

## Ridgeview Publishing Company

---

Creating the Kingdom of Ends: Reciprocity and Responsibility in Personal Relations

Author(s): Christine M. Korsgaard

Reviewed work(s):

Source: *Philosophical Perspectives*, Vol. 6, Ethics (1992), pp. 305-332

Published by: [Ridgeview Publishing Company](#)

Stable URL: <http://www.jstor.org/stable/2214250>

Accessed: 25/12/2012 22:43

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



Ridgeview Publishing Company is collaborating with JSTOR to digitize, preserve and extend access to *Philosophical Perspectives*.

<http://www.jstor.org>

# CREATING THE KINGDOM OF ENDS: RECIPROCITY AND RESPONSIBILITY IN PERSONAL RELATIONS<sup>1</sup>

Christine M. Korsgaard  
Harvard University

As the virtuous man is to himself, he is to his friend also, for his friend is another self.

Aristotle<sup>2</sup>

When we hold a person responsible, we regard her as answerable for her actions, reactions, and attitudes. We use the concept of responsibility in two contexts, the legal and the personal. We use it in the legal context when we must determine whether to punish someone for a crime or make him liable for another's losses. We use it in the context of everyday personal interaction, when we are pressed to decide what attitude we will take toward another, or toward some action or reaction of another. It is frequently assumed that these two uses are the same or at least continuous. Because I have doubts about this, and some worries about the appropriateness of using the notion in the legal context, I want to lay that use aside.<sup>3</sup> In this paper, my focus will be on our practice of holding people responsible in the context of personal relations.

I begin by offering an account of personal relations, derived from Kant and Aristotle, along with an explanation of why they require us to hold one another responsible. I then distinguish two views about what holding someone responsible involves. Specifically, I argue that to hold someone responsible is to adopt an attitude towards him rather to have a belief about him or about the conditions under which he acts. This view gives rise to a problem: if holding someone responsible is something that we *do*, why and how do we decide to do it? In the rest of the paper, I argue that Kant's theory of personal and moral relations provides some answers to this question.

## I. Personal Relations, Reciprocity, and Responsibility

In the British Empiricist tradition, the concept of responsibility has been closely associated with the ideas of praise and blame, and these in turn have

played a central role in its moral philosophy. In the theories of Hutcheson, Hume, and Smith, the approval and disapproval of others is the fundamental moral phenomenon, from which all our moral ideas spring.<sup>4</sup> There is something obviously unattractive about taking the assessment of others as the starting point in moral philosophy. One of the appealing things about Kant's ethics, by contrast, is that in it moral thought is seen as arising from the perspective of the agent who is deciding what to do. Responsibility is in the first instance something taken rather than something assigned. And this fact about the structure of his view is complemented by a fact about its content. Kant is not very interested in praise and blame and seldom mentions them. And when he does discuss issues of moral assessment, much of what he says favors a taking a generous attitude. His metaphysical view that we cannot know even our own most fundamental maxims (G 407/19) combines with a set of moral injunctions—to respect others, avoid scandal, and “never to deny the wrongdoer all moral worth” (MMV 462-464/127-129)—to give philosophical foundations to the Biblical injunction “Judge not.”<sup>5</sup>

But in a broader sense it is not possible for us to avoid holding one another responsible.<sup>6</sup> For holding one another responsible is the distinctive element in the relation of adult human beings. To hold someone responsible is to regard her as a *person*—that is to say, as a free and equal person, capable of acting both rationally and morally. It is therefore to regard her as someone with whom you can enter the kind of relation that is possible only among free and equal rational people: a relation of reciprocity. When you hold someone responsible, you are prepared to exchange lawless individual activity for reciprocity in some or all of its forms. You are prepared to accept promises, offer confidences, exchange vows, cooperate on a project, enter a social contract, have a conversation, make love, be friends, or get married. You are willing to deal with her on the basis of the expectation that each of you will act from a certain view of the other: that you each have your reasons which are to be respected, and your ends which are to be valued. Abandoning the state of nature and so relinquishing force and guile, you are ready to share, to trust, and generally speaking to risk your happiness or success on the hope that she will turn out to be human.

I borrow the idea that personal relations are characterized by reciprocity from both Kant and Aristotle, two of the very few philosophers in our tradition who have written about this topic. And it will be important to my argument that I hold along with them that the territory of personal relations is continuous with moral territory. That is to say, I accept their view that the forms of friendship, at their best at least, are forms of the basic moral relation among human beings—particular forms of that relation which have been rendered perfect of their kind. Aristotle holds that the most perfect human relation is the friendship of virtue, in which two people of good character share their lives and activities, and in particular, share those virtuous activities that make

their lives worth living. (NE IX.9 1169b28ff./1089-1090) And Kant holds that the ideal of friendship is that of “the union of two persons through equal mutual love and respect,” a relation in which the two basic attitudes we owe to one other as moral beings are realized in spontaneous natural sentiment (MMV 469/135). Characteristically, Aristotle holds that achieving such a relationship is a virtue, and Kant, that striving to achieve it is a duty. For friendship, Aristotle tells us, it is “not only necessary but noble” (NE VIII.2 1155a29-31/1059); and Kant echoes the thought: “friendship...is no ordinary duty but rather an honorable one proposed by reason.” (MMV 469/135)

Both define this perfect relation, as well as the less perfect variants of it, in terms of reciprocity, and both cite reciprocity as the reason why friendship is found above all among people who are good. For Aristotle, friendship is characterized by acknowledged reciprocal good will, in which each person loves the other for his own (the other’s) sake. (NE VIII.2 1155b28-1156a5/1059-1060) This requires trust in the other’s goodness, for as Aristotle says “it is among good men that trust and the feeling that ‘he would never wrong me’ and all the other things that are demanded in true friendship are found.” (NE VIII.3 1157a22-24/1062) Kant characterizes friendship in the *Lectures on Ethics* as “the maximum reciprocity of love” (LE 202). There he argues that friends exchange their private projects of pursuing their own happiness, each undertaking to care for the other’s happiness instead of his own. “I, from generosity, look after his happiness and he similarly looks after mine; I do not throw away my happiness, but surrender it to his keeping, and he in turn surrenders into my hands” (LE 203). This requires the maximum reciprocity of love because “if I am to love him as I love myself I must be sure that he will love me as he loves himself, in which case he restores to me that with which I part and I come back to myself again” (LE 202). The later account in *The Metaphysics of Morals* adds another element. Friendship in its perfection involves what Kant calls “the most intimate union of love with respect.” (MMV 469/135) While love moves you to pursue the ends of another, respect reminds you that she must determine what those ends are; while love moves you to care for the happiness of another, respect demands that you care for her character too. Kant means here the *feelings* of love and respect, for he is defining the friendship of sentiment, but this does not sever the tie to morality. Love and respect are the primary duties of virtue we owe to others. Although only the outward practices can be required of us, Kant makes it clear in many passages that he believes that in the state of realized virtue these feelings will be present. In one place he even defines love and respect as the feelings which accompany the exercise of our duties towards others (MMV 448/112; see also R 23-24n /19n). Feelings of sympathy, gratitude, and delight in the happiness of others are not directly incumbent upon us, but they are the natural result of making the ends of others our own, as duty demands. The feeling of respect, a still higher achievement, is the natural

result of keeping the humanity of others and so their capacity for good will always before our eyes. So this kind of friendship really is in Kant's eyes the friendship of *virtue*, the moral relation in a perfected form.

"When men are friends they have no need of justice," says Aristotle, and there are two ways to understand what he means. (NE VIII.1 1155a25-26/1059) The wrong way is to suppose that he is referring to an idea like Hume's of the "circumstances of justice": justice is only useful and so is only required when moderate scarcity holds among people who are only moderately benevolent.<sup>7</sup> Friends, because they are endlessly benevolent to each other, are not in the circumstances of justice and have no use for it. Now this clearly cannot be Aristotle's meaning, for he thinks that "the truest form of justice is thought to be a friendly quality" (NE VIII.1 1155a27-28/1059) and that "friendship and justice...seem to be concerned with the same objects and exhibited between the same persons" (NE VIII.9 1159b25-27/1068). Justice is, at its best, a kind of civic friendship. And indeed, friendship, like justice, is not primarily a matter of doing things *for* one another, but of doing things together. "Those in the prime of life it stimulates to noble actions—"two going together"—for with friends men are more able both to think and to act." (NE VIII.1 1155a14-16/1058) Aristotle sums up his account with these words:

And whatever existence means for each class of men, whatever it is for whose sake they value life, in *that* they wish to occupy themselves with their friends, and so some drink together, others dice together, others join in athletic exercises and hunting, and in the study of philosophy, each class spending their days together in whatever they love most in life; for since they wish to live with their friends, they do and share in those things which give them the sense of living together. (NE IX.11 1172a2-9/1093)

Justice isn't necessary between friends because the reciprocity (NE V.5-6) and unanimity (NE VIII.1; NE IX.6) characteristic of justice are already present. And this is because they want above all to act together. Kant would again agree. Kant thinks that justice is reciprocal coercion under a general will, made necessary by geographical and economic association (MMJ 232/36-37; 256/64-65). When we share a territory we may have a dispute about rights. But I may enforce my rights against you only on the understanding that you may enforce your rights against me, and in this way we make a social contract and constitute ourselves a state. (MMJ 315-316/80-81) Friendship is a free and uninstitutionalized form of justice, where the association is created by love rather than geographical necessity, and regulated by mutual respect rather than reciprocal coercion.

