I. A PATERNALISTIC THEORY OF PUNISHMENT

HERBERT MORRIS

The rational love of parents for their children then guides the parents' conduct so that their children may one day be fortunate enough to say with St. Paul, "... when I became a man, I put away childish things." A central drama of many lives is a result of imbalance in the relations between parents and children in this area—of being left too much on one's own or too little, of counting on one's parents too much, or of not being able to count upon them enough, of parental conduct that fosters too great dependence or conduct that imposes upon the child too great a personal responsibility, creating in the child not self-confidence but a sense of being alone and insecure in a threatening world.

Paternalism as a social phenomenon is prefigured in this elemental and universal situation of solicitous parental conduct that has its roots in our common humanity. But paternalism is of philosophic interest, not because of the way parents legitimately relate to their children—indeed there is oddity in describing this conduct as "paternalistic"—but rather because something like this practice is introduced into relations among adults. If our responses to adults mirror intrusive and solicitous parental responses to children we behave paternalistically.

Contemporary discussions of paternalism, understood in this way, proceed by focusing primarily on specific laws, laws that either prohibit or require certain conduct and that, arguably, have as their principal or sole reason for existence the good of those individuals to whom they are addressed. My focus in this paper is entirely different, for I consider paternalism, its meaning and its possible legitimacy, not in the context of specific laws prohibiting or requiring conduct, but rather with regard to the existence of a system of punitive responses for the violation of any law. I shall consider several issues and make a number of proposals. First, I define my particular version of a paternalistic theory of punishment. Second, I argue for, and consider a variety of objections to, this paternalistic theory. Third, I argue

NOTHING is more necessary to human life, and fortunately nothing more common, than parents' concern for their children. The infant's relatively lengthy period of helplessness requires that others nourish and protect it. And the child's existence as a vital being with an interest in the world, a capacity for eagerness and trust, and a sense of its own worth, all depend upon its receiving loving care, understanding, and attention. With time the normally developing child relinquishes its almost total dependence; it acquires the capacity to conceive of itself as an agent, to set out on its own, and to live in a world less dominated by its bodily needs and by its parents. Inevitably, this growth in competence and strength brings greater potential for self-harm, for the child's fantasies of its power and knowledge stand in marked contrast to the reality of its relative ignorance and vulnerability. In the ordinary course of events, the more powerful and knowledgeable parent often interferes with the child's choices in order to prevent harm and to bring about good and the reason for this is frequently, if the appropriate degree of parental selflessness is present, the child's own best interests, not primarily the interests of the parents or others.

Concern for the child often, of course, is manifested in allowing and encouraging experimentation just as it sometimes is in forceful intrusion. The child's developing individuality and sense of personal responsibility require that others encourage it a sense of its own power and competence, support its venturing out, and exercise judgment in forbearing from intrusion, permitting it to err and to learn some painful truths from painful consequences suffered. God commanded Adam and Eve but left them free to disobey, thereby providing evidence both of his love and respect. The Devil, preferring for humans a state of permanent infantilism, would, no doubt, have acted differently as Dostoevsky's Grand Inquisitor nicely illustrates.¹

¹ What is gained and what is lost by allowing a choice to disobey is also brought out in C. S. Lewis' engaging replay of the Adam and Eve myth in his novel Pelandra.
that the paternalistic theory I have constructed implies, in a more natural way than other common justifications for punishment, certain restrictions on the imposition of punishment.

II

Let us turn to the first topic. My aim here is to describe the paternalistic theory of punishment I later defend. I set out a variety of moral paternalism, for the good that is sought is a specific moral good.

First, then, in order to punish paternalistically we must be punishing. I assume that the human institution of punishment presupposes, of course among other things, that certain conduct has been determined to be wrongful, that what are generally recognized as deprivations are imposed in the event of such conduct, that these deprivations are imposed upon the wrongdoer by someone in a position of authority, that wrongdoers are generally made aware that the deprivation is imposed because of the wrongdoing, and that the context makes evident that the deprivation is not a tax on a course of conduct or in some way a compensation to injured individuals but rather a response to the doing of what one was not entitled to do.

I have placed a logical constraint on the concept of punishment that is not customarily explicitly associated with it. I have claimed that in order for a person to be punished there must be an intention—one normally simply taken for granted—to convey to the wrongdoer, and where it is punishment for breach of a community’s requirement, to others as well, that the deprivation is imposed because of wrongdoing. A communicative component is a defining characteristic of punishment and in part distinguishes it from mere retaliation or acting out of revenge where the goal of bringing about evil for another may achieve all that one desires. The paternalistic theory I present relies essentially on the idea of punishment as a complex communicative act—the components of which I hope will become clear as I proceed.2

