

Notice that a motive might be reliable in the noncognitive sense, but unreliable in the cognitive sense. And a motive might be reliable in the cognitive sense, but unreliable in the noncognitive sense. Thus, it might turn out that if a person is motivated by sympathy rather than duty, she will be more likely to actually perform dutiful actions, simply because duty is much less reliable than sympathy in the noncognitive sense. Indeed, it might turn out that, because of our natural constitutions, duty is very unreliable in the noncognitive sense, while sympathy is very reliable. And so, all considered, it would be more probable that a person motivated by sympathy will act dutifully than that a person motivated by duty will. Thus, if those who interpret C as having only to do with how *probable* it is that a given motive will produce dutiful actions, it might well turn out that sympathy, as well as many other motives, will do better than duty. In that case, Kant will have been quite wrong about the uniqueness of the motive of duty on these grounds.

However, if Herman is right, then the fact that it is more probable that a persons acting from motives other than duty will in fact perform dutiful actions may well be irrelevant.

## *II. C CRITICIZED*

Two criticisms of Kant's view show quite convincingly that if C were meant as a reliability condition, then he will fail to make his case. The first, as I mentioned above, agrees that a moral motive must be reliable, but argues that this does not exclude motives other than duty. Someone advancing this kind of criticism will probably think that Kant is in the grip of a crude picture of inclinations such as compassion or love. Kant, on this view, believes that such impulses are typically weak, changeable and easily overcome by other dark impulses. Consequently, although they are sometimes able to produce dutiful actions, sometimes they are not; it all depends on whether we are unlucky enough to find ourselves fraught with strong temptations. And this is why Kant would think that it is by mere fortuity that, for instance, Dorit's sympathy led her to perform the dutiful action of helping Jay.

The Kantian position is supposed to be that Dorit's sympathy, though it happened to lead to the dutiful action, was unlikely to have done so. Had Jay been a smelly, drooling oaf for instance, it probably would have withered. But the purely rational motive of duty, unruffled by such circumstances, is very likely to lead to the dutiful action. So, Kant apparently concludes, it is the morally preferable motive. But, as Lawrence Blum argues, this sort of argument criticizes altruistic inclinations

for how they actually operate ... whereas the rationalistic motives are not scrutinized for how they do operate, but are put forward only for how they are capable of operating ... Once one asks how a sense of duty actually

operates as a human motive, the great gulf which the Kantian view portrays between it and the altruistic emotions virtually disappears.<sup>16</sup>

While altruistic feelings *can* be weak, changeable, unavailable to act on, Blum contends, they need not be. And, after all, our interest in duty may be just as weak, changeable, and unavailable to act on. Dorit's motive of duty might also have withered had Jay been a foul smelling, drooling oaf. So there is no *special* reliability problem with inclinations such as benevolence. Let us call this the "no distinction" objection. Someone raising it accepts C, and so agrees that a moral motive must be *reliable* enough to overcome wayward temptations and produce dutiful actions. But it then argues that this gives no leverage to duty over other motives. On the one hand, the motive of duty sometimes fails to produce dutiful actions. So although benevolence is not always *actually* able to overcome contrary impulses, neither is the motive of duty. On the other hand, even if the motive of duty is *capable* of overcoming such wayward impulses, there is no reason to think that feelings such as benevolence are not equally capable. So C does mark a difference between feelings such as compassion, which are reliable, and feelings such as "liking" or "being favorably disposed toward", which are not. It does not, however, mark any difference between the motive of duty and altruistic motivation. In fact, it is difficult to see how there could be any significant difference between the two on this score. Duty might even fare badly in comparison.

Moreover, even if benevolence and compassion are not *perfectly* reliable in a cognitive sense, surely as a matter of fact they do single out mostly dutiful actions. The rationale of a person motivated by such interests describes actions as, for instance, helping others. Since helping others is typically a moral thing to do, such motives are

fairly reliable in the cognitive sense. But duty, even if *perfectly* reliable in a cognitive sense, may well be extremely *unreliable* in a noncognitive sense. Most people may well not care about duty very much. And so, as I mentioned above, a person who acts from duty may not be acting on a motive that makes it more probable *over all* that she will act dutifully. At the very least, Kant has failed to give any good argument that would show that benevolence, compassion and other motives are overall more unreliable than the motive of duty.