But it is not merely the narrow relation of political justice, but rather the moral relation generally, that friendship mirrors. For to join with others as citizens in the Kingdom of Ends is to extend to our inner attitudes and personal choices the kind of reciprocity that characterizes our outer actions in the political state. This is seen best in the way Kant uses the Formula of Humanity

to explain our duties to others.<sup>8</sup> In the positive sense, to treat another as an end in itself is to make her ends your own: "For the ends of any subject who is an end in himself must as far as possible be my ends also, if that conception of an end in itself is to have its full effect in me." (G 430/37) In the negative sense, to treat another as an end in itself is to respect her autonomy—to leave her actions, decisions, and ends to her own choice. But this respect gets its most positive and characteristic expression at precisely the moments when we must act together. Then another's right to choose becomes the "limiting condition" of my own. (G 431/37) If my end requires your act for its achievement, then I must let you make it your end too. Both what I choose and the way I choose it must reflect this constraint. You must be free to choose whether you will contribute to the success of my project or not. Kant says anyone engaged in a transaction with me must be able to agree with my way of acting towards him and to share in the end of my action. (G 430/37) If I force you to contribute to an end you have had no opportunity to decide for or against, or if I trick you into contributing into one end under the guise of soliciting your help with another, then I have used you as a mere means. Kant illustrates this with the example of the lying promise. If I ask you to lend me money, knowing I shall not be able to pay you back, I trick you into contributing to an end you have had no opportunity to choose. I make you think that the end produced by our transaction is my temporary use of your money, when in fact it is my permanent possession of it. Neither my way of acting nor the end produced by it are things that you are in a position to accept or reject, and this renders them morally wrong. Thus I must make your ends and reasons mine, and I must choose mine in such a way that they can be yours. But this just is reciprocity. Generalized to the Kingdom of Ends, my own ends must be the possible objects of universal legislation, subject to the vote of all. And this is how I realize my autonomy. Paradoxically if you like, my ends and actions are most truly my own when they are chosen under the restrictions of a possible reciprocal relation—a kind of friendship—with everyone.

I do not say this to join forces with those who believe that there could be no room in a Kantian life for personal as opposed to moral relations.<sup>9</sup> Nor, certainly, do I mean to suggest that being friends is just a matter of being good. My point is only that moral and personal relations are not different in kind. The difference between them is the difference between the degree of reciprocity that is *required* of us as one human being relating to another, and the degree of reciprocity that we are capable of when our relations are at their best. Anyone must tell the truth when the circumstances call for it, but between friends there is a presumption of intimacy, frankness, and confidence. Anyone must help another in need or emergency, but friends promote each other's projects as routinely as they do their own. Anyone must refrain from leading others into temptation; but friends help each other to

be good. The difference is the difference between the absolute moral requirements we must meet if human relations are to be decent at all, and the further reaches of positive virtue, where our relations with one another become morally worthy. Friendships are human moral achievements that are lovely in themselves and testify to the virtue of those who sustain them. To become friends is to create a neighborhood where the Kingdom of Ends is real.<sup>10</sup>

Kant's faith in the moral force of reciprocity shows up best when he believes that the basic moral relation is at risk. In both the *Lectures on Ethics* and the *Metaphysics of Morals* Kant gives inarticulate voice to the view that there is something morally troublesome, even potentially degrading, about sexual relations. It is important to understand that what bothers him is *not* the idea that one is using another person as a means to one's own pleasure. That would be an incorrect view of sexual relations, and in any case any difficulty about it, would, by Kant's own theory, be alleviated by the other's simple act of free consent. What bothers Kant is rather that sexual desire takes a *person* for its object.<sup>11</sup> He says: "They themselves, and not their work and services, are its Objects of enjoyment." (LE 162) And he continues:

Man can, of course, use another human being as an instrument for his service, he can use his hands, his feet, and even all his powers, he can use him for his own purposes with the other's consent. But there is no way in which a human being can be made an Object of indulgence for another except through sexual impulse: it is an appetite for another human being (LE 163)

Regarding someone as a sexual object is not like regarding him as an instrument or a tool, but more like regarding him as an aesthetic object. But in this case the attitude is not just appreciation but desire. (MMV 426/87)<sup>12</sup> Viewed through the eyes of sexual desire another person is seen as something wantable, desirable, and, therefore, inevitably, possessable.<sup>13</sup> To yield to *that* desire, to the extent it is really *that* desire you yield to, is to allow yourself to be possessed. The problem is how you can do that in a way that is consistent with respect for your own humanity.<sup>14</sup> And the solution rests in reciprocity:

If, then, one yields one's person, body and soul, for good and ill in every respect, so that the other has complete rights over it, and if the other does not similarly yield himself in return and does not extend in return the same rights and privileges, the arrangement is one-sided. But if I yield myself completely to another and obtain the person of the other in return, I win myself back, I have given myself up as the property of another, but in turn I take that other as my property, and so win myself back again in winning the person whose property I have become. In this way the two persons become a unity of will (LE 167)

The language of self-surrender and retrieval here is strikingly similar to that Kant uses elsewhere for both friendship and justice. In making the social contract, Kant says, we do not sacrifice part of our freedom for a particular purpose, but rather sacrifice all of our lawless freedom in order to regain

our freedom again, undiminished, under law. (MMJ 316/80-81) In the case of friendship Kant says I surrender my happiness completely into the hands of my friend, but that in loving me as he loves himself “he restores to me that with which I part and I come back to myself again” (LE 202).<sup>15</sup> This perfect reciprocity is the only condition under which the sexual relation is morally legitimate; and Kant thinks this condition is only possible in marriage, where the reciprocity of surrender has been pledged. Extramarital sex is forbidden only because the woman, as Kant supposes, does not then have the same rights over the man that he has over her. Of course marriage as it has usually existed has hardly been a solution to *this* problem. The equality necessary for reciprocity is far more likely to be distanced even further by marriage, which has usually given the husband rights over his wife *additional* to those that accrue from the superior social position he has held as a man. Kant admits as much in the *Metaphysical Elements of Justice*, asserting that an unequal marriage is not a marriage in his sense at all. Thus marriage as it has been practiced in most societies has not sanctified but rather degraded sexual relations. (MMJ 278-279)<sup>16</sup> But perhaps the most startling ramification of Kant’s view emerges in what he says about incest. As strong as our natural aversion to it may be, and however risky and therefore *conditionally* wrong it is from a reproductive point of view, incest is only morally wrong in itself, *unconditionally*, in one case: the case of parent and child. And this is because, according to Kant, the equality of respect required for reciprocity cannot and should not be achieved in that relation. (LE 168)

Which brings me back to my topic. The relations of reciprocity are relations that obtain between free and equal persons. As such, they call for mutual responsibility for two important reasons. In order to make the ends and reasons of another your own, you must regard her as a source of value, someone whose choices confer worth upon their objects, and who has the right to decide on her own actions. In order to entrust your own ends and reasons to another’s care, you must suppose that she regards you that way, and is prepared to act accordingly. People who enter into relations of reciprocity must be prepared to share their ends and reasons; to hold them jointly; and to act together. Reciprocity is the sharing of reasons, and you will enter into it only with someone you expect to deal with reasons in a rational way. In this sense, reciprocity requires that you hold the other responsible.

It is certainly a concomitant of holding someone responsible that you are prepared for blame, resentment, and the other reactive attitudes.<sup>17</sup> If my friend fails me in a serious way and I do not blame her, shrugging it off as I would the misdemeanors of a child or a pet, then I was not holding her responsible after all, and probably I was holding myself back. But it is a mistake to make these reactions central. Blame is important, not as a tool of training or the enforcement of social norms, but as an expression of the tenacity of

disappointed respect. At its best, it declares to its object a greater faith than she has in herself. Yet still it is not central. The willingness to take a chance on some form of reciprocity is the essence of holding someone responsible.

