The Philosophical Review, XCIV, No. 1 (January 1985)

INVOLUNTARY SINS

Robert Merrihew Adams

I. The Issue

"You have heard that it was said to the ancients, 'You shall not commit murder, and whoever commits murder shall be liable to judgment.' But I say to you that everyone who is angry with his brother without cause shall be liable to judgment."[1] The ordinary moral consciousness of our civilization supports the view ascribed to Jesus in these familiar lines. Exaggerated or senseless anger, an anger that is not justified by a proportionate provocation, is morally offensive; and one who is guilty of it is liable to blame.

Yet this view seems inconsistent with another thesis that commands very wide support—the thesis, namely, that we are ethically accountable only for voluntary actions and omissions. For anger is not in general voluntary. It is apt to be manifested in voluntary behavior, and we may make voluntary efforts to control it; but we cannot choose to be angry or not as we normally can choose to sit or stand. We may be very angry when we are trying hard not to be, and a conscious effort to be angry when we think that our dignity demands it is quite likely not to be successful. Even if we have learned to suppress conscious episodes of anger by redirecting our attention, we may still be profoundly affected by an unconscious anger. This absence of control is quite normal, moreover. It is not an exceptional failure of a natural power, like the paralytic's inability to move his limbs.

How then can we be blamed for anger? My answer is that the thesis that we are ethically accountable only for our voluntary actions and omissions must be rejected. There are involuntary sins,[2]

---

[1]This is a translation of the traditional text of Matthew 5:21–22. The qualification, "without cause," is omitted by some of the best ancient manuscripts and by most modern editors. This textual variation is surely not an accident. It reflects a substantive disagreement: is *all* anger sinful? We need not go into that issue here, however. For my present purpose it is better to focus on the weaker, more qualified claim of the traditional text.

[2]I use 'sin' as a convenient term to mean any action, omission, or state

and unjust anger is only one of them. Among the others are jealousy, hatred, and other sorts of malice; contempt for other people, and the lack of a hearty concern for their welfare; or in more general terms, morally objectionable states of mind, including corrupt beliefs as well as wrong desires. My aim in this paper is to defend and develop this view of the matter. And the first task (in Sections II to IV) will be to consider three alternatives to my answer to the question of how we can deserve blame for bad states of mind.

II. Voluntary and Involuntary Manifestations

One alternative is to hold that what is to be blamed in anger, jealousy, and contempt is not an involuntary state but the pattern of voluntary behavior in which that state is apt to be manifested. On this view we are not accountable for a motive unless we have voluntarily consented to it by acting on it. But the ethics of the heart cannot be reduced to an ethics of conduct in this way.

Obviously there are types of voluntary action that are characteristic of anger, for example, and that are morally objectionable. But that is not all that is wrong about unjust anger. If I am furiously angry at you without proportionate cause, but refrain from hitting you, insulting you, yelling at you, or speaking ill of you, no doubt my self-restraint is a credit to me; but the anger remains a moral fault. It is also a fault that is apt to do harm, casting a pall over relations between us, and making it impossible for me to think fairly about you when it may be important to do so. And I am to blame for such consequences of my unjust anger, even if they are consequences simply of my being angry and not of any voluntary action or omission of mine.

The insufficiency of any theory that regards states of mind as blameworthy only in their voluntary manifestations is particularly glaring in the case of *self-righteousness*. It sometimes happens that two people sin approximately equally against each other although one of them offends greatly against the other in voluntary actions

---

that is inherently blameworthy. I have no wish to disavow the theological connotations of the term, but do not mean to base any argument on them in this paper. By 'involuntary' I mean simply *not voluntary*, not necessarily *against the subject's will*.

whereas the behavior of the second party is almost impeccable. The latter's offense is self-righteousness. He wants to be above the other person, to be in the right against the other person. The passion with which he clings to this superiority is what energizes all his wit and will-power to do his duty as he sees it—and to see his duty as he has done it. He thinks he would like the other to be a better person, but he would hate to lose his position of being in the right. He is more interested in that position than in friendship or the larger ends of morality. And of course something of this arrogant and self-centered attitude comes through to the other person, with damaging effect.

The offense of the self-righteous person is typically not in what he voluntarily chooses to do, but in the motivation and attitude with which he usually does what he ought to do. In choosing to do his duty as he sees it, he is acting on the bad motive of self-righteousness as well as the good motive of conscientiousness. But it would be misleading to accuse him of "consenting" by this act to the bad motive; for we do not think he ought to have chosen not to do his duty as he sees it. What we think is rather that he ought not to have had the bad motive in the first place.

This is not to say that behavior that flows from self-righteousness is likely to be totally indistinguishable from that which flows from a purer conscientiousness. But a self-righteous contempt may be manifested in facial expressions and other cues of which the self-righteous person is unaware, much more than in choices; in what he never even thinks of saying, much more than in what he decides not to say. If something that he voluntarily says is contemptuous, he may not realize that it is; and in that case he is not voluntarily contemptuous, since we voluntarily do only what we know we are doing. His saying it is still an offense, but what we chiefly blame in it is the attitude behind it. His blameworthiness would not necessarily be less if his tongue were so carefully mild that only totally involuntary cues revealed his contempt. If his self-righteousness is subtle, his voluntary actions may include many efforts to take a humble stance. He may be profuse in his apologies. But he is not sincere in them; they are a strategem to keep him in the right, and he is congratulating himself on his goodness all the while. This of course is felt by the other person.

The self-righteous one faces a dilemma, whether he realizes it or

not. He can either pretend or not pretend to be humble rather than self-righteous. If he does not pretend to be humble, he will be more openly self-righteous, and that is offensive. If he does pretend, he is insincere, and that is offensive too, and will probably be unsuccessful. Whatever he does in the matter will be wrong, so long as his motive is to be in the right, to be above the other person. And having this motive is not voluntary. He may not even know that he has it, and he probably could not get rid of it just by deciding to give it up. He is not able voluntarily to avoid offending. According to the doctrine that all sin must be voluntary, it would seem that he is not to blame. And yet I think he clearly is to blame, not because of his voluntary choices but because of his self-righteous attitude. We will have a lop-sided view of the guilt in human relationships if we do not recognize this.

III. Control and the Voluntary

Both the alternatives to my position remaining to be considered are attempts to understand blameworthy states of mind (and not just actions manifesting them) as voluntary. In the present section we shall consider the view that they are directly or inherently voluntary, and not involuntary at all. The suggestion that they are not directly but indirectly voluntary, in the sense of being traceable to one's own previous voluntary acts and omissions, will be discussed in the next section.