A central theme in paternalism is to justify one’s conduct out of a concern for the good of another. And so a paternalistic theory of punishment will naturally claim that a principal justification for punishment and a principal justification for restrictions upon it are that the system furthers the good of potential and actual wrongdoers. This contrasts with views—though many of the practices supported may be the same—that it is justice that requires that guilty persons be punished or that it is the utility to society that requires punishment. The theory I put forward emphasizes what retributivist and utilitarian theories largely, if not entirely, ignore, that a principal justification for punishment is the potential and actual wrongdoer’s good. The theory should not, however, be confused with “reform” or “rehabilitative” theories. First, these theories may be based, not on consideration of what promotes the good of actual and potential wrongdoers, but on what promotes value for society generally. A reform theory, further, may countenance responses ruled out under the paternalistic theory proposed in these pages. And, finally, reform theories usually fail to address the issue of how instituting a practice of punishment, meaning by this both the threat of punishment and its actual infliction, may promote a specific moral good and this is a central feature of the theory I propose.3

I also assume that paternalistic measures characteristically involve disregard of, indeed conflict with, a person’s desires. Giving a person what they want and being motivated to do so for that person’s good is benevolence not paternalism. And so, if a longing for punishment were characteristically the way in which people responded to the prospect of its imposition, there would, I think, be no role for a paternalistic theory regarding the practice, for it would simply be a practice that generally supplied people with what they acknowledged wanting. We may speak meaningfully of a paternalistic theory of punishment for two reasons: first, punishment by its nature characteristically involves a deprivation that individuals seek to avoid, with the implication that there is some conflict between what people want and what they get; second, the practice is such that the desires of a person at the time of the deprivation are not determinative of what they receive. Thus, while there are obviously persons guilty of wrongdoing who desire punishment, this fact will not affect either its being punishment that is meted out to such a person or the punishment being possibly based on paternalistic consideration, for what is customarily viewed as a deprivation is being imposed independently of the individual’s desires.

Most importantly, the theory I am proposing

3 See generally Walter Moberly’s splendid The Ethics of Punishment (London: Faber and Faber, 1968), particularly pp. 201 ff.

3 The reform theories discussed by H. L. A. Hart and found to be unacceptable as answers to the question what could be “the general justifying aim of punishment” differ, then, from the theory developed in these pages. Hart’s change of mind in the notes to his collection of essays is occasioned by consideration of theories that still differ markedly from the one I propose. See Punishment and Responsibility (New York and Oxford: Oxford University Press, 1968), pp. 24–27, 240–41.
requires that the practice of punishment promote a particular kind of good for potential and actual wrongdoers. The good is a moral one, and it is, arguably, one upon which all morality is grounded.

What is the character of this good? It has a number of component parts but it is essentially one’s identity as a morally autonomous person attached to the good. This statement obviously needs explanation. First, it is a part of this good that one comes to appreciate the nature of the evil involved for others and for oneself in one’s doing wrong. This requires empathy, a putting oneself in another’s position; it also requires the imaginative capacity to take in the implications for one’s future self of the evil one has done; it further requires an attachment to being a person of a certain kind. The claim is that it is good for the person, and essential to one’s status as a moral person, that the evil underlying wrongdoing and the evil radiating from it be comprehended, comprehended not merely, if at all, in the sense of one’s being able to articulate what one has done, but rather comprehended in the way remorse implies comprehension of evil caused. A person’s blindness about such matters—this view assumes—is that person’s loss. The Devil’s splendid isolation is his hell.

Of course, this element of the good makes it apparent that for this theory, as with other moral justifications for punishment, that the rules defining wrongdoing, the rules whose violation occasions punishment, themselves meet certain minimal moral conditions. I assume, and do not argue for the view, that attachment to the values underlying these rules partly defines one’s identity as a moral being and as a member of a moral community, that it gives one a sense of who one is and provides some meaning to one’s life, and that the price paid for unconcern is some rupture in relationships, a separation from others, a feeling ill at ease with oneself, and some inevitable loss of emotional sustenance and sense of identity. I further assume that attachment to these values is a natural by-product of certain early forms of caring, understanding and respect and that the practice of punishment applies to those with such an attachment and not to those who because of some early disasters in primary relationships might value nothing or possess values we might attribute to the Devil.

Second, it is a part of the good that one feel guilt over the wrongdoing, that is, that one be pained at having done wrong, that one be distressed with oneself, that one be disposed to restore what has been damaged, and that one accept the appropriateness of some deprivation, and the making of amends. Not to experience any of this would be to evidence an indifference to separation from others that could only, given the assumptions I have made, diminish one as a person.

Third, it is also part of the good that one reject the disposition to do what is wrong and commit oneself to forbearance in the future. I assume that this makes possible, indeed that it is inextricably bound up with, one’s forgiving oneself, one’s relinquishing one’s guilt, and one’s having the capacity fully to enter into life.

Finally, it is part of the good that one possess and vividly retain a conception of oneself as an individual worthy of respect, a conception of oneself as a responsible person, responsible for having done wrong and responsible, through one’s own efforts at understanding and reflection, at more clearly coming to see things as they are with a deepened attachment to what is good. This conception of oneself is further nourished by freely accepting the moral conditions placed upon restoring relationships with others and oneself that one has damaged.

It is a moral good, then, that one feel contrite, that one feel the guilt that is appropriate to one’s wrongdoing, that one be repentant, that one be self-forgiving and that one have reinforced one’s conception of oneself as a responsible being. Ultimately, then, the moral good aimed at by the paternalism I propose is an autonomous individual freely attached to that which is good, those relationships with others that sustain and give meaning to a life.