I mean in these words both to acknowledge the affinity of my position with P.F. Strawson's in "Freedom and Resentment" and to notice one point of difference. Strawson also emphasizes the employment of the concept of responsibility in everyday personal relations. But he tends to focus more on the effect of attributions of responsibility on our sentiments than their effect on our practices. His topic, as he describes it, is "the non-detached attitudes and reactions of people directly involved in transactions with each other;...the attitudes and reactions of offended parties and beneficiaries; of such things as gratitude, resentment, forgiveness, love, and hurt feelings."<sup>18</sup> I want to focus less upon the exchange of benefits and harms, and the feelings that result from that exchange, and more upon the willingness to act in concert. But my point is similar to his. In everyday personal interaction, we cannot get on without the concept of responsibility. And therefore we cannot rest with the view that agents take responsibility for their own actions but can refrain from judging others. For a Kantian, this means it is necessary to say more than Kant himself did about what, on his view, is involved in determining when and whether to hold people responsible.

## II. Theoretical and Practical Conceptions of Responsibility

Attributions of responsibility may be understood in either of two ways, which I will call theoretical and practical. Construed theoretically, responsibility is a characteristic of persons. Construed practically, holding one another responsible is something that we do, the more or less deliberate adoption of an attitude. In what follows I will distinguish these two ways of understanding attributions of responsibility, and show that according to Kant we must understand attributions of responsibility in a practical way. I believe that this view of responsibility is implicit in our actual practices, and therefore that, on this point at least, Kant's account can make us more transparent to ourselves.

Responsibility is construed theoretically by those who think that it is a fact about a person that she is responsible for a particular action, or that there is some fact about her condition either at the time of action or during the events which led up to it which fully determines whether it is correct to hold her responsible. It is a fact, say, that she could have done otherwise, or that she could have avoided the condition which made it impossible for her to do otherwise. Similar although somewhat more complicated claims would be made about the person's reactions and attitudes: facts about the person settle the question whether she is accountable for them. Deciding whether to hold someone responsible is a matter of assessing the facts; it is a matter

of arriving at a belief about her. It seems probable that we arrive at this model by a certain route: we think about legal responsibility first, and we suppose that in that case we *must* find facts which can settle the matter, and then we imagine that personal responsibility is an extension of this.

Responsibility is construed practically by those who think that holding someone responsible is adopting an attitude towards her, or, much better, placing yourself in a relationship with her. While of course facts about the agent and about her condition at the time of the action *guide* your decision whether to hold her responsible, they do not fully *determine* it. It is important to see that the facts still do provide guidance, for a practical conception need not be envisioned as completely voluntaristic. On either a theoretical or a practical conception, we will, when deciding whether to hold someone responsible, say such things as "he is very nervous about the interview he has tomorrow" or "he's been hurt so often that now he can never trust a woman." But in a practical conception these considerations appear in the role of practical reasons for not holding the person responsible rather than as evidence that he could not have helped what he did. When responsibility is viewed this way, we need not suppose that there is a *fixed degree* of nervousness or past heartbreak beyond which someone is *in fact* no longer responsible for the way he acts and reacts; deciding whether to hold him responsible is therefore not a matter of determining whether this fixed degree has been reached. A resulting feature of the practical conception which I take to be one of its virtues is that it distances the question whether to hold someone responsible from the question whether he acted voluntarily. I do not believe there is a stable relationship between the voluntariness of an action or attitude and the appropriateness of holding someone responsible for it. If a bad action is found to have been involuntary in some straightforward way, we will withdraw blame, we may also do this if the person is under severe emotional stress. But there is neither need nor reason to reduce the second kind of excusing condition to the first and say that people under severe emotional stress *cannot* control themselves. We do not need to understand a form of debilitation as a form of impossibility in order to make allowances for it; we need only to know what it is like. Conversely, we may well blame people for involuntary attitudes or expressions, because we blame people for lack of control itself. If you cannot repress a victorious grin on learning that your rival has met with a gruesome accident, you ought to be blamed, precisely on that account. The impulse to reduce all excusing conditions to claims about the voluntary comes from the theoretical conception of responsibility, which demands an answer to the question whether one could have done otherwise or not. On the practical conception excuses need not completely determine our decisions about whether to hold people responsible. If the decision to attribute responsibility is practical, it may be reasonable to make it partly on the basis of other kinds of considerations: in particular,

which reciprocal relations you already stand in or plan to stand in or hope to stand in to the person in question.

Construing responsibility practically opens up possibilities that would not make sense if responsibility were a fact about the person. It is because we both accept and avail ourselves of these possibilities that I claim that we implicitly understand attributions of responsibility practically in everyday life. For instance, it may be perfectly reasonable for me to hold someone responsible for an attitude or an action, while at the same time acknowledging that it is just as reasonable for someone else not to hold the same person responsible for the very same attitude or action. Perhaps it is reasonable for *you* to forgive or overlook our friend's distrustful behavior on the grounds that he has suffered so much heartbreak, but not for me, *not* because I fail to appreciate how hurt he has been, but because I am the woman whose loving conduct is always met with distrust.<sup>19</sup> Again, if deciding whether to hold someone responsible is something that we do, it is something that we may in turn be held responsible for. Holding someone responsible can be insensitive or merciless; failing to hold someone responsible can be disrespectful or patronizing. Moral requirements will apply to our attributions of responsibility, just as Kant believes they do.

Consider, for instance, the appropriate reaction to a case where one is disappointed in friendship. Kant thinks the perfect friendship I described earlier, characterized by *feelings* of equal mutual love and respect, is impossible to achieve. But he does think we can achieve what he calls "moral friendship." (MMV 471-473/138-140) The form of reciprocity central to this relation is the frank conversation, the sharing of sentiments, of which Kant believes we all stand in need. Like other reciprocal relations it calls for good character on the part of the participants, because it is hedged with dangers—ranging from the crude risk that you will tell your secrets to an unreliable person who will publish them, to the more subtle risk that your confidences will be met with disrespectful attitudes.<sup>20</sup> I do not want to share my ambitions with someone who is inwardly amused by my vanity, nor whisper my temptations to someone who will place a harsh construction on them. One who consents to receive my confidences is committed to avoiding the vices of mockery and calumny, serious failures of respect in the Kantian catalogue (MMV 466-467/131-133). And I will blame her if she fails in these ways, without regard to the available evidence of her character or of the circumstances in which it was formed. Her circumstances must have been very bad indeed, or her failures very frequent, before I may decide it was simply *my* error to trust her. For in deciding this I write her off as a person, and I do this at my own moral peril.

I suppose that most of us have at one time or another had the experience of being tempted to "write somebody off." The extent to which we do this is a matter of degree, and hopefully we do not go so far as to give up treating

the person with the most basic forms of moral decency. But we may avoid interaction, as far as possible; we may choose to execute our projects in the company of others; where interaction is necessary, we may come to treat the person as an obstacle to be worked around. In an extreme case we may cease to have reactive attitudes altogether, or at least we may scold ourselves, as for irrational feelings, when we have them. "You know that she always ends up infuriating you. Why don't you just stay out of her way?" Taking such attitudes towards others seems disrespectful, but it can certainly sometimes be tempting all the same. How do we decide what to do in such a case? On a theoretical construal of responsibility, we simply ask whether the person is in fact responsible for the offensive behavior, and treat her accordingly. On a practical construal, we must discover moral and practical reasons that will guide us to the right attitude. Kant's theory of moral and personal relations, I believe, can show us where these reasons are to be found.

### III. Kant's Two Standpoints

I will approach these issues in a roundabout way, however. I begin by discussing the way Kant reconciles free will and determinism, and by showing how his reconciliation gives rise to some apparent problems about holding people responsible. Kant's theory of moral and personal relations show us how he might have resolved one of these issues, and how we might resolve the other.

Kant's solution to the problem of freedom and determinism is clear enough in outline, however much philosophers may disagree about what it means. We must view ourselves from two standpoints, from which we appear as members of two different "worlds." (G 452/53-54) Complete causal determinism holds in the phenomenal or sensible world, the world of things as they appear to us; but we cannot know that it holds in the noumenal world, the world of things as they are in themselves. Indeed, since we must suppose that there are some undetermined first causes, or free agencies, which generate the appearances, we must suppose that things which exist in the noumenal world are free.<sup>21</sup> Insofar as we regard ourselves as "intelligences," the spontaneity of reason induces us to attribute a noumenal existence to ourselves. (G 452/53; C2 42-43/43-44) Insofar as we consider ourselves to be intelligent *agents*, then, we must regard ourselves as free: indeed, completely and transcendently so. Yet at the same time we must view our actions, like all phenomena, as fully determined.<sup>22</sup>

Despite Kant's strictures against trying to envision what occurs on the boundary between the two worlds, it is natural to want a picture that reconciles these two views of ourselves. At one point in the *Critique of Practical Reason*, Kant supplies the beginning of such a picture. He proposes that we should think of ourselves, and also that we do think of ourselves,

as if we created our own characters. Although a person may know that his actions are determined in the phenomenal world, Kant says:

.the same subject . is conscious also of his existence as a thing-in-itself...determinable only by laws which he gives to himself through reason In this existence nothing is antecedent to the determination of his will, every action, and ..even the entire history of his existence as a sensuous being, is seen...only as a consequence.. of his causality as a noumenon From this point of view, a rational being can rightly say of any unlawful action which he has done that he could have left it undone, even if as an appearance it . was inescapably necessary For this action and everything in the past which determined it belong to a single phenomenon of his character, which he himself creates.. (C2 97-98/101)

Kant then applies this picture to our attributions of responsibility:

From this point of view.. judgments may be justified which. seem at first glance to conflict with equity. There are cases in which men...have shown from childhood such depravity. that they are held to be born villains and incapable of any improvement of character; yet they are judged by their acts, they are reproached as guilty of their crimes, and, indeed, they themselves find these reproaches as well grounded as if they...were just as responsible as any other men This could not happen if we did not suppose that whatever arises from man's choice . has a free causality as its ground the vicious quality of the will ..is the consequence of. freely assumed evil and unchangeable principles. (C2 99-100/103)

Here one's life is regarded as the phenomenal representation or expression of a single choice, the choice of one's character or fundamental principle. This choice must be understood as occurring outside of time, in the noumenal world. The choice is the one described in the first book of *Religion Within the Limits of Reason Alone*: the choice of how incentives are to be ordered in one's most fundamental maxim, the choice between morality and self-love. (R 36/31) As Kant sees it, human beings are subject to certain incentives—impulses which present themselves to us as candidates, so to speak, to be reasons for action. Among these are our desires and inclinations, as well as respect for the moral law. Kant believes that we are not free to ignore such incentives altogether. Instead, our freedom consists in our ability to rank the incentives, to choose whether our self-love shall be governed by morality or morality shall be subordinated to self-love. This fundamental choice then governs our choice of lower-order maxims. The fundamental choice is an act—in the *Religion* Kant calls it an intelligible act—and it is ultimately this intelligible act that is imputable to us, and makes our phenomenal actions imputable to us. (R 31-32/26-27)

When first exposed to Kant's view, one may be tempted to try to picture how and where the choice of one's character enters the processes which ultimately issue in action. Suppose, with violent oversimplification, that it is a law of nature that children raised in certain conditions of poverty and

insecurity tend to become somewhat selfish as adults, and suppose that such a childhood has had this effect on Marilyn. Are we to say to her: "Your childhood insecurity gave you an incentive to be selfish, but it is still your own fault if you elevate that incentive into a reason?" Then we are thinking that Marilyn's freedom inserts itself in between the causes in her background and their ultimate effect.<sup>23</sup> Or are we supposed to think that, in her noumenal existence, Marilyn *wills to be* a selfish person? Or, to get even fancier, should we think that in her noumenal existence Marilyn wills the law of nature that deprived children become selfish adults? Obviously, if we try to picture *how* Marilyn's freedom is related to the forces that determine her, we must imagine it either inserting itself somewhere into the historical process, or standing behind the laws of nature from which this historical process necessarily follows. And both of these pictures seem crazy.<sup>24</sup>

And of course they are crazy. Kant's response to this problem is to maintain that the question should not be asked. To ask how freedom and determinism are related is to inquire into the relation between the noumenal and phenomenal worlds, a relation about which it is in principle impossible to know anything. But our understanding of what this response amounts to will depend on how we understand the distinction between the noumenal and phenomenal worlds, and the related distinction between the two standpoints from which Kant says we may view ourselves and our actions.

This is a large issue which I cannot treat here in a satisfactory way, I shall simply declare my allegiance. On a familiar but as I think misguided interpretation, the distinction between the two worlds is an ontological one; as if behind the beings of this world were another set of beings, which have an active and controlling relation to the beings of this world, but which are inaccessible to us because of the limits of experience. According to this view, we occupy both worlds, and viewing ourselves from the two standpoints we discover two different sets of laws which describe and explain our conduct in the two different worlds. We act on the moral law in the noumenal world, the law of self-love in the phenomenal world. This view gives rise to familiar paradoxes about how evil actions are even possible, and how we could ever be held responsible for them if they were.<sup>25</sup>

On what I take to be the correct interpretation, the distinction is not between two kinds of beings, but between the beings of this world insofar as they are authentically active and the same beings insofar as we are passively receptive to them. The "gap" in our knowledge exists not because of the *limits* of experience but because of its essential *nature*: to experience something is (in part) to be passively receptive to it, and therefore we cannot have experiences of activity as such.<sup>26</sup> As thinkers and choosers we must regard ourselves as active beings, even though we cannot *experience* ourselves as active beings, and so we place ourselves among the noumena, necessarily, whenever we think and act. According to this interpretation, the

laws of the phenomenal world are laws that describe and explain our behavior. But the laws of the noumenal world are laws which are *addressed to us* as active beings; their business is not to describe and explain at all, but to govern what we do.<sup>27</sup> Reason has two employments, theoretical and practical. We view ourselves as phenomena when we take on the theoretical task of describing and explaining our behavior; we view ourselves as noumena when our practical task is one of deciding what to do.<sup>28</sup> The two standpoints cannot be mixed because these two enterprises—explanation and decision—are mutually exclusive.<sup>29</sup>

These two ways of understanding the noumenal/phenomenal distinction yield very different interpretations of Kant's strictures against trying to picture the relation between the noumenal and phenomenal worlds. On the ontological view, the question how the two worlds are related is one which, frustratingly, cannot be answered. On the active/passive view, it is one which cannot coherently be *asked*. There is no question that is answered by my descriptions of how Marilyn's freedom interacts with the causal forces that determine her. For freedom is a concept with a practical employment, used in the choice and justification of action, not in explanation or prediction; while causality is a concept of theory, used to explain and predict actions but not to justify them.<sup>30</sup> There is no standpoint from which we are doing both of these things at once, and so there is no place from which to ask a question that includes both concepts in its answer.

So, if I am myself Marilyn, and I am trying to decide whether to do something selfish, reflections on the disadvantages of my background are irrelevant. I must act under the idea of freedom, and so I must act on what I regard as reasons. Being underprivileged may sometimes be a cause of selfish behavior, but it is not a reason that can be offered in support of it by a person engaged in it. So although we do not necessarily say of Marilyn: "her background gave her some tough incentives to deal with, but still it is up to her whether she treats them as reasons," that is what she must say to herself. I say that we do not *necessarily* say this, because, as I am about to argue, whether we say it depends on whether we have decided to enter into reciprocal relations with her and so to hold her responsible. But in that case, it is better regarded as something we say *not about* but *to* her. The second-person grammatical form, so rarely privileged in philosophy, is exactly right here, for if anyone besides Marilyn has the right to make this judgment, it is her friends, those with whom she interacts. On the other hand, if I am not Marilyn's friend but a social scientist who is trying to understand and explain her behavior, then my business is not to try to justify her conduct, and for my purposes the causal explanation which makes her selfish actions seem inevitable is the right one to pursue.

The two worlds, or the two views of the world we get from the two standpoints, may seem strangely incongruent, but it is important to see that

there is no contradiction. The incongruity simply follows from the fact that we stand in two very different relations to our actions: we must try to understand them, but we must also decide which ones to do.

#### **IV. Practical Grounds for Holding People Responsible**

But we cannot just leave the matter there. For there are contexts in which we have to mix considerations derived from the two standpoints, and make a moral assessment of someone's action, on the basis of a theoretical explanation of what she did. This occurs when we are making judgments about responsibility: when we must decide whether, for instance, someone is to be exonerated, excused, forgiven, blamed, or not held responsible for a bad action at all.

There are really two problems here. First, given that we can view people and their actions either way, or from either standpoint, what reason do we have for settling on the practical point of view, and holding people responsible, at all?<sup>31</sup> Second, even if we can discover such a reason, won't Kant's view be intransigent? For if we do regard people as free agents, fellow citizens in the Kingdom of Ends, then it seems as if we must treat them as transcendently free and so as completely responsible for each and every action, no matter what sorts of pressures they may be under. Yet the obvious fact is that we live in neighborhoods which are at different distances from the Kingdom of Ends, and it seems merciless to give this obvious fact no weight. But it also seems as if the only option Kant provides is to switch to the theoretical standpoint and regard candidates for forgiveness as if they were no more responsible for their actions than small children and animals. The very idea of an action's being excusable or forgivable or understandable seem to bring together explanatory and justificatory thoughts. The doctrine of the two standpoints seems to keep such thoughts resolutely apart.

In response to the first problem, why we hold people responsible at all, it is important initially to separate two issues. One is the issue of holding yourself responsible for your own actions in the context of deliberative choice, and the other is the issue of holding other people and your self at other times responsible. On Kant's view, we first encounter the idea of freedom when we are deciding what to do. We encounter it in the necessity of acting under the idea of freedom, and in the commands of the moral law.<sup>32</sup> At the moment of decision, you must regard yourself as the author of your action, and so you inevitably hold yourself responsible for what you do. It is only when you think about the actions of other people, and when you think about your own actions at other times, that you can view them from either standpoint. You can take up the position of the social scientist, and regard actions as psycho-social phenomena that need to be explained. Or you can put yourself in the other person's shoes as a decision-maker, and think about

what it is like to *choose* or to *do* an action of that kind.