The most plausible case for regarding desires, for example, as directly voluntary is based on a certain conception of the will. The noun 'will' has been used in a variety of senses. It has often been used quite broadly, to signify a more or less rational appetitive or conative faculty whose functions include desiring—and perhaps loving, hating, liking, and disliking—as well as the initiation of voluntary motion.[3] So why shouldn't we say that desires are inher-

[3]Medieval philosophers, conceiving of the will in this way, regarded love and hate as acts of the will, and in that sense as voluntary. 'Appetitive faculty' is a Scholastic term for a faculty having such functions as are mentioned in the text. What is meant is not merely a faculty of physical or instinctive appetite. That would be a "sensitive appetitive faculty." But the will was conceived of as a "rational appetitive faculty," which responds to intellectual data.

ently voluntary, on the ground that they are operations of the will, conceived as such a faculty?

I have made a decision to use 'voluntary' in a sense that is not so broad as to correspond to this broad sense of 'the will'. When I say that desires are not in general voluntary, I do not mean to deny that they can be ascribed to the will in that sense. Desires can certainly be ascribed to the faculty of desiring, if we are going to speak of faculties at all. This suggests that the question whether desires are directly and inherently voluntary is at least partly verbal rather than factual.

It is not only a verbal issue, however, and the problem of involuntary sins cannot be disposed of simply by defining all the operations of a rational appetitive faculty as "voluntary." The chief reason why this is not sufficient is that, at least for most modern thinkers, the claim that we are ethically accountable only for what is voluntary is not to be understood in terms of faculty psychology, but means that we are responsible only for what is within our voluntary *control.* That is the widely supported thesis that I am disputing. And the obvious fact that desires can be ascribed to the faculty of desiring does nothing to show that they are within our control in the relevant sense. Indeed it is clear that we cannot normally control our desires, and the other states of mind that concern us here, as we can control the voluntary motions of our bodies. For this reason I shall continue to use 'voluntary' and 'involuntary' in a sense for which control and not faculty psychology is the test. Whether all the states of mind for which we are directly accountable, ethically, are to be ascribed to an appetitive faculty, is another question, which will not be treated systematically in this paper—though in Section VI some states of mind will be discussed for which I think we are directly accountable and of which it is not obvious that they can rightly be ascribed to an appetitive, rather than exclusively to a cognitive, faculty.

It is not a simple matter to explain the nature of this control which is the decisive criterion of the voluntary for most modern thought.[4] A really adequate explanation would require a complete

[4]In resting the conception of the voluntary on this sort of control, rather than on faculty psychology, the predominant tendency in modern thought

theory of human action, and more. Here we shall have to be (provisionally) content with much less. But the importance of the concept to my project demands at least a brief account. I think the following is roughly right, though it may be controversial.

To say that something is (directly) within my voluntary control is to say that I would do it (right away) if and only if I (fully) tried or chose or meant to do so, and hence that if I did it I would do it *because* I tried or chose or meant to do it, and in that sense voluntarily. (This definition may need refinements to deal with atypical cases, but I think they are not of crucial importance for the difference that concerns us between voluntary actions and involuntary states of mind.[5]) 'Control' suggests a controlling of which that which is in our control is the object; and the key to this conception is the idea of a trying, choosing, or meaning of which what is voluntarily done is (or would be) the object. Several points should be noted about what I mean here. (1) I am claiming that in a voluntary action the action itself is an object aimed at by the agent in a way that is aptly described as "trying, choosing, or meaning." (2) I do not mean to commit myself as to whether, in those cases in which the voluntary deed is actually done, the trying, choosing, or meaning is a separate act, or only an aspect of the way in which the deed is done. (3) But I do assume that it is possible to (fully) try or choose or mean to do something, and not do it—through paralysis,

---

has departed from the medieval views mentioned in note 3. I suspect that the Protestant Reformers played an important part in this change. See especially the first (1521) editions of Philip Melanchthon's *Loci Communes*. But this is not the place to develop a history of the subject.

[5]One refinement that could be considered would be a proviso that doing x is not directly but only indirectly within my voluntary control if I could try successfully to do x right away, but only by trying successfully to do something else as a means to doing x. We would have a case in point if I could count on making myself angry, at once, by thinking about my job. But perhaps my definition, unrefined, already excludes my anger in such a case from direct voluntary control. For it does not seem that in this case I would become angry only if I tried to; it does not even seem that I would become angry through thinking about my job only if I tried or chose or meant to become angry. For I might think voluntarily about my job only because I had to think about it in order to gain my bread; and then my anger would be only an unintended (though possibly foreseen) consequence of my voluntary action, and not something that I tried or chose or meant to produce.

for example, or some other failure in execution. And in general such failures in execution would show the deed not to have been within the agent's voluntary control. (4) There is a kind of failure in execution, however, which does not show the deed not to have been within the agent's voluntary control, and that is the kind that is due to weakness of will.[6] There is a sense in which one can "try" to stand one's ground in the face of danger and fail through cowardice, or "try" to resist temptations of pleasure and fail through intemperance. But those are cases in which one does not *fully* try or choose or mean (at the crucial moment) to stand or resist; that is precisely a part of what is meant by calling them cases of weakness of will. It is only where the action or omission was within one's voluntary control that weakness of will is needed as an explanation. (5) I do not mean to be committing myself one way or the other on the issue of whether voluntariness is compatible with determinism. I leave open the question whether an indeterminist proviso should be added to the definition of voluntary control. (6) Trying, choosing, or meaning can itself be regarded as voluntary, as an operation of the will, so to speak, that is characteristic of voluntary action. It may perhaps also be said, in an extended sense, to be within our voluntary control, though it is itself the controlling, rather than an object of control.

Desires and emotions are not (directly) within our voluntary control in the sense I have indicated. For (a) they are not cases of trying, choosing, or meaning, as is illustrated by the fact that if I simply desire to do something, it remains a question whether I will try or choose or mean to do it. And (b) it is rarely true that one would have a particular desire or emotion right away if and only if one (fully) tried or chose or meant to have it. In having desires and emotions, indeed, we often are not *aiming* at having them at all, and

[6]Another sort of failure in execution, about which we may hesitate to say whether it shows the deed to have been out of the agent's voluntary control, is that due to inattention to detail. The inattention is a voluntary omission, and one would have succeeded if one had paid attention. On the other hand, it may be that the failure of the enterprise as a whole was not due to any lack of trying to succeed in it. One did not fully try to attend to the detail, but that may be due to an error of judgment rather than to half-heartedness or weakness of will. Perhaps we should say that such an error of judgment causes the success of the larger project to "escape from" one's voluntary control.

usually not in a way that is aptly described as "trying or choosing or meaning" to have them. Of course we sometimes do try to have or acquire a desire or emotion; and over a period of time such efforts may be fairly effective. But they do not normally succeed immediately. Wanting, and trying without immediate success, to have better motives is a familiar and painful experience.