The second objection to the reliability interpretation of C is that it places the bar too high. We are being too harsh to demand that Dorit's motive be reliable, at least in the measure that Kant's view seems to require. This second objection is well summarized by Holly Smith's rhetorical query:

Would we want to say that [an agent's] present act is not praiseworthy just because there are possible alternative circumstances of this kind in which his morally good desire would not prevail? This is completely implausible.<sup>17</sup>

Smith does not directly present this as a criticism of Kant's own view. But it is clear that she takes it to be one way of understanding his views, and it nicely sums up this concern with such an understanding. A similar concern is voiced by Bernard Williams. He argues that Kant did not regard our sentiments as possible sources of moral motivation because they are contingent, changeable features of agents.<sup>18</sup> But this is an unreasonable view, he argues, since no motive can be anything other than a contingent, changeable feature of agents. This objection, then, either leads us back to the former view, that under any reasonable conception of reliability, there is no distinction to be made between duty

and many other motives; or it simply concludes that reliability of motivation is not a relevant condition of moral worth. Given the latter, if Dorit does what she ought from a benevolent desire, we should not think less of her action because, had she faced threats or mouth-watering temptation, it would not have prevailed. Instead, we should be satisfied with the fact that her desire was benevolent and sufficed in the circumstances as fortune delivered them.

This objection could, if it wished, even admit that a rational interest in duty is alone the only motive that is reliable in either a cognitive or noncognitive sense. Its contention is rather that it is “completely implausible” to accept this as the standard by which to distinguish motives which confer moral worth on actions from those which do not. Call this the "unreasonable demand" objection. From its perspective, we should settle for less than reliability. And when we do, we find that we can again see no advantage for the motive of duty.

These objections to reliability as a condition of moral worth will lead us, if we are sympathetic to Kant, to adopt Herman’s interpretation of C, which focusses simply on the content of our maxims. On her view, even motives that reliably lead to right actions may still lead to them completely by accident. Consider, for instance, unadorned sympathy. This, as a motive, is directed toward a certain set of actions (such as those that promote another's well-being) that is not necessarily coextensive with the set of moral actions. Left on its own it therefore has no scruples about prompting actions which promote the well-being of others but conflict with duty. So sympathy leads to dutiful actions only by accident--that is, when promoting another’s well-being also happens to be dutiful. Now given that only the motive of duty prompts us to perform all and only dutiful actions, C,

thus interpreted, appears to give leverage to the motive of duty over other kinds of motivation.

Nor does this seem to be an unreasonable demand. Herman represents the concern with reliability as tied to the consideration of counterfactual situations in which competing interests are stronger, rather than of the probability that a given motive will lead to a dutiful action. She thus argues that a non-accidental connection between one's motive and the dutifulness of one's action does not require that one's motive be able to produce the dutiful action in alternative circumstances in which contrary inclinations are stronger--that it be reliable in a noncognitive sense.<sup>19</sup> So Herman agrees with the unreasonable demand objection in holding that it is unreasonable to require that the motive of a morally worthy action should be noncognitively reliable. Indeed, even if our interest in duty tends, relatively speaking, to be ever so faint, in favorable circumstances (such as when there happens to be no wayward temptations) it could still lead one to perform one's duty, and when it does, one's action would have moral worth. In her view, the counterfactual performance, and so noncognitive reliability, of the motive of duty concerns only the agent's virtue, not the moral worth of her action.<sup>20</sup> However unlike Smith and Williams, Herman denies that motives other than duty can contribute to the moral worth of an action. For other motives are not internally connected to the dutifulness of one's actions. So it is no accident that an agent's motive led her to the dutiful action just in case it was an interest in duty, and nothing else, that led her to act. This appears to avoid setting the bar too high while maintaining a distinction between duty and other motives.

But while Herman's account avoids some of the problems of the reliability interpretation, it faces difficulties nonetheless. For why should it be *morally significant* that one's motive is internally connected to dutiful actions? The reliability interpretation at least gives us an intuitive sense of why meeting it might be morally important: If it is important that one's actions be dutiful, it is important that one's motives be reliable in actually producing them. Yet on Herman's view, it is not morally important that one's motives be reliable in producing right actions. It is only morally important that they be internally connected with the dutifulness of one's actions. But why?

Second, her account does not seem adequate as an interpretation of Kant's own views. For in his view a moral motive must "produce actions which accord with the law" more than merely "now and then".<sup>21</sup> But Herman's account does not show that the motive of duty would produce dutiful actions more than merely now and then. For she does not think that the motive of duty has to lead us to dutiful actions even were competing inclinations stronger. Yet it seems that the motive of duty will produce dutiful actions more than merely now and then only if it *would* do so. So how can it be true both that a motive leads to the dutiful action more than merely "now and then", *and* that it would not have led to it had contrary inclinations been stronger?<sup>22</sup>

It therefore appears that if, against Herman, we think that it matters whether our motives lead to dutiful actions more than merely now and then, then we will have to accept the reliability interpretation of C, and, consequently, the loss of support for the claim that the motive of duty is unique. But if, with her, we reject this, we appear to have no rationale for C, nor any way to make good Kant's claim that the motive of duty does not lead to dutiful actions merely now and then.