Now it seems clear that you cannot restrict the concepts of freedom and responsibility to yourself in the context of deliberative choice. If you did, you would think that the only free agent in the world is me-right-now. But the moral law, which according to Kant presents itself to you in exactly these moments, commands that you treat everyone as an end in himself (C2 29-30/29). Unless you hold others responsible for the ends that they choose and the actions that they do, you cannot regard them as moral and rational agents, and so you will not treat them as ends in themselves. Indeed, unless you regard others and your future self as moral agents, there will be no content to your duties at all, for all duties (according to Kant) are owed either to other persons or to the enduring self (MMJ 241/47; MMV 442-444/105-107). The moral law, announcing itself as the law of your will, would be without content or application. Your relations to other people, and to your self at other times, would be, at best, like your relations to small children and the other animals. But there is more at stake here than just whether you have any duties, for you cannot enter into *any* reciprocal relations with people whom you do not hold responsible. Nor can you do this if you do not take responsibility for your own actions at other times, since relationships after all are enduring things.

This is why our reaction to Derek Parfit's nineteenth-century Russian nobleman is that he's wrong, and in particular, that he wrongs his wife. The story goes like this. Parfit's Russian nobleman is now, in his youth, a socialist, and plans to distribute large portions of his inheritance, when he comes into it, to the poor. But he also anticipates that his attitudes will become more conservative as he grows older, and so that he may not think this is the right thing to do, when the inheritance is actually his own. So he asks his wife to hold him to the promise he makes *now*, to distribute the land, even if he tells her *then* that he has changed his mind.<sup>33</sup> Parfit makes it clear that the case is not like that of Ulysses binding himself to the mast to resist the Sirens' song. The young nobleman does not anticipate that he is going to become irrational, that his judgment will be clouded, or that he will be out of control. He merely believes that he is going to think differently than he does now. This case illustrates my point well. The young nobleman's attitude towards his own future attitudes is essentially a *predictive* and theoretical one, and, because it is so, he abdicates the kind of responsibility that is necessary for reciprocity: the kind of responsibility that enables people to act in concert. His way of making himself do the right thing is not to take responsibility for doing so, but to give the responsibility to his wife. This may be one way to form the "united will" that Kant says is necessary in marriage, but it is not the right way. The Russian nobleman leaves his wife alone in the standpoint of practical reason, where people who are married must stand together. Her decision is not, as Parfit says, which of these two men, older and younger,

is her real husband, the man she loves, the man she has married. Nor, for that matter, would *that* be just a question about how she feels about them or what she thinks of them. She cannot be married to the older man, later, unless she holds him responsible, and takes him at his word. She cannot be married to the younger one, now, because he has already abandoned her. And further than that: to the extent that it is important to this woman's sense of her own identity, morally and personally, that she is his wife, he leaves her without anything clear to be, and so without anything clear to do. You cannot act in concert with one who does not act in concert with himself. Where our relations are constitutive of our ongoing identities, those with whom we have them must have ongoing identities too.<sup>34</sup>

So if you only apply the concepts of freedom and responsibility to yourself at the moments of deliberative choice, you do not have any sort of recognizable moral life at all. No Kingdom of Ends on earth can be sought or realized if responsibility is restricted to its original home in the first person deliberator's perspective.

But notice that all of the reasons I have just given are moral and practical ones. I have been suggesting that holding people responsible is something that we do for moral reasons. The reason we must view another as a fellow rational person rather than as a psycho-social phenomenon is not that he is *in fact* one of these things rather than the other. In fact, he is both. That another is responsible is what Kant calls a postulate of practical reason: a belief or attitude that can be formulated theoretically, but is practical and moral in its basis. (C2 132-134/137-139) We hold others responsible in the same way that, according to Kant, we "will that there be a God", because it is a condition of our obedience to the commands of the moral law. (C2 143/149) Or, when a more personal relation is at stake, because it is the condition of our submission to the imperatives of love.

No doubt this way of putting it makes it all sound more deliberate and voluntary than it really is. We do not, of course, simply decide whether to hold other people responsible in general; reciprocal relations and the attitudes that characterize them are, as Strawson argues, too deeply imbedded in the framework of human life to "come up for review", and reactive attitudes, or at least the feelings that accompany them, cannot always be helped.<sup>35</sup> But as Strawson himself observes we do make these decisions in particular cases, and even more frequently we make decisions about whether to identify with our reactive feelings or not. If I have decided not to hold someone responsible, I may view my rage at him as mere inevitable emotion, like the rage provoked in everyone except saints by recalcitrant home appliances and fractious infants. Still, it might be better to put my point a different way. The idea is not that we deliberately decide to hold people responsible in general, but that our commitment to this view of others and our commitment to the moral life issue together from the standpoint of practical reason. Holding others

responsible is an inevitable concomitant of holding ourselves so, both in particular personal relations and in more general moral ones. To share our ends and reasons is to share the standpoint from which those ends and reasons are generated. The citizens of the Kingdom of Ends make their decisions in congress; the noumenal world is, above all, a place that we occupy *together*.

## V. Mitigating Moral Judgment

Now while this explains why we hold others responsible, and why our doing so has and must have a practical basis, it does not solve the problem of what now appears to be Kant's intransigence. The moral command that we hold others responsible seems as absolute as it would be if we had theoretical knowledge that they were indeed transcendently free. Kant does not separate the grounds for holding people responsible in general, from the grounds for holding them responsible for particular actions. And so it seems as if holding someone responsible in general amounts to holding her responsible for everything she does. The flexibility with which I credited the practical account of attributions of responsibility does not seem to follow readily from Kant's view.

Some of the things Kant says, however, suggest that there is room for such flexibility. I will discuss two kinds of considerations, mentioned by Kant, which may be used to guide our decisions whether to hold people responsible for particular actions and reactions, and in particular, to mitigate the intransigence that seems required by the commitment to treating others as persons.

The first consideration springs from what I call Kant's practical compatibilism. Although Kant endorses both free will and determinism, he is not a theoretical compatibilist. Kant does not believe that these two things can be reconciled from a single point of view, as his contempt for Leibniz's *automaton spirituale*, which he says has "the freedom of a turnspit," shows. (C2 97/100-101) And yet this does not stop him from adjuring politicians that "a good constitution is not to be expected from morality, but conversely, a good moral condition of a people is to be expected only under a good constitution." (PP 366/112-113) Nor does it stop him from detailing a theory of moral education designed to awaken our sense of our own autonomy.<sup>36</sup> To the extent, or in the sense, that Kant believes that virtue can be taught, or made to flower by a good constitution, he must believe that it can be caused.<sup>37</sup>

Readers of Kant may want to deny this, for in the *Groundwork*, Kant says that insofar as we are members of the world of sense, our actions "must be viewed as determined by other appearances, namely, desires and inclinations." (G 453/54) But this remark is actually somewhat misleading. Insofar as we view our actions as phenomena we must view them as causally determined, but not necessarily as determined by mere desires and inclinations. We can still view them as determined by moral thoughts and

moral aspirations; only from this point of view, those must themselves be viewed as determined in us. For instance, I might explain someone's doing the right thing by saying that she did it because she values humanity as an end in itself, and I might in turn explain that fact by showing how she received a moral education. And, for that matter, I might explain how that kind of education is possible by appealing to a psychological or even psychoanalytic theory, such as Freud's, of how human beings develop a conscience or superego. A deterministic account can be a deterministic account of moral motivation itself—it does not have to bypass morality and pretend we do everything for the sake of happiness. The element of truth in what Kant says is that a deterministic account necessarily leaves out what is distinctively good about moral motivation. From a merely theoretical and explanatory point of view moral interest is on a footing with inclination. We may imagine the cynic saying: "it doesn't really matter how she came to treat humanity as an end in itself. It is what she likes to do, so she is still pursuing her own happiness." When moral motivation is viewed theoretically, it can be distinguished from inclination only by its content. Its special *source*, in the agent's autonomy, does not show up.