Painful, and at times harmful, as our lack of direct voluntary control over our desires and emotions may be, I do not mean to suggest that it is a tragic flaw in our nature. Would it be better if we were able, more reliably and with more immediate effect, to govern such states of mind in accordance with our ethical beliefs and larger purposes? I doubt it, for several reasons. One reason (which there is not space to explore here) is that in many cases such control would undermine the sincerity or genuineness of the desire or emotion produced.

A more obvious reason is that our impulses are often better than our principles. In many cases the voice of duty or of wisdom turns out to have spoken in the uncontrolled feeling or desire rather than in the principled conscience or the rational calculation. If our emotions and desires could not put up a stout resistance against our apparently moral or rational judgments about them, we would be easy prey to fanaticism and doctrinaire folly.

Here (without committing myself to any interpretation of the *Republic*) I may make use of Plato's device of likening the soul to the state. I am rejecting the view that the soul ought to be like a state autocratically ruled by a philosopher-king, with the practical reason orchestrating and determining the desires and feelings and taking no advice from them. What I think is and ought to be the order of the soul is not a pure democracy, and certainly not mob rule, but something more like the American system of representative government with "divided powers," with opposing tendencies and competing interests retaining an independent voice and influence—with state and federal governments, legislative and executive and judicial branches, public and private sectors, all acting in their own right in ways that directly affect the moral character of the nation and of its relations with other peoples. It is important for the individual, as for the state, to be able to *act* fairly consistently over time in accordance with rationally coherent policies subjected to ethical reflection. But it is also important for the individual, as for the state, to have potential sources of dissent within—

to have, as it were, organs that can take positions that the chief executive wishes they would not take. The ever present possibility of internal conflict is not only a vexation and a potential hindrance to resolute action; it is also a wellspring of vitality and sensitivity, and a check against one-sidedness and fanaticism.

In trying to come to terms with the fact of inner conflict, we may be tempted to restrict our moral responsibility and our moral selfhood to the part of us (so to speak) that deliberates and resolves about our lives. This would be as much a mistake as it would be to hold our nation, as such, responsible only for the acts of its President, and not also for the decisions of its courts and the activities of its business corporations which its constitution allows to operate with considerable independence of the government. Moral credit and discredit is reflected on a nation by the kinds of dissenters that it has and lacks. Our desires and emotions, though not voluntary, are responses of ours, and affect the moral significance of our lives, not only by influencing our voluntary actions, but also just by being what they are, and by manifesting themselves involuntarily. Who we are morally depends on a complex and incompletely integrated fabric that includes desires and feelings as well as deliberations and choices.

IV. Voluntary and Involuntary Roots

If a state of mind is not directly within our voluntary control, it might still be *indirectly* within our voluntary control. That is, there might be (directly) voluntary steps we could take that would predictably, over time, affect our state of mind in ways that we would choose. Among philosophers interested in the topic of this paper, probably the commonest opinion has been that we are accountable for wrong states of mind only insofar as they are consequences (that should have been anticipated) of our own voluntary acts and omissions in the past—and thus only insofar as they were indirectly within our voluntary control. On this view I can rightly be blamed for a desire or emotion only insofar as voluntary actions or omissions of mine that led to it were blameworthy.[7]

[7]For attractive presentations of ideas of this sort, by philosophers who emphasize the role of the emotions in the moral life, see Edward Sankowski, "Responsibility of Persons for Their Emotions," *Canadian Jour-*

The first thing to be said about this theory is that it is right in supposing that we ought to try to improve our motives by voluntary action. We ought to cultivate good motives and try to root out bad ones. Many philosophers assume, however, that these duties of self-culture exhaust our obligations in this matter. Believing that all ethical principles must be *action*-guiding, they can take the question, "What states of mind ought we to have?" as an ethical question only by reading it as "What states of mind ought we to *try* to develop and maintain in ourselves?"[8] It is one of my principal contentions that this is mistaken—that many involuntary states of mind are objects of ethical appraisal and censure in their own right, and that trying very hard is not all that is morally demanded of us in this area. We ought not only to *try* to have good motives and other good states of mind rather than bad ones; we ought to *have* good ones and not bad ones. On my view the ethics of motives, and more generally of states of mind, has a certain independence, and is not merely a department of the ethics of actions. The subject of ethics is how we ought to live; and that is not reducible to what we ought to do or try to do, and what we ought to cause or produce. It includes just as fundamentally what we should be for and against in our hearts, what and how we ought to love and hate. It matters morally what we are for and what we are against, even if we do not have the power to do much for it or against it, and even if it was not by trying that we came to be for it or against it.

The issue between me and the theory that I am rejecting here is how fully one's responsibility for a bad state of mind must be accounted for by wrong voluntary actions and omissions in the past by which one has caused it, or failed to prevent it, in oneself. The reader is asked to imagine a case in which to examine this question. Suppose you have just realized that you are ungrateful to someone who has done a lot for you—perhaps at great cost to herself. Far from responding to her sacrifices with love and gratitude, you have made light of them in your own mind; and if the truth be told, you

---

nal of Philosophy, 7 (1977), pp. 829–840, and L. A. Kosman, "Being Properly Affected: Virtues and Feelings in Aristotle's Ethics," in Amélie Oksenberg Rorty, ed., *Essays on Aristotle's Ethics* (Berkeley, Los Angeles, and London: University of California Press, 1980), p. 112f.

[8] I have given a brief critique of this preconception in "Motive Utilitarianism," *Journal of Philosophy*, 73 (1976), p. 474f.

actually resent them, because you hate to be dependent on others or indebted to them. Surely this attitude is blameworthy. Must we assume that you have caused it, or let it arise, in yourself by actions that you have voluntarily performed or omitted?

What have you done that you ought not to have done? Suppose you have acted ungratefully toward your benefactress. Perhaps you have even shown resentment toward her, and the resulting estrangement may have contributed to your present ungrateful state. But why did you act ungratefully in the first place? Because you already *were* ungrateful. That ingratitude is as shameful as your present ingratitude, and was not caused by your subsequent ungrateful behavior.

What have you left undone that you ought to have done? You have not begun sooner to struggle against this ingratitude. But it would not be correct to say that you have thereby voluntarily consented to the bad attitude. For voluntary consent, as ordinarily understood, implies knowledge; and you did not realize that you had a problem in this area. How then can you be blamed for not having fought against your ingratitude? You are not to be blamed for not taking steps to rid your house of termites when you had no reason to suspect that there were any termites in the vicinity. Of course ingratitude is different from termites. You should have known of your ingratitude. Why didn't you? Presumably because you did not want to recognize any shameful truths about yourself—because at some level you cared more about having a good opinion of yourself than about knowing the truth about yourself. And that's a sin too, though not a voluntary one. Thus the search for voluntary actions and omissions by which you may have caused your ingratitude keeps leading to other *in*voluntary sins that lie behind your past voluntary behavior.