In the following sections, I will explain how we can preserve the internal connection interpretation. First, I will explain in what sense C can show that the motive of duty does not lead to dutiful actions merely “now and then” without being interpreted as establishing the overall reliability of a motive. Second, I will argue that the reliability interpretation is anyway inconsistent with Kant’s claims about the value of a good will. Finally, I will show how the internal connection interpretation fits very well with these claims, and makes good sense as a condition of morally good action.

### *III. RELIABILITY AND PURITY*

In holding that a person’s motive for a dutiful action must not have led to it by accident, C surely does seem to require that it be true that it would have led to the dutiful action under other more difficult circumstances. But it does not follow from this, as I think Herman assumes, that therefore C would have to be interpreted as being concerned with the overall reliability of a motive. For whether a given motive is reliable over all depends both on its content and such things as its strength, availability and invariability.<sup>23</sup> This is why C avails itself of two interpretations. Suppose, for instance, an agent was motivated by duty alone to do her duty, and so her rationale for acting would have been simply that it was her duty. Yet also suppose that duty would not have prevailed over other motives under other circumstances. The internal connection interpretation can be seen as holding that if a motive which produced a dutiful action would have led to the dutiful action in any circumstance were *it to have prevailed over all other motives*, then it was no accident that the person performed the dutiful action. This is surely the case here. For this agent is simply motivated to do her duty. She did not, if she was motivated by duty alone, intend to do her duty only if there were no stronger competing motives. We

might say that, on the internal connection interpretation, the fact that her maxim does not motivate her reliably is irrelevant to the fact that her motive here and now would have led her to do what is her duty under any circumstance. Alternatively, suppose an agent was so strongly motivated by her sympathy or benevolence in dutifully helping someone that her sympathy would have prevailed under any circumstance. We could imagine circumstances having been different in such a way that her sympathy was still engaged, and would have prevailed, but her action would not have been dutiful.

In former case, the internal connection interpretation will say that the agent's motive did not lead to the performance of her duty by accident. Notice that this is so even if the overall probability that she would have done her duty would have been higher if she were motivated by sympathy (given that her sympathy is both cognitively and noncognitively reliable) than it would have been were she motivated by duty (since her interest in duty is not very reliable in a noncognitive sense, though in a cognitive sense it is perfectly reliable). The point is that, in the latter case, the agent intended to do what sympathy indicated, and the fact that it was her duty was not what led her to act. And were it to prevail over all other motives, under certain circumstances this would have led to an action which violated her duty. Indeed, suppose she were to live in a world in which no such conflict would arise. Suppose, that is, that in some world everything that sympathy leads her to do is the same as what it is dutiful for her to do. Then, because she is so strongly motivated by sympathy, in that world she always does what is dutiful. Yet according to the internal connection interpretation, any such motive would be based on a maxim which would advocate violating a demand of duty under some possible circumstance, and so will still fail C. And it will fail because it is a contingent fact that

the circumstances in this world are not those which the maxim picks out as being suitable for violating the demands of duty.

For example, suppose we think that Dorit's decision to help Jay was determined by the principle to help others, to some extent and degree, when they are in need of it--a Kantian imperfect duty. But later we come to suspect that, had he been a foul smelling, drooling oaf, she would not have helped. We would then have two alternative explanations by which to revise our initial evaluation. Either we retract the original supposition that she acted on the dutiful principle (since it is a contingent feature about Jay that he is not a foul smelling, drooling oaf), or conclude that, although she acted on a principle with a moral content, it would not have prevailed under those trying circumstances. In the latter case, we simply judge that she lacks fortitude in her determination to stick to moral principle under any circumstance. But for all that, according to the internal connection interpretation her action was still morally good. In the former case, however, we will be thinking that Dorit did not act from duty at all, and so her action was without any moral worth, even if it is more reliable as a motive. Perhaps we will conclude that her maxim contained an "escape clause" to deal with cases where duty involves helping foul smelling, drooling oafs. It seems, then, that we can indeed see C as focussing simply on the content of a motive, even if it is concerned with whether a motive would have produced a dutiful action under more difficult alternative circumstances.