Kant's practical compatibilism suggests that it may be reasonable, when we are deciding whether and when to hold people responsible, to take into account such things as upbringing and education. Depending on the particular circumstances, the fact that someone has had a good moral education may provide a special reason either for forgiveness or for blame, and our decisions about whether to hold him responsible may be governed accordingly. Or it may by itself, quite apart from prediction, provide a special reason for holding someone responsible. When the community has done all it can to make someone good, then there may be no further outlet for respect for humanity, than to blame him if he goes wrong.<sup>38</sup>

Another kind of consideration comes from Kant's iterated demand, in the *Metaphysical Principles of Virtue*, for generosity of interpretation. As I mentioned at the beginning of my discussion, Kant believes that we cannot know people's most fundamental or intelligible characters. But he censures contempt, calumny, and mockery as much for their disrespectful and ungenerous nature as for their lack of a theoretical basis. (MMV 462-468/127-133) He says, for instance, "One should cast the veil of philanthropy over the faults of others, not merely by softening but also by silencing our judgments." (MMV 466/132) Our theoretical estimate of another person's character may be set aside in favor of our respect for the humanity within him. The reproach of vice, according to Kant,

...must never burst out in complete contempt or deny the wrongdoer all moral worth, because on that hypothesis he could never be improved either—and this latter is incompatible with the idea of man, who as such (as a moral being) can never lose all predisposition to good. (MMV 463-464/129)

Kant compares this to the duty, when someone makes an error, not just to deem him stupid but to try to determine how the mistaken view could have seemed reasonable to him. We are to do this in part in order to “preserve the mistaken individual’s respect for his own understanding.” (MMV 463/129) But regarding a person as stupid or making her errors seem reasonable are not our only options in these cases. Sometimes we can best preserve someone’s self-respect, as well as our own respect for her, not by making her errors seem reasonable, but by laughing them off as the result of transitory emotion or exhaustion. The same is surely true in the moral realm. Respect for someone’s humanity is not always best expressed by holding him responsible for each and every action. It may be better to admit that even the best of us can just slip. Indeed Kant’s own doctrine of moral progress, in *Religion Within the Limits of Reason Alone*, has this implication. The phenomenal expression of a noumenally good will is not perfect action in all cases, but progress towards the better. (R 47-48/43) If an anomalous action intrudes into a course of steady progress in virtue, we might find it in our hearts simply to dismiss it as atavistic or transient, or sometimes without any explanation at all. We simply say “He isn’t himself.”

## VI. Conclusion

On the whole, Kant’s view is that we must always hold ourselves responsible, and that we should as far as possible always hold other people responsible. But this is not because people’s noumenal freedom is known to us as a theoretical fact. It is because of the respect which the moral law commands us to accord to the humanity in every person. We hold one another responsible because this is essential to our interactions with each other as *persons*; because in this way we together populate a moral world. We may disagree with Kant about some of the details of how respect for humanity is best expressed, but his theory captures the essential idea that attributions of responsibility have a practical basis. To view people theoretically, as objects of knowledge, is to view them as part of the world that is imposed upon us through the senses, and, to that extent, as alien. But insofar as we are noumena, or active beings, we join with others in those intersubjective standpoints which we can occupy together, either as thinkers or as agents. When we enter into relations of reciprocity, and hold one another responsible, we enter together into the standpoint of practical reason, and create a Kingdom of Ends on earth.

## Notes

1. I have many people to thank for help with this paper. Ken Simons provided extensive and helpful comments which prompted me to make a number of

revisions. Sidney Axinn, Charlotte Brown, Dan Brudney, and Jay Schleusener read and commented on various versions. I presented the paper to several philosophy departments and found all of the discussions helpful; special thanks are owed to audiences at UCLA, the University of Vermont, and the University of Michigan. A short version of the paper, entitled "Holding People Responsible," was presented at the VIIIth International Kant Congress and is forthcoming in the proceedings of that meeting.

2. *Nicomachean Ethics* IX.9 1170b 6-7, with parentheses removed. From the translation by W.D. Ross in *The Basic Works of Aristotle*, ed. Richard McKeon (New York: Random House, 1941), p. 1090. Henceforth cited parenthetically in the text as NE followed by the Bekker page, column, and line references and then the page number from the translation.
3. My reasons for these doubts will become apparent in the course of the paper, although I will not discuss them in the text. If the argument of this paper is correct, the decision whether to hold someone responsible is governed by a variety of considerations, rather than determined wholly by facts about the person. One might think that the legal use of the concept of responsibility requires that the issue of whether a person is responsible be determinable by such facts. Did he understand what he was doing? Does he know right from wrong? If so my view might cause difficulties for it, unless the legal use is not as continuous with the moral use as some believe. However, it is important to notice that my doubts concern the *particular* uses to which the concept of responsibility is sometimes put in our legal system. In a general and philosophical way, the justification of the penal system may rest on our will, as social contractors, to hold one another responsible. But *this* legal use of the concept of responsibility admits of the moral and practical foundation I describe in this paper, and indeed probably requires it. We have no general reason to believe that our fellow citizens are for the most part rational and moral people, who only occasionally go haywire or fall into sin. If I am right, we do have a general reason to hold them responsible: it is *because* they are our partners in the social contract.
4. This is clear from the structure of their theories. But for some more specific statements, see for example the opening paragraph of Francis Hutcheson's *Inquiry Concerning Moral Good and Evil* (in D. D. Raphael, *British Moralists*, (Hackett, 1991) Volume I, p. 261); Hume's statement of the central question of his moral philosophy on p. 456 of *The Treatise of Human Nature* (ed. L. A. Selby Bigge and P. H. Niddich. Oxford: Clarendon Press, 1978). The complaint applies less straightforwardly to Smith, whose theory in general is more sensitive to the perspective of the agent than those of his predecessors. But see, for instance, the opening lines of I.i.5 of *The Theory of Moral Sentiments* (Indianapolis: Liberty Classics, 1982), p. 23.
5. Matthew 7.1. Where I have cited or referred to Kant's works in this paper I have inserted the reference into the text. As is standard, in each case except the *Critique of Pure Reason* and the *Lectures on Ethics*, the first page number refers to the Prussian Academy Edition of Kant's works (*Kants gesammelte Schriften. Preussische Akademie der Wissenschaften*: Berlin, 1900-1942), and the second to that of the translation used. The following abbreviations are used:

- G     *Grounding for the Metaphysics of Morals*. (1785) Prussian Academy Edition Volume IV; James Ellington's translation in *Immanuel Kant. Ethical Philosophy*. Indianapolis: Hackett, 1983. Although I have used Ellington's translation, I refer to the *Grundlegung* as the *Groundwork*.
- C1    *Critique of Pure Reason*. (1st ed.1781, 2nd ed.1787) Page numbers of

- the A and B editions are followed by those of the translation by Norman Kemp Smith. New York: Macmillan, St. Martin's Press, 1965.
- C2 *Critique of Practical Reason*. (1788) Prussian Academy Volume V; Lewis White Beck's translation. Indianapolis: Bobbs-Merrill Library of Liberal Arts, 1956.
- LE *Lectures on Ethics*. (1775-1780) edited by Paul Menzer from the notes of Theodor Friedrich Brauer, using the notes of Gottlieb Kutzner and Chr. Mrongovius; translated by Louis Infield. London. Methuen & Co., Ltd., 1930; rpt. New York, Harper Torchbooks, 1963; current rpt: Indianapolis, Hackett Press.
- MMV *The Metaphysical Principles of Virtue*. (1797) Prussian Academy Volume VI; James Ellington's translation in *Immanuel Kant: Ethical Philosophy*. Indianapolis: Hackett, 1983.
- MMJ *The Metaphysical Elements of Justice*. (1797) Prussian Academy Volume VI, John Ladd's partial translation. Indianapolis: Bobbs-Merrill Library of Liberal Arts, 1965.
- PP *Perpetual Peace* (1795) Prussian Academy Volume VIII, translation by Lewis White Beck in *On History*, edited by Lewis White Beck. Indianapolis: Bobbs-Merrill Library of Liberal Arts, 1963
- R *Religion Within the Limits of Reason Alone* (1793) Prussian Academy Volume VI; translation by Theodore M. Greene and Hoyt H. Hudson. La Salle, Illinois: Open Court, 1934. rpt. New York, Harper Torchbooks, 1960
6. We have two somewhat different uses of the term "responsible." When we say someone is responsible for an action or attitude, we imply that she is a candidate for praise or blame. But when we say someone is a responsible person, we imply that she is reliable, resourceful, trustworthy, and self-controlled. The notion I want is a combination of these but more like the second. we think of the person as someone who should be regarded as reliable and trustworthy and so forth, and *therefore* as a candidate for praise and blame.
  7. See David Hume, *An Enquiry Concerning Human Understanding* (in *Enquiries Concerning Human Understanding and Concerning the Principles of Morals* by David Hume. ed. L.A Selby-Bigge, 3rd ed with text revised and notes by P.H Nidditch Oxford: Clarendon Press, 1975). pp. 183-192 I borrow the term "circumstances of justice" from John Rawls in *A Theory of Justice*. (Cambridge: Harvard University Press, 1971): pp. 126ff.
  8. These remarks obviously assume a particular reading of Kant's Formula of Humanity, according to which what is involved in treating someone as an end-in-itself is respecting her as a rational being, whose choices confer value on their objects, and whose actions must be left to her own autonomous decision I defend this reading in two articles, "Kant's Formula of Humanity" *Kant-Studien*, Band 77, Heft 2 (April 1986) pp. 183-202, especially pp. 197-200; and "The Right to Lie: Kant on Dealing with Evil" *Philosophy and Public Affairs*, Volume 15, Number 4 (Fall 1986) pp. 325-349, especially pp 330-337.
  - 9 See note 15 for some remarks on this point.
  - 10 Here, as several readers have pointed out to me, I am obviously discussing very close and intimate friendships, and saying things that do not hold of less personal but still particular relationships. In these cases perhaps the right thing to say is that reciprocity is heightened, but only in a certain sphere of activity The members of a committee or a department, for example, must take action and make decisions together, and this involves a commitment to treating each another's contributions to these decisions as responsible ones and each other's wishes about them as

having weight This is a heightened form of reciprocity, although only within a delimited sphere But within this sphere what is involved is like friendship. The comparison of factionalized departments to unhappy marriages is a good one. When reciprocity breaks down, and the entity is held together only by formal institutional mechanisms, not only its pleasantness but also its moral character deteriorates.