In truth there is something odd about the search for voluntary faults to explain our responsibility for wrong states of mind. If someone says to me that I am incapable of feeling gratitude, or that I do not sincerely care about my moral character, or that although I act rightly I do so only because I think it the most effective way to get what I want from other people, this claim about my feelings or motives is already an ethical indictment; and if it is true, I stand condemned. There is no need to search for guilty actions or omissions of which I may be accused.

And if they are found, the blame that I deserve for them may be much less than I deserve for my state of mind. The morally imprudent voluntary omissions, for example, by which a person has failed to pay the price to extricate himself in time from a situation that has left him embittered, cynical about morality, and full of racist resentments, may be less gravely blameworthy than the attitudes to which they have led. Indeed we might think them blameless, a successful gamble, if the sequel had not left the person so corrupted. These considerations confirm the intuitively plausible judgment that what we chiefly blame in the present immoral state of mind is not the imprudence of the previous voluntary omissions.

V. Repentance and Responsibility

I have often met with the objection that it is unfair to hold people responsible for having the right states of mind and not just for trying to have them. Many feel that a person who through heredity or a bad upbringing or the weakness or corruption of human nature finds herself in a bad motivational state, but struggles bravely to get out of it, deserves at least as good a moral rating as the person who does not have to fight against those motives because she has never had them. Maybe she *is* no less meritorious on the whole, deserving praise for her efforts as well as blame for her bad motives. But that is not inconsistent with my position.

I wish to emphasize, rather than deny, that there are things that we can and should voluntarily do to improve states of mind that are not directly within our voluntary control. If we practice behaving appropriately, for example, after a while we are likely to feel more appropriately too, at least in many cases. We could seek out a counselor who has helped others change in the direction we think would be right for us. If we hold religious beliefs they will probably indicate prayers or other steps to take in search of better motives. Experience yields no guarantee of complete success for any of these voluntary endeavors; and the connection between the effort and the improvement is often obscure and variable, providing a measure of empirical support for religious claims that the improvement depends on divine grace. Still we are responsible for taking or not taking those voluntary steps that we can.

The question before us is whether that exhausts our responsibility in the matter. I think it does not, and I will argue that a closer look at the struggle against bad states of mind actually supports my contention that we are accountable for those that we have and not just for voluntary efforts to improve them.

There must be, in the first place, a motive for the struggle. Going through the (voluntary) motions is not enough for a serious effort. Many a counselor must have said to a client, "I cannot help you unless you *want* to change." What is called for here is not merely actions but a desire. And having or lacking the desire is not voluntary in the sense that concerns us.

So far is it from being the case that our moral merit or demerit for having or lacking good or bad motives depends on our voluntary efforts to get good ones and exclude bad ones, that the reverse of that is closer to the truth. We do give people credit for trying to improve their desires and attitudes. In this we assume, however, that their motivation is conscientious. Suppose I were trying hard, without success, to get rid of certain ethically offensive attitudes, but you discovered that I was making this effort only because I was curious to see if I could succeed in it. Would you still give me credit for trying? I think you shouldn't. In this case I would deserve censure for the frivolous amorality of my attitude toward my own moral improvement, as well as for my other bad attitudes. My not desiring good attitudes for their own sake, or at least from some good motive, is an involuntary state; yet it is both reprehensible in itself and sufficient to destroy all the moral credit I would otherwise get by my voluntary efforts to reform.

In the second place, the struggle against a wrong state of mind in oneself is normally a form of *repentance,* which involves self-reproach. At the center of such a process is one's *taking responsibility* for one's state of mind.[9] Your ingratitude (to return to a previous example) is not a voluntary action; but if you take responsibility for it you also do not see it as something that just happens to you, like a toothache or a leak in your roof. You see it as an opposition that you yourself are making, not voluntarily but none the less really, to the generosity of the other person and to your own position as a recip-

[9]I mean taking responsibility for *having* the attitude now, as well as for doing something about it in the future.

ient of love and assistance. In repentance you repudiate this opposition, not as an evil existing outside the inner circle of your selfhood, but as your own; and you reproach yourself for it.

The deepest reason for accepting this responsibility, if I am right, is that it is rightly ours. It is important for a correct ethical appreciation of one's own life. To refuse to take responsibility for one's emotions and motives is to be inappropriately alienated from one's own emotional and appetitive faculties. But we are also interested here in another reason for accepting responsibility for our states of mind—namely, that it is useful for moral improvement to do so.

Those who insist that ethics must be *action*-guiding may object that it is useless to take responsibility for states of mind, or to reproach ourselves for them, or to regard them as wrong, except insofar as we can control them by voluntary action. But this objection reflects a gross overestimate of the role of will-power in moral change. Certainly there is an important place for voluntary efforts at reform. And it is notorious that recognizing that one's mental state is bad does not always cure it. But seeing one's wrong desires and attitudes, recognizing their blameworthiness, and accepting responsibility for them do seem commonly to have a beneficial effect on one's state of mind—probably as much effect as effort. The acceptance of responsibility is important to repentance because it enlists the desire to satisfy one's responsibilities in support of the desire to change; whereas if one says, "I'm not to blame for my ingratitude because I can't help it," one takes some of the pressure off oneself by seeking refuge in an excuse. And the part of one that wants to be ungrateful (if I may speak somewhat metaphorically) will all too naturally read the excuse as granting a right to be ungrateful.

The acceptance of responsibility for one's states of mind can also be important for interpersonal reconciliation, which is intimately linked with repentance. If a bad attitude of mine has contributed to poisoning my relationship with another person, it is surely not useless for me to take responsibility for the attitude and blame myself for it. So long as I refuse to do that, I will see the other person as more responsible than he is for whatever is wrong with our relationship; whereas if I admit my wrongness, my loss of ill-founded pride is apt to facilitate reconciliation. Whether or not reconcilia-

tion can be attained, indeed, and whether or not I can root out my bad attitude, I owe it to the other person to recognize my guilt for it in order to respond to the moral situation as it really is.

VI. Cognitive Sins

As far as the blameworthiness of a bad state of mind is concerned, it does not matter very much whether you know that you have it. That you have never noticed your ingratitude or your jealousy toward so-and-so is no excuse. And as for self-righteousness, by its very nature it is apt to be wholly or partly hidden from the one who is corrupted by it. For the typically or consummately self-righteous person is not merely pretending to others, but has persuaded himself, that he is better than he really is. His self-righteousness includes a blindness to faults such as self-righteousness in himself. It would be far too soft on self-righteousness to accept ignorance that one has a bad attitude as an excuse for having it.