But it might be objected that in fact Kant cannot separate reliability from purity of content in the way that I am suggesting for the following reason: Suppose we think that Dorit acted on a principle to help others when they are in need, even when it is

inconvenient, unpleasant, and so on. But when we further reflect on the ways in which circumstances might have been different--say, helping would have meant giving up a one-time opportunity to fulfill a personal "ground project"--we decide that she would then have been moved little by her principle. One might insist that we have to construe the content of her maxim as the principle she would have acted on, had she considered difficult circumstances such as this. That she did not actually consider them at the time of acting is irrelevant, for if the notion of a maxim is to be useful, it must take into account the fact that we very rarely actually consider what we would do were circumstances more difficult. And in this case, had she considered the more difficult circumstance, she would have acted on a principle that made an exception for it. Further, suppose that she did consider it, claimed that she would still do the dutiful thing, yet the truth is otherwise. Surely we must conclude that her "dear self" had simply clouded her self-knowledge. Her real principle is impure, and she is, like the rest of us, simply not in a position to determine with certainty that it is.<sup>24</sup>

Although there is textual evidence to support this objection, I think that there is enough against it to show that Kant does, after all, view strength of motivation as a separate concern from its purity. For he not only claims that a good will--a will motivated by duty--may be "entirely lacking in power to carry out its intentions",<sup>25</sup> but also that weak motivation "can indeed coexist with the best will", being indicative only of a lack of virtue.<sup>26</sup> So even if we may be uncertain of the purity of our own or another's intent, Kant thinks that purity would still not guarantee the reliability of motivation to carry through with it. Indeed, Kant needs a dimension of strength in order to discuss virtue at all, since he takes it to be "the moral strength of a man's will in fulfilling his duty".<sup>27</sup>

Moreover, it seems sensible to view things this way. For suppose that Dorit had considered the possibility that helping Jay could conflict with a ground project, and yet seemed quite sincere in her commitment to act on moral principle. We need room both to acknowledge Dorit's sincerity yet still wonder whether, had things been different, she was really strongly motivated enough to have carried through.

But now someone pressing the "unreasonable demand" objection may want to shift ground. She may rephrase her objection and insist that it is unreasonable, excessively puritanical, to hold that the maxim of a morally worthy action must be to do the dutiful action no matter how things might have been different. For instance, principles such as "I will keep my promises, except when doing so will significantly harm others" or "except when doing so will endanger my life or the life of those close to me", or even "except when doing so leads me to violate a 'ground project' which gives my life meaning", endorse violating one's duty whenever it conflicts with certain important interests. But what is morally wrong with such principles? Surely it would be a mark of fanaticism foreign to even most die-hard Kantians to stick to one's guns under such circumstances.

But it is not at all clear that all such principles really endorse violating one's duty under certain circumstances. They seem, rather, to simply acknowledge that under certain circumstances, the action in question would not be one's duty at all. For, moral skepticism aside, what else could an "important" interest mean in this case but a *morally* important interest? A principle of keeping promises whenever it is morally required that also acknowledges circumstances in which it is not, is quite different from one of not keeping one's promises under certain circumstances *even if* it is morally required. The

first looks to moral reasons to ground the exclusion of certain circumstances, whereas the second does not and cannot.

Whether such exceptions are consistent with the nature of moral demands depends on what the proper account of such demands is, but this issue can be left on one side here. Suppose, for instance, duties were determined by impartial sentiments of approval, to take one non-Kantian account. “Keep promises under all circumstances” would not be a duty under such an account just in case there were circumstances in which an impartial spectator would not approve of it. But what C requires is simply that a morally worthy action be motivated by a maxim to perform the morally required action under any circumstance in which it is morally required—a principle to not make exceptions *given* that one’s action is morally required (here, when an impartial spectator would approve of it). To the extent that the refocussed unreasonable demand objection has force, then, it is because it has shifted its attack away from C and onto Kant’s supposed rigorous conception of what duty demands. This is a crucial point that I will emphasize in the next sections.

The internal connection interpretation of C, then, is this: A motive avoids leading to a dutiful action by accident when and only when its content advocates doing one’s duty no matter how circumstances might have been different, and so would lead to the dutiful action under any circumstance in which it prevailed. In one sense, this clearly singles out the motive of duty without being too demanding of a condition. Only the motive of duty will advocate the dutiful action no matter how circumstances might have been different, for that is the content only of the motive of duty. Benevolence, for example, will advocate benevolent actions no matter how circumstances might have been different. But

benevolent actions under certain circumstances can violate duty. So the motive of duty is unique. Moreover a person may act from duty and yet it not motivate her reliably enough, in a noncognitive sense, to have prevailed under any circumstance. So there is nothing demanding in this requirement.