11. Sometimes Kant unfortunately changes his ground and says the problem is precisely that we don't want the other person *qua* that person, but only *qua* member of a particular gender (LE 164) This is nonsense, and spoils what I take to be of interest in his point.
12. In the *Groundwork*, Kant suggests that in the Kingdom of Ends everything either has a market price, an affective price, or a dignity. Ordinary commodities have market prices, art objects have affective prices, and human beings have dignity (G 434-435/40) Thus my suggestion in the text is that Kant is not worried that sexual desire reduces its object to something with a market price, but to something with an affective price This suggests two further reflections. The first is interpretive Whatever has a price, Kant claims, can be replaced by something else as its equivalent. This is already an odd thing to say about art objects, but it may explain why he was driven to make the bizarre claim mentioned in note 11 above: that we do not desire another as a person but as a member of a gender. The second is more general Many people seem to be more skeptical about the respectability of offering yourself as a direct object of enjoyment than about the respectability of offering your services, especially, of course, if you are a woman. Actresses, entertainers, and models have often been regarded as disreputable characters, while cleaning ladies, nurses, and sales clerks are not thought thereby to degrade themselves. People may even have the obscure feeling that the character actor is more respectable than the movie star, and in this case Kant's analysis fits; for what the movie star offers for our delight is not her talents but simply herself. The view, perhaps surprising but not completely at odds with our intuitions, is then this: being useful is no threat to your dignity, but being delectable is. I do not say this to criticize movie stars, of course, but rather to urge that they are unusually dependent upon the good will and delicacy of their audiences.
13. Again Kant spoils his point, by making an oddly metaphysical-sounding argument that the lover only wants your sexuality but that "It is not possible to have the disposal of a part only of a person without having at the same time a right of disposal over the whole person, for each part of a person is integrally bound up with the whole" (LE 166) But perhaps the argument that sexual love wants its object to be entirely at its disposal can still be made, and made on more interesting grounds than the ones Kant appeals to here Pursuing this line of thought might have forced Kant to admit that the problem he is concerned with here is more of a problem about sexual love than about casual sexual encounters.
14. It is clear from the way Kant sets the problem up in the *Lectures on Ethics* that he sees the problem as arising, so to speak, from the point of view of the sexual object (LE 164) This point should be detachable from the familiar view, which he also sometimes seems to have in mind, that this fact makes the morality of sexuality more of a problem for a woman.
15. In public discussions of this paper, several people pointed out that more needs to be said about the sense in which one is restored to oneself in these relationships. Lawful freedom is not the same as lawless freedom; the condition to which one is restored is not the same. Kant makes this clear in a rather forceful way when he says that marriage produces a unity of will. The kind of reciprocity I am discussing here is not mere exchange, from which one can walk away What is

exchanged is a part of one's practical identity, and what results is a transformation of that identity. Kant's account of marriage is clearly based on Rousseau's account of the social contract, in which "each person gives himself whole and entire" and "in giving himself to all, each person gives himself to no one. And since there is no associate over whom he does not acquire the same right that he would grant others over himself, he gains the equivalent of everything he loses..." Rousseau certainly thinks that this produces a change of identity, since he says it is what transforms a human being from "a stupid, limited animal into an intelligent being and a man." (Rousseau, *On the Social Contract*, trans. Donald A. Cress. Indianapolis. Hackett, 1983: pp. 24-27) This aspect of Kant's view of personal relations has a number of striking implications, among them some that address contemporary criticisms of Kant. From a feminist perspective, Kant has sometimes been accused of denying that personal relationships can be constitutive of identity. See for instance Sally Sedgwick, "Can Kant's Ethics Survive the Feminist Critique?" (*Pacific Philosophical Quarterly* 71 (1990): pp. 60-79, especially p 74) And it has also been argued that his ethics requires that the moral agent be completely impartial among persons in some undesirable way. See for instance, Bernard Williams, "Persons, Character and Morality" in *Moral Luck* (Cambridge: Cambridge University Press, 1981): pp 1-19, especially pp. 16-18. In my view Kant's theory of personal relations provides grounds for challenging both of these views. I hope to pursue these points sometime.

16. Not translated in Ladd. Kant does not draw *this* conclusion, of course. But he comes close. For he goes on to raise the obvious question whether the marriages of his time, which declare the husband to be master, are real marriages, and to assert absurdly that so long as the inequality is really only based on the natural superiority of the man's faculties it is no inequality at all. Both the feebleness and the moral irrelevance of this excuse for inequality suggest the conclusion in the text.
17. In *The Possibility of Altruism* (Princeton: Princeton University Press, 1970-83), Thomas Nagel argues that resentment, for instance, involves the thought that the person resented had a *reason* to act differently than he did. If this is right, and personal relations essentially involve the sharing of reasons, it is clear why personal relations especially involve such reactive attitudes.
18. See Strawson, *Freedom and Resentment and Others Essays* (London. Methuen & Co., 1974), p 4.
19. In his discussion of Gauguin in "Moral Luck," Bernard Williams suggests that, even if we accept Gauguin's success in painting as a justification for his desertion of his family, his family need not do so. Williams thinks that this is because you can do something justified and yet leave some people with a justified complaint. Leaving aside that question, on my view we may at least say this. Given Gauguin's belief in his vocation, we may find his desertion of his family understandable and forgivable—just another instance of the strains which the institution of marriage places on the moral life—while *his wife* certainly need not find in this a reason for forgiveness at all. See Williams, "Moral Luck" in *Moral Luck* (Cambridge: Cambridge University Press, 1981): pp 20-39, and especially pp. 36-37.
20. There are others, of course. For instance, one who knows you well may use his knowledge to manipulate you psychologically. And there is also the simple risk that while you are opening your heart, the other is holding back. Few things are as disconcerting as the discovery that someone in whom you have confided a certain kind of secret or thought or feeling has secrets or thoughts or feelings of a similar kind, which she has not in turn shared with you. This may make you feel exposed, watched, or objectified. You do not need to think that she was

- spying on or judging you in order for this to hurt; the bare failure of reciprocity is enough.
21. We must suppose this, more specifically, to avoid falling into the third antinomy. (C1 A444 & B472-A452 & B480/409-415)
  22. In this sketch of Kant's view I skate over the differences between Kant's accounts of how we arrive at the idea of our own freedom in the Third Section of the *Groundwork for the Metaphysics of Morals* and in the *Critique of Practical Reason*. In the *Groundwork*, Kant's emphasis is on our consciousness of the spontaneity of reason in the production of ideas in general; in the *Critique of Practical Reason*, it is on our awareness of the moral law and of our ability to act from it (the Fact of Reason), which he says reveals our freedom to us (See the references in the text and C2 30-31/30-31) I believe that Kant revised his argument because the spontaneous production of ideas only places us among the noumena as thinkers. To be among the noumena as *agents*, we must be able to *act from* pure ideas, and for this, the positive conception of freedom which is found only in the categorical imperative, as well as our ability to act from that conception, are necessary
  23. This account, which of course is not Kant's, resembles the more traditional rationalist account: incentives incline but do not determine the will. Kant does think that this is how we must regard our own incentives from the practical point of view
  24. It is important to say that the claim is only that it is crazy to regard *Marilyn's* noumenal will, taken by itself, as standing behind the laws of nature. Whether Kant thinks that all rational wills taken together should be regarded as standing behind the laws of nature is a different question altogether.
  25. Kant's language in *Groundwork III* could certainly lead one to believe that he holds this view, and it is this same language which gives rise to the paradoxes mentioned. If we always choose morally in the noumenal world, and if our noumenal choices govern our phenomenal ones, how do bad actions ever occur? And if they do occur, since they cannot be attributed to our noumenal will, how can we be held responsible for them? It is possible that at the time of writing the *Groundwork* Kant had not sufficiently distinguished (what I take to be) his own view from the one under discussion here. I discuss this further in note 29. I discuss the paradoxes about the possibility of evil and responsibility for evil in "Morality as Freedom" in *Kant's Practical Philosophy Reconsidered*, ed. Y. Yovel (Kluwer Academic Publishers, 1989), especially pp. 35-40.
  26. The knowability of pure activity or power is an important theme in modern philosophy, taken up by thinkers as diverse as Descartes and Hume. In the *Second Meditation*, Descartes argues that although we cannot "imagine" ourselves as pure thinkers, that is the role in which we *know* ourselves best (i.e., most free from skeptical doubt) (*Meditations on First Philosophy*, trans. Elizabeth S. Haldane and G.R.T. Ross in *The Philosophical Works of Descartes*, Volume I. Cambridge: Cambridge University Press, 1911, p. 153). Hume, who thinks we get all of our ideas from the senses and therefore cannot have ideas of what we cannot imagine or envision, supposes that we do *not* know ourselves as active thinkers. He tells us that "The uniting principle among our internal perceptions is as unintelligible as that among external objects, and is not known to us any other way than by experience." (*A Treatise of Human Nature*, cited in note 4 above, p. 169). But the view comes out most clearly in the *Dialogues Concerning Natural Religion*, in remarks like "But the ideas in a human mind, we see, by an unknown, inexplicable economy, arrange themselves as to form the plan of a watch or a house. ." (Part II, p. 146) and "We have indeed, experience of ideas, which fall