It is true that a person who knows he has a bad attitude will rarely be assigned the same blame as if he did not know it. But that is not because the blameworthiness is increased by knowledge or diminished by ignorance as such. It is rather that our evaluation is affected by what he does with his knowledge. If he does nothing voluntarily to stand against his bad attitude, we are apt to blame him more than if it never crossed his mind that he had it; but if he is struggling to rid himself of it, we will probably blame him less than if he were oblivious to it. And the blame for a bad attitude may be especially severe in the case of a person who continues to believe that he does not have it in spite of the evidence of its presence that has been pointed out to him.

All of this is additional proof that the blameworthiness of states of mind is not dependent on voluntariness. For one of the definitive marks of the voluntary is that if one does something voluntarily one knows that one is doing it. The tradition that attaches blame exclusively to the voluntary has therefore accepted ignorance as an excuse. It has not blamed us for doing things if we did not know that we were doing them or did not know that they were wrong.

This statement must be qualified, however. For it has usually been admitted that ignorance itself can be culpable (that is,

blameworthy), and that culpable ignorance is no excuse—or is only a partial excuse, depending on the degree of culpability of the ignorance. And if we have a bad state of mind without knowing it, this ignorance will commonly be culpable. The unawareness of one's own self-righteousness is itself an offensive piece of self-righteousness, and ignorance of one's own hostility or ingratitude can constitute a lack of integrity.

I do not think this is the chief reason why states of mind that are wrong are not excused by our ignorance of them. A culpable lack of self-knowledge is only part of what is offensive in typical cases of self-righteousness or ingratitude. And the blameworthiness of the rest of the offense is independent of whether and how one is ignorant of it. But instead of pursuing this line of thought let us turn to a question that opens up a whole new department of involuntary sins. How can ignorance be culpable? For knowledge and ignorance seem in general to be involuntary states. They do not even fall, with intention and desire, in the category of appetitive or conative states, but rather in that of cognitive states.

It is clear, however, that ignorance of one's faults can be blameworthy; and it is not the only cognitive state of which this is true. There are many cognitive failures that we regard as morally reprehensible.[10] Some examples are: believing that certain people do not have rights that they do in fact have; perceiving members of some social group as less capable than they actually are; failing to notice indications of other people's feelings; and holding too high an opinion of one's own attainments. These failures are not in general voluntary. *Trying* to pay attention to other people's feelings will not necessarily be successful, if one is insensitive or afraid of emotions. And *trying* to assess one's own abilities and accomplishments accurately may not keep one from thinking too highly of oneself, if one is vain. We do give people credit for trying in these matters, but we still regard the failure to notice other people's feelings or one's own deficiencies as a fault—and a fault that lies within the domain of ethics.

The traditional way of trying to accommodate such facts as these to the doctrine that only the voluntary is blameworthy is to say that

[10]I have developed this point, partly on the same and partly on different lines, in "The Virtue of Faith," *Faith and Philosophy,* I (1984), pp. 4–6.

cognitive errors (and specifically ignorance) are culpable only insofar as one has contributed to them by voluntary actions and omissions of one's own. In a very interesting recent discussion of the issue, Alan Donagan maintains a position of this type: "Ignorance . . . is culpable if and only if it springs from negligence—from want of due care" (or "from want of due consideration," as he puts it elsewhere).[11] Negligence, in this context, is a voluntary omission of actions that one ought to have performed and that would have cured or prevented the ignorance.

Donagan points out that false beliefs about moral principles may be due to bad education rather than to negligence, and concludes, "A graduate of Sandhurst or West Point who does not understand his duty to noncombatants as human beings is certainly culpable for his ignorance; an officer bred up from childhood in the Hitler *Jugend* might not be."[12] I disagree with Donagan's conclusion. The beliefs ascribed to the graduate of the Hitler *Jugend* are heinous, and it is morally reprehensible to hold them (even if one has no opportunity to act on them). No matter how he came by them, his evil beliefs are a part of who he is, morally, and make him a fitting object of reproach. He may also be a victim of his education; and if he is, that gives him a particular claim to be regarded and treated with mercy—but not an exemption from blame.

There is also something questionable about the role of the idea of "due consideration" of moral matters in Donagan's position. *Is* there a standard (however imprecise) of due consideration of ethical issues to which we hold rational agents in general? Is it a sin to graduate from college without having taken a course in ethics (or two or three) if it would not have been an enormous personal burden to take one? I don't think so—much as the general acceptance of such a claim might improve the employment prospects of philosophers.

Consider, at the other extreme, the professional moralist who has spent a lifetime of painstaking work refining and testing his principles, but who without conscious hypocrisy has developed an

[11]Alan Donagan, *The Theory of Morality* (Chicago: University of Chicago Press, 1977), pp. 130 and 134. The first of these passages has to do with "ignorance of fact," the second with ignorance of moral truths; Donagan says essentially the same thing about both.

[12]*Ibid.*, p. 135.

utterly bizarre system of ethics that suits his own misanthropy and taste for eccentricity. Is he inculpable for the inhumanity of his opinions if he holds them sincerely and aspires to live by them? Surely not. Yet we can hardly say that his errors proceed from "want of due consideration."

Cowardly beliefs provide another example in which the culpability of a belief seems not to depend on negligence. James Wallace has argued persuasively that while a charge of cowardice can sometimes be rebutted by showing that one believed that what one did not do would not have been worth the risks involved, one will still be convicted of cowardice if one held that belief because one was excessively afraid of the risks involved.[13] It seems that an overestimate of the dangers involved in an otherwise desirable course of action is apt to be culpable if the error is rooted in excessive fear. But errors of this sort do not seem in general to be due to negligence. For the cowardly are not less *careful* than the brave in estimating dangers.

This example illustrates the point that false beliefs and other cognitive failures are often culpable at least partly because of other *in*voluntary sins that are manifested in them (such as excessive fear in this case). Indeed blameworthiness of cognitive failures commonly arises more from this factor than from duties of (voluntary) self-culture that one has failed to perform (which is not to say that it must be explained entirely by either or both of these causes). If I have failed to recognize another person's feelings, for instance, how much I should blame myself depends much more on whether I have the respect and concern that I should have for that person than on whether I was conscientiously trying to notice his or her feelings.