The question is whether there is some reason to restrict C as applying to the content alone, apart from the non-cognitive reliability of a motive for the purposes of evaluating the moral worth of an action. Surely, one might insist, both purity *and* reliability matter to moral worth, even if they are separable issues. The reason C should be taken to concern only purity of motivation has to do with the point of C itself, to which I will now turn.

#### *IV. MOTIVATING C*

Notice first that if we think of the motive of duty as Kant eventually does, that is, as a purely rational interest, then C does not help Kant to show that moral motivation is purely *rational* motivation. For, as I mentioned above, C is independent of his account of duty. Recall that in setting out what a “moral content” was, we noted that a motive has a moral content not only if it one’s rationale describes one’s action as being dutiful, but also if it describes it in terms which (accurately) define actions as duties. But it is a substantive issue between Kant and other moral theorists what the proper analysis of duty is. What I want to point out here is that Kant could be *right* about the connection between the motive of duty and the dutifulness of actions (as construed by the internal connection interpretation), yet *wrong* about what it is that makes an action a duty. Let me explain.

The general intuition behind the internal connection interpretation is simply: A motive non-accidentally leads to the performance of an action of a certain kind only when

it is an interest simply in performing the same kind of action that it led the agent to perform. So it is no accident that one's motive led to a helping action only if it was simply an interest in helping that motivated one to do it. It is no accident that one's motive led to a harming action only if it was simply an interest in harming that motivated one to do it. Likewise, it is no accident that one's motive led to a dutiful action only if it was simply an interest in acting dutifully that motivated one to do it.

But suppose someone, call her a “sympathy theorist”, holds that moral requirements are defined as what sympathy leads her to do. Details aside, she could advance right along with Kant the views that (a) only when it is no accident that one’s motive leads to dutiful actions is it of moral worth, and (b) it is no accident only when one's motive was to perform the dutiful action. She simply has a different view of what duty is, and therefore what the motive of duty is. Thus, the motive of duty, on the sympathy theory, is simply the motive of sympathy. And when she acts on sympathy when doing her duty--that is, the action that her sympathy leads her to do--it is no accident that it led her to perform the dutiful action. This shows that C by itself does not put Kant's position into conflict with alternative accounts of duty. All accounts are still live options, and our selection of one must be made on other grounds.

This also indicates what the rationale for the internal connection interpretation of C is. The rationale for the reliability interpretation, as we have already seen, is quite clear. Were actually producing dutiful actions a our desideratum, then we would be concerned with the reliability of a motive in doing this.<sup>28</sup> But in fact this is *not* Kant's desideratum. For he does not think, as a consequentialist does, that our aim should be to maximize anything, even the performance dutiful actions. Indeed, as we have already seen, it is not

at all obvious that the motive of duty has any advantage over other motives in terms of reliability. Moreover, whether it has such an advantage is an empirical question. It would only be on the basis of our experience with various kinds of motives, surely not on the basis of an *a priori* investigation, that we could pronounce them unfit. So it is quite implausible to suppose that Kant's desideratum was to find a motive that reliably produces right actions. Moreover, the reliability interpretation could only attribute an *instrumental*, and so *extrinsic*, value to the motive of duty--its value as a producer of dutiful actions. Now surely if anything is central to Kant's account, it is that morally worthy actions express the goodness of the good will, an *intrinsic* good.<sup>29</sup> So reliability really could never have been the proper interpretation of C if it is to have any role in determining whether an action has moral worth--that is, whether it expresses an intrinsically valuable will.

What is the point, then, of showing that the motive of duty alone is internally connected to the dutifulness of one's actions? I think that now we can see that its point is simply this: When there is an internal connection between an agent's motive in performing a particular kind of action and the particular kind of action that she in fact performs, then that action *expresses her motive*. And since her will is in part her motives, that action equally *expresses her will*. Thus, a helping action, a kind of altruistic action, expresses an altruistic will only when an altruistic motive, or a motive to perform altruistic kinds of actions, led to it. Likewise, a harming action, a kind of malicious action, expresses a malicious will only when a malicious motive, or a motive to perform malicious kinds of actions, led to it. In such cases, there is an internal connection between the agent's motive and the kind of action she performs. Now what a good will expresses

in her dutiful actions is the motive to act morally--or a motive to perform dutiful kinds of actions. Thus, a dutiful action expresses a good will only when the motive of duty led to it--when there is an internal connection between the agent's motive and the dutifulness of her action. And this is simply because a good will is a will motivated by duty.