into order, of themselves, and without any known cause..." (Part IV, p. 162; quoted from the edition by Norman Kemp Smith, Macmillan Library of Liberal Arts, 1947). Kant's move here as everywhere is to find a path between empiricism and rationalism, using what is right in both positions. Hume is correct in tying what we can know to what can be represented. The world must show itself to us before we can apply the concepts that give us understanding. But he is wrong in thinking we can only *have ideas* of the sorts of things we can know. What we can think is not exhausted by what we can know: our concepts do not all come from sensible intuition. Descartes is right in insisting that we can *think* about our activity. But he is wrong to suppose that we *know* ourselves as thinkers and agents. Our agency, although not knowable, is *intelligible*, and we must think of it. (C1 A538 & B566/467)

27. These remarks apply to the moral law, on the practical side, and to the regulative principles of reason, on the theoretical or speculative side. Something more complex must be said about the constitutive principles of the understanding, an issue which I here leave aside.
28. Including, in the theoretical or speculative realm, deciding how to proceed with our investigation or theory construction. In fact, when describing and explaining our behavior we must view ourselves both ways, since we appear in the role of thinker as well as that of object thought about.
29. The reader may wonder whether I am suggesting that Kant was simply wrong in the *Groundwork* when he said that insofar as we are members of the intelligible world we necessarily will according to the moral law, and that if we were only members of that world we would will always according to that law (G 453/54). The answer is no, but here I think it is significant that in the *Groundwork* Kant uses the language of "intelligible" and "sensible" rather than that of "noumenal" and "phenomenal"; and also that he changes his language in the *Critique of Practical Reason*. As I understand these terms, the noumenal world is the intelligible world insofar as it is thinkable. If we think of noumena at all, we must think of them as acting in the only way that is intelligible to us, which is according to the laws of freedom. But at the same time we must always admit the possibility that the noumenal world is unintelligible to us. The trouble with the way Kant phrases the argument in *Groundwork III* is that it can make it sound as if the normative force of the moral law followed from its descriptive application in the noumenal world: "Now if I were a member of only that world [the intelligible world], all my actions *would* always accord with autonomy of the will. But since I intuit myself at the same time as a member of the world of sense, my actions *ought* so to accord." (G 454/55) If we suppose, naturally but incorrectly, that the normativity of morality enters the scene with the "ought", Kant seems to be deriving a normative sensible "ought" from a descriptive intelligible "is". But he is not, for the laws of the intelligible world are normative through and through. The moral law characterizes noumena insofar as they are intelligences (insofar as we can think of them) because acting according to it is the only thing it makes sense for them to *do*, and this is already a normative point.
30. This is slightly overstated, since Kant does think that insofar as we are free we think of ourselves as the causes of our action, and this idea plays an important role in his ethics at various crucial moments. But since he insists that free causality is an idea without a theoretical employment, the point still holds. (C2 49/50, 56/57-58; 133-136/137-142)
31. Perhaps I should make it clear that the question I am asking here concerns the way we make this decision in a case where it is *already* clear that we *can* view the creature and its actions in either of these two ways. Kant thinks we can do

this whenever the actions are performed by a human being. I am not concerned here with what justifies that view—that is, I am not discussing the question why we think that human beings are candidates for being held responsible while the other intelligent animals, who make some use of reason and with whom we may enter into some forms of relationship, are not. This is an important question, but it requires a separate treatment.

- 32 This remark again straddles the accounts in the *Groundwork* and in the *Critique of Practical Reason*, since I think that both elements are involved in Kant's best explanation of how we come to think of our own freedom. See my "Morality as Freedom," pp. 39-40, cited in note 25 above.
33. See Derek Parfit, "Later Selves and Moral Principles" in A. Montefiore, ed. *Philosophy and Personal Relations* (London: Routledge & Kegan Paul, 1973, pp. 145ff.; and *Reasons and Persons* (Oxford: Oxford University Press, 1984) pp. 327-328.
34. I discuss the practical construction of our own identities in "Personal Identity and the Unity of Agency. A Kantian Reply to Parfit" *Philosophy and Public Affairs*, Volume 18, Number 2 (Spring, 1989): 101-132. The issue of whether relationships can be constitutive of identity is touched on in note 15 above.
35. "This commitment [the natural human commitment to ordinary inter-personal attitudes] is part of the general framework of human life, not something that can come up for review as particular cases can up for review within this particular framework." Strawson, "Freedom and Resentment", cited in note 18 above.
- 36 Kant's theory is spelled out in the "Methodologies" of the *Metaphysical Principles of Virtue* (477-484/145-153) and of the *Critique of Practical Reason* (151-163/155-168), as well as in his book *Education*. (1803, trans. Annette Churton. Ann Arbor: University of Michigan Press, 1971).
- 37 Kant denies that we can have a duty to promote the moral perfection of others, on these grounds "For the perfection of another man as a person consists precisely in his being able to set his end for himself according to his own concepts of duty. And it is a contradiction to require (to make it a duty for me) that I ought to do something which no one except another himself can do." (MMV 387/44) But this, again, is overstated. Granted, that it would be both disrespectful to you, and unfair to me, to hold me responsible in a general way for your moral character. Yet it is clear that we have a duty to provide for the moral education of our children, and, Kant himself insists, our intimate friends (MMV 470/136) Choosing ends on another's behalf is as impossible as it would be disrespectful, but putting others in a good position to choose ends for themselves, and to choose them well, is the proper work of parents, teachers, friends, and politicians; providing for someone's moral education as well as nurturing her self-respect is an important part of the way we do this
- 38 Nor is Kant unaware of the more direct educational benefits of holding others responsible, for he reminds us that "Examples of respect shown to others may also incite in them an endeavor to deserve it." (MV 466/132) In *Ethics and the Limits of Philosophy* (Cambridge: Harvard University Press, 1985), Bernard Williams writes, "The institution of blame is best understood as involving a fiction, by which we treat the agent as one for whom the relevant ethical considerations are reasons .. This fiction has various functions. One is that if we treat the agent as someone who gives weight to ethical reasons, this may help to make him into such a person." (p. 193) It is presumably this form of "recruitment into the deliberative community," to use Williams's phrase, that he has in mind when he writes "The purity of morality conceals not only the means by which it deals with

deviant members of its community, but also the virtues of those means.” (p. 195) Williams thinks that “the fiction of the deliberative community is one of the positive achievements of the morality system” but adds “As with other fictions it is a real question whether its working could survive a clear understanding of how it works” (193-194) I want to make two comments about these remarks. First, the view of persons we adopt from the practical point of view will seem “fictional” (if that is supposed to suggest some form of inferiority) only to those who privilege the theoretical standpoint and its concepts, or at least believe that all our concepts should be congruent with those. This suggests a certain view of what concepts in general are for. No doubt theoretical concepts are more firmly aimed at tracking the truth, but tracking the truth is not the primary business of ethical concepts, as Williams would certainly agree. In any case the term “fiction” is one adopted from the theoretical standpoint, and relativized in an obvious way to the purposes of theoretical reason. My second point concerns recruitment into the deliberative community. Kant himself apparently thought that we *can* understand how holding people responsible works—and even, as the quotation above suggests, that we can take notice of its more strategic benefits—and yet go on doing it. Of course it is a delicate business to manipulate someone into morality while maintaining the essentially non-manipulative attitude that morality demands. But, as Kant’s remarks about error at MMV 463/129 (quoted in Section V) show, he rightly perceives this to be a quite general problem about education.