The connection between cognitive sins and other wrong states of mind is a two-way street. Cognitive failures may be culpable at least in part because they *manifest* a bad attitude, as I have just noted. But it is also true that a morally offensive attitude may be *constituted* in part by an offensive belief. Thus the false belief that certain people lack certain rights is not just a *consequence* of the bad attitude of disrespect or contempt for those people; it is part of what *con-*

[13]James D. Wallace, *Virtues and Vices* (Ithaca, N.Y.: Cornell University Press, 1978), pp. 67–74.

stitutes disrespect or contempt for them. As such it is as vicious morally as any other aspect of contempt.

VII. Blame

In a recent discussion with which I am otherwise largely in agreement, Lawrence Blum holds that emotions, feelings, values, and attitudes are appropriate objects of "moral criticism," but not of praise or blame. On this view most racist feelings and beliefs, for instance, are inside rather than outside one's selfhood for purposes of moral evaluation, and we may "think poorly" of a person for having them, but must not *blame* him for them because they are not voluntary. "To say that someone is to blame for something, seems to [Blum] to imply that he could have brought it about through his will that he did otherwise."[14]

I have found a number of philosophers sympathetic with Blum's point of view. They grant much of what I have been contending for. Bad attitudes, morally inappropriate emotions, corrupt ethical beliefs—all are ethical *faults,* they agree. But they balk at the idea of "blaming" people for them, or for anything but voluntary acts and omissions.

Perhaps for some people the word 'blame' has connotations that it does not have for me. To me it seems strange to say that I do not blame someone though I think poorly of him, believing that his motives are thoroughly selfish. Intuitively I should have said that thinking poorly of a person in this way *is* a form of unspoken blame. I am not sure, however, that there is only a verbal dispute here, about the use of the word 'blame'. Blaming is a type of response to faults in oneself or in others. So there may be some substantive disagreement about how it is appropriate to treat us in view of our involuntary faults.

I do not believe that exactly the same responses are appropriate to involuntary as to voluntary sins. In particular, only voluntary acts and omissions are rightly *punished* by the state, or even by a university or a club.[15] A wicked motivation, or a lack of sincere

[14]Lawrence A. Blum, *Friendship, Altruism, and Morality* (London: Routledge & Kegan Paul, 1980), p. 189.

[15]It is less easy to say whether God, or parents of small children, can appropriately punish involuntary states of mind. Perhaps what may seem

remorse, may perhaps be allowed to aggravate the punishment for a misdeed; but it is not right to punish someone solely for an emotion or attitude or belief that has not issued in some forbidden voluntary conduct. I am not sure why this is so—but then I am also not sure why punishment is appropriate when it is appropriate.

Punishment is only one kind of response to wrong, however; and it presupposes some special conditions. For instance, punishment is appropriately imposed only by someone with authority to do so. There are other, less specialized responses to wrong; one of them is *reproach.* The line between punishment and reproach is not always sharp. In some cases it may be unclear whether a scolding, for example, is a punishment or merely a reproach. In either case it is clear enough that scolding is blaming—but that is getting a little ahead of the argument. Persons in authority may both punish and reproach an offender for the same offense. But no authority is needed to reproach. In many cases children may rightly reproach their parents, and subordinates their superiors. Punishment is most at home in something like a courtroom. Reproach is at home in a much wider variety of relationships in which one is affected by, or appropriately cares about, another person's behavior or attitudes.

I have two chief claims about reproach. The first is that reproaching is a form of blaming.[16] Blame is therefore at least as widely appropriate as reproach is. In particular, blame is appropriate in many cases in which punishment is not.

My second claim about reproach is that people may rightly be reproached (and therefore blamed) for involuntary as well as for voluntary moral faults (indeed, more severely reproached for some involuntary than for many voluntary faults). This is undoubtedly connected with the fact that *anger* often seems an appropriate as well as natural response to offensive states of mind—appropriate

---

like punishment in these cases might be better understood under the less juridical categories of anger, taking offense, alienation, reproach, and discipline. Sankowski, op. cit., pp. 829–831, also holds that emotions, as such, though subject to various other "sanctions," are not subject to punishment.

[16]I do not mean to suggest that blaming is always an activity like punishing or reproaching. That you have always blamed somebody for something can be true simply on the basis of beliefs and feelings you have had.

at least in the sense that it need not betray any failure to appreciate the situation.[17] The following reproaches clearly express blame for involuntary states; yet each of them would be appropriate, and deserved, in some circumstances.

"You don't really feel sorry for what you've done!"

"It really hurts my feelings that you don't feel anything about what I'm going through."

"It's arrogant of you to think you have a right to do that."

Perhaps it will be objected that the blame expressed by these reproaches is not *moral* blame. We blame ourselves, and may be to blame, for consequences of our sheer incompetence, for example, morally inculpable though it be; and that is not moral blame. It seems very odd, however, to say that blame incurred for a *moral* fault and its consequences, as in these cases, is not moral blame. I think it would feel like moral blame to most of us.

I do not mean to advocate, or even defend, a censorious approach to moral faults. But judgments about censoriousness must be distinguished from judgments about blameworthiness. There are several grounds on which it might be wrong to reproach someone even if the reproach is deserved. There are situations in which it is none of our business to reproach someone, though he richly deserves it. Whether you have a right or duty to reproach another person may depend on your stake in the matter, or your relation to him, or on whether he has (explicitly or implicitly) invited you to correct him. The rights of privacy and community are such that in general it is likelier to be our business to reproach someone for an action or inaction than for a bad state of mind. But it would be a mistake to infer that the former must be more blameworthy than the latter.

It may also be thought unseemly to reproach anyone for a bad motive that we all share, on the principle that "people who live in glass houses should not throw stones." Perhaps this is to some extent a principle of self-protective expediency rather than of morality; perhaps one has moral standing to reproach another person for a fault that one shares, provided one is struggling against it. But even if we should adhere to the principle without qualification, it

[17]On the role of anger (and of guilt feelings) here, cf. Sankowski, op. cit., p. 830.

would not follow that a "normal" degree of selfishness in our motives, for example, is not something blameworthy for which we could appropriately be reproached by a morally perfect person.

Another important ground for not reproaching someone for a wrong state of mind is repentance. It will be much less appropriate to harangue an offender with reproaches if he is already reproaching himself for his fault, and struggling hard, though with meager success, to rid himself of it, than if he is doing neither of these things. The reason for this, in my opinion, is that we ought in general to be treated better than we deserve, and we have a corresponding duty to be merciful. No matter how hard I am struggling against a bad motive that I still have, another person who is freshly hurt by an involuntary manifestation of it is within his rights in reproaching me for it; and the reproach is deserved. It may even be helpful in establishing a general recognition of the moral facts of the situation. But to go on beyond a certain point in reproaching someone for a state of mind that he is trying hard to change, and for which he is reproaching himself, serves no good purpose. It is unmerciful and vindictive.