Of course, in his examples, Kant excludes motives of pursuing self-preservation, others' well-being, and our happiness as moral motives.<sup>30</sup> But he leaves room for advocates of any alternative theory of what makes actions duties to respond. Thus, the sympathy theorist can claim that a moral agent as set out in her theory expresses a motive to act dutifully *if* we assume her account of what makes an action a duty. In other words, if a good will is a sympathetic will, and so its principle (i.e., "duty") is "to perform the sympathetic action", then the sympathetic actions to which the motive of "duty" (i.e., sympathy) leads would be of moral worth. They would be because they would express a "good will", or the motive of "duty", as defined by the sympathy theorist.

The rationale for the internal connection interpretation is therefore this: If our motive leads to the morally right action by accident, then our action does not express the motive to do what is morally right. And if our action does not express this motive, then it does not express a good will. But since the only unqualified and intrinsic good is a good will, then our action does not express what is unqualifiedly and intrinsically good. If moral goodness is simply unqualified and intrinsic goodness, then such an action does not express the only thing that is morally good. Hence, a morally worthy action is simply an action that expresses what is morally good--a good will. So *a morally worthy action must have a motive that is internally connected with the dutifulness of one's action because only then does it express a good will, a will to do what is dutiful.* And this will be so

*regardless* of how reliable this motive is in producing dutiful actions. If this is correct, then a full defense of Kant's account of moral goodness crucially hangs on defending his views that the only unqualified or moral good is a will which consists of the motive to do what is moral or dutiful, and in turn that what defines what is moral or dutiful are the *rational* principles of conduct argued for later in the *Groundwork*. Whether the motive of duty *as Kant understands it* is unique depends on whether duty, the principle of a good will, is indeed the CI.

#### V. THE ROLE OF C IN GROUNDWORK I

Kant's argument in *Groundwork* I obviously deserves a fuller treatment than I can give it here. Nevertheless, I wish to highlight how this account of Kant's position on the motive of duty fits into it. Kant's claims about the worth of acting from duty come mostly before his so-called "second proposition" which marks a new step in the argument to the CI--the account of what duty is.<sup>31</sup> Before the second proposition, Kant is arguing without the benefit of this account. Indeed, it is crucial to notice that in the examples prior to the second proposition duty is left unanalyzed. Thus, his position here is simply this: Only actions from duty (undefined) lead non-accidentally to dutiful actions (undefined). Any other ground for doing one's duty will be in accidental accord with it. Now this piece of our "ordinary moral outlook" does require that a good will express itself only in actions motivated by a conception of them as duties, or as morally required.<sup>32</sup> But what counts as such a conception is a matter left for the move toward a more "philosophical" outlook which comes with the analysis at the second and third propositions. So C, coming before this move, does not rule out any proposals for the nature of duty.

Such proposals are supposed to be ruled out by the second and third propositions. At the second proposition, Kant argues that, since the “material” of the maxim of duty, its end or purpose, cannot be the source of the unqualified value of the will it motivates, it must be its form.<sup>33</sup> For the maxim of someone motivated by duty is to do her duty no matter with what end it conflicts. Serving these ends cannot, then, be the source of its value, since it may well thwart them. Kant then argues at the third proposition that the form of moral demands is the form of a law. Thus, acting from duty is acting from an interest in the form, rather than material, of this maxim, and this is acting from an interest in its lawfulness--that is, from respect.

So Kant concludes *before* the second proposition that self-preservation, sympathy and happiness are all in accidental accord with duty. Dutiful actions performed from them have no moral worth (failing, as they do, to express the motive of duty, and so a good will). Yet someone could respond, before the second proposition, by giving content to the idea of duty in terms of one of these interests. The response would take this form: “But X is not in accidental accord with duty, because, in my view, our duty is to do whatever X indicates”--where “X” is some interest. In that case, the left over interests will provide ends our maxim may serve, but the fulfilling of which will be in accidental accord with “duty” as it is now defined. So one can *agree* that the principle of a good will is duty, and so only actions from duty express a good will, yet *disagree* with Kant's account of duty. For Kant's response must look forward to the argument after the second proposition, not backward to C.