These last comments may also help to deal with the objection, which I have encountered, that it is oppressive to assign blame to states of mind as well as to acts and omissions.[18] The appropriate purpose of reproach, and of judgments of blame, directed at others or at oneself, is not to crush us but to lead us to repentance, and to acknowledge moral realities. In religious terms, there is need for the forgiveness of sins. To the extent that one sees a problem of oppressiveness here, moreover, it would be catering to self-righteousness to try to solve it by leaving the burden of blame on the shoulders of the shoplifter and the car thief while lifting it from the self-seeking motive and the contemptuous attitude that may involuntarily poison the conscientious deed.

VIII. Towards a Theory of Responsibility

As one who holds that there are involuntary sins, I have sometimes been challenged to explain what are the limits of our respon-

[18]Cf. H. H. Price, "Belief and Will," reprinted in Stuart Hampshire, ed., *Philosophy of Mind* (New York: Harper & Row, 1966), p. 113: "If we are to be allowed, or even encouraged, to blame [people] for the way they direct

sibility for our own condition. If I lack natural aptitude for music or athletics, am I accountable for that? Surely not. I am not to blame for it even if it is a great disappointment to my parents. Likewise pneumonia, blindness, mental retardation, and a broken arm are not in themselves blameworthy, though they are certainly undesirable conditions. If I have caused a physical disability in myself deliberately or by some negligence, I may be to blame for it; but even in that case my disability is not, like a bad motive, something that it is a new offense to have for an additional day. These facts are easily explained if it is agreed that we are ethically accountable only for what is voluntary. But what is to keep us from being to blame for every undesirable state we are in, if involuntary states can be blameworthy?

It is not exactly necessary for me to meet this challenge. We could know that we are accountable for some involuntary states and not for others without knowing what (besides our accountability) distinguishes the two kinds of states from each other. Indeed I have less confidence in any general theory about this than I have in many particular judgments to the effect that this involuntary state is a sin but we are not to blame for that one. Nonetheless I have some principles to suggest, somewhat tentatively, for a theory of responsibility.

All the involuntary sins we have been considering are, in the first place, states of mind. They fall in many psychological categories: desires, emotions, attitudes, beliefs and other cognitive states; the list is not intended to be exhaustive. And the lack of an appropriate state in any of these categories can also be blameworthy. All of these are states of mind; and perhaps the reason why pneumonia, blindness, fractured limbs, and lack of athletic talent cannot be blameworthy in themselves is that they are not states of mind (and of course also not voluntary acts or omissions).[19]

---

their thoughts, as well as for their actions, there will be a perfect orgy of moral indignation and condemnation, and charity will almost disappear from the world." I owe this reference to Edward Sankowski, "Love and Moral Obligation," *The Journal of Value Inquiry*, 12 (1978), p. 109, whose response seems to be in agreement with me that regarding a state of mind as blameworthy is not incompatible with a charitable response to it.

[19]We can be blamed for bodily movements and states such as a grimace or a supercilious expression, even when they are not voluntary; but that seems to be due to the states of mind that they express.

Among states of mind, moreover, we are concerned only with those that are directed toward an object—that have, in other, more technical philosophical words, an "intentional object." We are angry *at* someone, glad *of* something, sorry *about* an event; we have attitudes *toward* persons and things; we notice states of affairs; and beliefs have propositions as their objects. All these mental states are directed toward an object in the relevant sense. Whether a state of mind that is not so directed can be blameworthy in itself, I am not sure. Low intelligence and lack of musical talent cannot—though it may be odd to speak of them as "states of mind" at all. But what about a generalized anxiety or irritability? I will not try to answer that question here. I will be content if I can give even the main lines of a theory of our responsibility for states of mind that have intentional objects. (For convenience I will treat the *lack* of a state *s* as having the same intentional object as *s*.)

One of the more attractive ideas behind the use of voluntariness as a criterion of responsibility is that we cannot be morally responsible for anything unless some of its springs or causes lie within us in an appropriate way. I believe this idea can be retained in a theory of involuntary sins. To this end we may try to frame a criterion of our accountability for states of mind in terms of the way in which they arise in us. My suggestion is that among states of mind that have intentional objects, the ones for which we are directly responsible are those in which we are responding, consciously or unconsciously, to data that are rich enough to permit a fairly adequate ethical appreciation of the state's intentional object and of the object's place in the fabric of personal relationships. Among the states of mind for which we would not be accountable under this criterion are simple feelings of hunger and thirst, insofar as they are primitive responses to physical stimuli, and many states of mind in young children whose experience is not yet rich enough for adequate appreciation of their objects.

In relating responsibility to the possibility of ethical appreciation I am to some extent following the lead of traditional theories of responsibility for voluntary evil, which hold us accountable only when our intellect noticed or should have noticed the badness of the act. But whereas the traditional theories are concerned with *conscious* recognition of the badness of the act, my criterion demands only that the data to which we are responding be rich

enough to *permit* recognition of the relevant values. And I think it would not be plausible to limit our responsibility for states of mind to cases in which we are or should have been conscious (so as to be able to say) that we are responding to those data. We expect of our desires and feelings an ethical sensitivity that exceeds what we can articulate.

My criterion is one that cannot be applied without considerable exercise of moral judgment. What ethical appreciation would be "fairly adequate?" And when are the data "rich enough?" There will be disagreements on these points, and some of them will be substantive ethical disagreements.

One tricky point is the place of intellectual understanding among the data. In view of my argument about cognitive sins, I would not say in general that exposure to enlightened ethical views is among the data required. I take it the imaginary Hitler *Jugend* alumnus discussed above (in Section VI) has rich enough data in his evidence of the humanity of the noncombatants in question, even if he is never told that they have rights. This will normally be true even if he has never met a member of the race or ethnic group to which the noncombatants belong; it is enough to know that they are human beings. On the other hand, I am prepared to grant, for example, that some conception of a preferable, workable alternative system may be part of the data needed for a fairly adequate appreciation of the injustice of a social or economic system, and that one's experience and education my leave one innocently unable to imagine such an alternative—though in any particular case it would be hard to know that one's imagination was not constrained by a culpable fear or complacency or attachment to privilege.

It is important to note that a state of mind can satisfy this criterion of accountability, and also be undesirable in some way, and yet not be ethically blameworthy—not because we are not responsible for it, but because it is not ethically wrong. I am not prepared to offer a complete theory of the conditions under which states of mind are right or wrong. But my point can be illustrated by the consequences of some of the most plausible principles about rightness and wrongness of desires. It is in general blameworthy to want, for its own sake, something that is intrinsically bad from a moral point of view, such as the suffering of another person. But it

is not in general wrong to want something that is only accidentally bad. For this reason it is not always wrong to want to do something that it would be wrong to do; for the object of the desire may not be inherently bad. Suppose you wake up wanting to go to the beach, but judge that you have a duty to go to work instead. Your desire for surf and sand is inconvenient on this occasion, and it would presumably be wrong to act on it; but the desire itself is not blameworthy, as going to the beach is not inherently or generally wrong. Conversely it is not necessarily innocent to want to do something that it would be right to do. Even if capital punishment were right, for example, a desire to inflict violence on other human beings that would lead one to seek a career as executioner would be reprehensible. If the desire to kill is wicked and the desire for seaside recreation is innocent, it is not because we are more responsible for the former than for the latter, but because the former's object but not the latter's is intrinsically bad from a moral point of view. (I do not mean to commit myself to the view that that is the *only* ground on which desires can be ethically blameworthy.)

Thus far I have said nothing about the relation of moral responsibility to the metaphysical issue of determinism. I do not think it poses any special problem for my position. That is, it poses no greater threat to our responsibility for our states of mind than to our responsibility for our actions.

As is well known, there are two chief alternative views of the relation of the issue of determinism to our moral responsibility for our actions: the compatibilist and the incompatibilist. According to *compatibilists* an action is free, and the agent may therefore be morally responsible for it, if and only if its causes lie within the agent in the appropriate way—and the appropriate way is roughly that the agent did it because she chose to do it. Qualifications may be added to this criterion to rule out specific sorts of determination of the agent's choice, as for instance by hypnosis; but it is *not* ruled out that the agent's choosing to do the action was completely determined by previous causes, as for instance by some combination of heredity and environment. On this view free action, and moral responsibility, are compatible with complete causal determinism.

Incompatibilists, of course, deny that freedom and moral responsibility are compatible with determinism. They hold that an action is free, and the agent may be morally responsible for it, if and only if

the agent did it because she chose to do it *and* was not completely determined by any causes to choose to do it rather than not. If we believe in incompatibilist free agency, we will very likely account for it in terms of *agent causation.* That is, we need not say that the choice just happened, without a cause. We may hold that the choice, or the free action, is caused by the agent herself—that by virtue of her power of free choice she is a cause of the action or the choice, even though nothing causally determined her to choose one way rather than the other.

Now it seems to me that both compatibilist and incompatibilist treatments are also possible for the relation of the issue of determinism to our moral responsibility for our states of mind. Not that the two analyses of the notion of free action can be directly applied to states of mind; but analogous theories about states of mind can be constructed, according to which moral responsibility is in the one case compatible, in the other case incompatible, with complete causal determinism.

According to the *compatibilist* theory what is required if I am to be morally responsible for a state of mind is that it be mine, and that its causes lie within me in the appropriate way. My principal suggestion as to what way that is is that the state must arise in me in response to data that are sufficiently rich in certain respects, as indicated above. As in the case of act-compatibilism, we may want to add qualifications to the effect that the state must not be due to hypnosis or insanity or some other condition that keeps me from "being myself." On this view I am still to blame for a wrong state of mind even if it is the inevitable result of my heredity and environment, or the unforeseeable result of past voluntary actions that I had no reason to believe were wrong. "What does it matter who or what is in the causal chain further back?" the compatibilist will ask. "If (on the basis of rich enough data) *I* am for the evil and against the good, that is condemnation enough, and I cannot pretend that it is someone or something else that is judged."

The *incompatibilist* view, on the other hand, is that if I am to be morally responsible for a state of mind, it must not only meet the requirements of the compatibilist theory but must also be a state in which I am spontaneous in such a way that I was not completely determined by any causes to have it. It may be somewhat untraditional to speak of spontaneity in this context, but I find it rather

plausible. Phenomenologically, it does not seem to me that desires, emotions, and moral convictions just happen in me, or assail my soul from the outside, like an ache in my ankle. They are not voluntary actions, but they seem to be ways in which *I* am responding to something. It might be misleading to speak of "agent causation" here, since in an important sense I am not performing an action in being in the state. But the notion of agent causation is a fragment of a larger, pre-Humean conception of causation, according to which effects are produced not only, nor perhaps even primarily, by events, but also (or maybe, at bottom, only) by substances, acting in accordance with their powers. So the incompatibilist might speak here of "substance causation," and hold that we as substances, endowed with the power to respond to certain objects with either love or hate, acceptance or rejection, are the causes of the involuntary states for which we are morally responsible.

This theory of substance causation suggests that our relation to our desires, emotions, and beliefs may be more like our relation to our voluntary choices than we might have supposed; and I do mean to suggest that. But desiring is still not the same as choosing; and the indeterministic spontaneity postulated by the theory does not imply that our desires, emotions, and beliefs are within our voluntary control after all. Because voluntary control is a matter of what we could do by trying, choosing, or meaning to do it, that implication would hold only if the way in which we caused ourselves to be in such a state of mind were by trying, choosing, or meaning to be in it. And substance causation in this theory is not to be understood as working in that way. My causing of my desire for comfort, for example, is simply a causal relation between me (a substance) and the state of desiring in which I am. It does not involve an aiming at the desire which could be called "trying, choosing, or meaning" to have the desire. In spontaneously desiring comfort I aim at comfort, the object of the desire; but I do not necessarily aim at having the desire itself. We use the notions of trying, choosing, meaning, and direct voluntary control precisely to mark an empirically evident difference between actions and most states of mind—that is, to mark a way in which we govern the former and cannot govern the latter. A plausible theory of substance causation will not obliterate this difference.

I am not certain which of the two theories of moral responsibility (compatibilist or incompatibilist) is correct, either for actions or for states of mind (though metaphysically I am inclined to favor indeterminism and substance causation). Whichever way we go, we can claim that the springs or causes of everything for which we are responsible lie within us in an appropriate way. I have been trying to show that that condition of responsibility can be satisfied for involuntary states of mind as plausibly as for actions, on either deterministic or indeterministic assumptions.[20]

University of California, Los Angeles
Center of Theological Inquiry, Princeton

[20]I am indebted to numerous auditors and readers—particularly to Marilyn McCord Adams, Lawrence Blum, Jean Hampton, Herbert Morris, Charles Robbins, Ivan Soll, Allen Wood, and the editors of *The Philosophical Review*—for comments that have contributed to revisions in this paper. My greatest debt of this sort is to Thomas E. Hill, Jr., for many fruitful discussions of these topics. I am grateful to the Center of Theological Inquiry for fellowship support in the form of a resident membership in the Center at Princeton during one of the periods in which the paper was written.