Suppose, for instance, someone holds that duty is “whatever sympathy advocates”. A good will would still be one that decides its actions based on their dutifulness. In

deciding them on this basis, it is determining whether sympathy advocates them. It would then be no accident that sympathy leads a person to perform dutiful actions. How could it be? An action is a duty if and only if her sympathy advocates it. Kant's response to this account would not--*could* not--be that sympathy only accidentally leads to dutiful actions, that its maxim has no moral content. It would be, rather, an argument that sympathy cannot be definitive of one's duty, because it does not advocate actions in the form of a law.

Kant's argument, then, is roughly this: He begins by asking, What is the motivating principle of a good will? If a will that is unqualifiedly good is one good under any circumstance, and this is a will that wills moral actions under any circumstance, then we are directly before the issue which C proposes to settle. For we think that more is required for expressing a morally good will than simply willing a dutiful action in a particular circumstance. The will that willed a dutiful action in *this* circumstance must be such that it would have willed the dutiful action under *any* circumstance. It must, in other words, be no accident that *this* will willed *this* dutiful action. And only if the principle of this will is the principle of doing one's duty will the action it motivates express unconditionally good willing.

This then raises the issues of the second and third propositions. Since the good will is intrinsically good, its principle cannot be one which makes it only of extrinsic value--as it would be were the principle simply a recipe for producing this or that end. Its value must come from its willing in the form of a law. And that form, Kant argues, the form of lawfulness, is "universality". Therefore, the principle of an unconditionally good will is simply the principle of willing in a universal form--that is, the CI.<sup>34</sup>

Thus, even though Kant sometimes seemed to claim that reason's power to motivate was colossal, having a capacity to trump any other sensuous interest, and so an unsurpassed reliability in producing moral actions, this is not the issue he is pressing in the discussion of moral worth.<sup>35</sup> Such a point would not help him to arrive at a statement of the supreme principle of morality, since the reliability of a principle, as motive, is irrelevant to its conferral of intrinsic value on the actions it motivates. The discussion of moral worth simply leads us to the doorstep of the more important considerations of what kind of practical principle could play the role of duty.

Whether Kant's argument succeeds in establishing that duty is a purely rational demand is of course wide of my interest here. But notice that, so far as C is concerned, those who think morality is not a demand of reason not only have no ground for complaint, but can cheerfully agree with Kant's claims before the second proposition. For whatever the content of a moral will, it is both a plausible and agreeable doctrine to hold that morally good actions express such that will, be it sympathetic or purely rational.

## *VI. CONCLUSION*

Surprisingly, one can agree with Kant that the principle of a good will is duty and that only actions from duty have moral worth, yet reject Kant's account of duty. Indeed, Kant's rejection of motives other than duty leaves on one side all controversy over whether reason or sentiment grounds moral demands, and to that degree at least, some champions of sentiment who may bristle at this have mistaken the target. The real battle is over what account of moral demands we give; little hangs on whether acting from an interest in them is morally good. C simply follows from accepting the idea that judgments

of moral goodness focus on an agent's motives or will, offering a plausible account of the conditions under which the content of one's will is expressed in one's actions.

Kant eloquently claimed that a good will "would still shine like a jewel for its own sake as something which has its full value in itself" even if it is "entirely lacking in power to carry out its intentions".<sup>36</sup> My account has taken Kant at his word. The fact that a will might not have prevailed in leading to the right action under more difficult circumstances does not in the least diminish its goodness. That it motivated the agent to do her duty is enough for her action to express that goodness of will. This of course does not imply that reliability is unimportant. It implies only that reliability is a separate concern.

## NOTES

1. An ancestor of this paper was improved with the help of those who attended its presentation at a session of the 1993 Eastern Division Meeting of the American Philosophical Association in Atlanta, Georgia. I am especially grateful for the comments of Christine Korsgaard, which helped me to clarify my views, especially on the argument of *Groundwork I*. Others whose comments and suggestions have improved this paper include Simon Blackburn, David Cummisky, Thomas E. Hill, Jr., Richard McCarty, Peter Markie, Andrews Reath, Geoffrey Sayre McCord, Keith Simmons, Alexander von Schönborn, and Arnulf Zweig. Work on this paper was supported by a 1994 Research Council Summer Research Fellowship from the University of Missouri, Columbia.

2. G 4:390. I use H.J. Paton's translation of the *Groundwork of the Metaphysics of Morals* (G), (New York: Harper & Row, 1964); L.W. Beck's translation of the *Critique of Practical Reason* (CPrR), (Indianapolis: Bobbs-Merrill, 1956); M. Gregor's translation of the *Metaphysics of Morals* (MM), (New York: Cambridge University Press, 1991); and the T.M. Green & H.H. Hudson translation of *Religion within the Limits of Reason Alone*. (R) (New York: Harper & Row, 1934). All page numbers refer to the volume and page numbers from the Prussian Academy edition, except the *Religion*, which refer to the translator's page numbers.

3. For instance, see G 4:398 (quoted below); G 4:411: "Mixed moral philosophy, compounded of feeling and inclination and at the same time rational concepts...can only guide us by mere accident to the good" (although Kant here identifies duty with a *rational concept*, for reasons below, I take it that it is not its rationality that is required for a non-accidental relationship between motive and dutifulness of action); R 26: "For when

incentives other than the law itself (such as ambition, self-love in general, yes, even a kindly instinct such as sympathy) are necessary to determine the will to conduct conformable to the law it is merely accidental that these coincide with the law"; CPrR 5:83: "No other subjective principle must be assumed as incentive, for thought it might happen that the action occurs as the law prescribes, and thus in accord with duty but not from duty, the intention to do the action would not be moral".

4. In this, I am following Barbara Herman in "On the Value of Acting from a Motive of Duty", *The Philosophical Review* XC, 3 (July 1981): 359-382.

5. See Williams, Bernard, "Morality and the Emotions" in *Problems of the Self*. (New York: Cambridge University Press, 1973); and Blum, Lawrence. *Friendship, Altruism and Morality*. (Boston: Routledge & Kegan Paul Ltd., 1980).

6. See Smith, Holly, "Varieties of Moral Worth and Moral Credit", *Ethics* 101 (1991): 279-303.

7. Op. cit. Herman.

8. G 4:398

9. There are other ways, however, that one might make the distinction. For instance, we might distinguish a motive's intentional object from its strength as well. All that is required is that there be some discernible cognitive or semantic dimension to motives.

10. Allison provides an illuminating overview of the most prominent interpretations of Kant's account of maxims in op. cit. Allison, 1990: 85-94.

11. For a recent example of this interpretation, see Paul Guyer, *Kant and the Experience of Freedom*. Cambridge: Cambridge University Press, 1993; pp. 337-355.

12. See op. cit., Blum, Chapter 2, for an extensive discussion of various kinds of unreliability; also op. cit., Williams, p. 228.
13. Op. Cit. Herman. Actually, she believes that “the moral fault with the profit motive is that it is unreliable”, and that a moral motive should “guarantee that the right action will be done”, but that this is only part of the story. (op. cit., Herman, p. 363) Later, she appears to reject reliability altogether, however. (pp. 368-71)
14. Ibid., p 365.
15. Specifying (iii) is also in line with Kant’s setting aside of actions contrary to duty because they raise no issue as to whether they could have been performed from duty or not. (G 4:397)
16. Op. cit., Blum p. 34.
17. Op. cit., Smith, p. 290
18. Op. cit., Williams, pp. 227-8.
19. Op. cit., Herman, pp. 369-70.
20. Ibid.
21. G 4:390
22. This is one of Paul Benson’s main points in “Moral Worth”, *Philosophical Studies* 51, 1987, pp. 365-82.
23. Blum makes little of this same sort of distinction: “it is true that in particular cases acting from sympathy or compassion will not lead to the promotion of beneficence ...this does not mean that the connection between the sympathy (as a motive) and the beneficence is merely an accidental or contingent one” (op. cit., Blum, p. 30). But in order to make good the claim that the connection between sympathy and dutiful actions is

not contingent, he would have to argue--dubiously, I would think--that the object of sympathy is necessarily moral.

24. See, e.g., G 4:407-8; MM 6:392-3

25. G 4: 394.

26. MM 6:408

27. MM 6: 405

28. Herman makes a similar point. (op. cit., Herman, p. 366). The problem is that she does not go to say what point Kant has in applying C that he does.

29. G 4:393-395

30. G 4:397-9

31. G 4:399

32. In her comment on my paper, Christine Korsgaard noted that Kant thought that, in viewing our actions as duties, we must view them as lawful. My point is that Kant argues for this position, and that argument comes after, not before, the second proposition.

33. G 4:400

34. This move in the argument is perhaps the most problematic. See, e.g., Bruce Aune's *Kant's Theory of Morals* (Princeton: Princeton University Press, 1979): 29-30, 86-9; and Henry Allison, "On a Presumed Gap in the Derivation of the Categorical Imperative". *Philosophical Topics*, Vol. 19 No. 1 (1991): 1-15.

35. See, e.g., G 4:410.

36. G 4:394