The theory I propose claims that the potential of punishment to further the realization of this moral good is one principal justification for its existence. From the perspective of this form of paternalism there must be full respect in the design of the practice of punishment for the individual’s moral and intellectual capacities. The good places logical and moral constraints on the means that it is permissible to employ to realize it. This is the principal reason that I earlier emphasized the communicative aspect of punishment, for on this theory we seek to achieve a good entirely through the mediation of the wrongdoer’s efforts to understand the full significance of the wrongful conduct, the significance of the punishment being imposed, and the significance of acceptance of that punishment. Thus, unacceptable to this theory would be any response that sought the good of a wrongdoer in a manner that bypassed the human capacity for reflection, understanding, and revision of attitude that may result from such efforts. Any punitive response to a fully responsible being, then, and it might be no more than the giving of an evil-tasting pill or some form of conditioning, that directly
in some causal way, with or without the agent's consent, sought to bring about a good, say, instantaneous truth or aversion to acting violently, would be incompatible with this constraint. There is, then, a good to be achieved but one cannot, logically or morally, be compelled to obtain it. Throughout there must be complete respect for the moral personality of the wrongdoer; it is a respect also, as I later argue, that must be given despite the wrongdoer's consent to be treated otherwise.

It is evident that this paternalistic goal is not to make people feel less burdened or more content. Once the good is achieved, these may be likely results; they are not, however, what is sought. It is important, too, to recognize that this good differs markedly from those particular goods associated with specific paternalistic legislation. It is not one's health; it is not even one's moral health with respect to any particular matter that is sought to be achieved; it is one's general character as a morally autonomous individual attached to the good.

III

What might be said in favor of such a theory and what might be objections to it? Two major issues will be considered. First, can a plausible case be presented that punishment is connected with the good as I have defined it? Second, is there anything morally offensive or otherwise objectionable, as there often is with particular legislation, in having as one's goal in limiting freedom, the person's own good?

Let us direct attention again to the relationship between parent and child with which I commenced this essay and in which paternalistic-like elements seem clearly and appropriately present. The range of situations here is very great. Sometimes parents coercively interfere to protect the child from hurting itself, sometimes to assure its continued healthy growth, sometimes so that the child will learn to move about comfortably in a world of social conventions. But sometimes, of course, coercion enters in with respect to matters that are moral; certain modes of conduct are required if valued relationships among individuals within the family and outside the family are to come into existence and be maintained.

Slowly such values as obedience, respect, loyalty, and a sense of personal responsibility are integrated into the young person's life. This results to a considerable degree—of course not entirely and in differing degrees in different stages of development—from the child's conduct sometimes meeting with unpleasant responses. Written vividly upon children are lessons associated with some loss or some pain visited upon them by those to whom they are attached. It is important for my purposes that a difference in the significance of the painful responses be noted. The pain experienced by the child subjected to a parent's anger or disapproval only has the significance of punishment if the parent deliberately visits upon the child some pain because of the perceived wrongdoing. The parent's spontaneous anger or disapproval or blame cause the child distress. They may motivate future compliant conduct. They may arouse in the child guilt. They are not, however, by themselves requital for wrongdoing and by themselves do not relieve guilt. My view is that punishment has some special and logical relationship to wrongdoing and to the possibility of a child's acquiring the concept. Because of this relationship, punishment is connected with the good that I have described in a way that blame or disapproval by themselves are not.

First, because of punishment children come to acquire an understanding of the meaning of a limit on conduct. Logically connected with the concept of wrongdoing is the concept of a painful response that another is entitled to inflict because of the wrongful conduct. Second, a punitive response conveys to children the depth of parental attachment to the values underlying the limit. Just as children know from experience that they are disposed to strike out when they or what they care for are injured, so they come to appreciate the seriousness of their parents' attachment to the limit and to the values supported by its existence by the parents' visiting some pain upon them. The degree of punishment, then, conveys to the child the importance parents attach to their child's responding to the limit and promotes in children, not just an appreciation that something is wrong, but how seriously wrong it is. It conveys, too, the significance of different degrees of fault in the doing of what is wrong. Further, particular punishments that are chosen often communicate to children the peculiar character of the evil caused by their disregard of the limit, the evil to others and the evil to themselves. Thus, even young children will find it particularly fitting to penalize a cheater by not permitting, for a time at least, further play, for such punishment conveys the central importance of honesty in the playing of the game and one's placing oneself outside the community of players by dishonesty. "If you will not abide by what makes this segment of our lives together possible, suffer the consequence of not being here a part of our lives."

Finally, punishment "rights the wrong." It has, in contrast to blame and disapproval, the character of closure, of matters returning to where they were before, of relationships being restored. Just as a limit being placed upon conduct serves to provide a bounded, manageable, world for the child, so the punitive response to a breach defines a limit to separation that is occasioned by wrongdoing. The debt is paid, life can go on.

The young hero in Styron's SOPHIE'S CHOICE gives into a desire for an exciting ride with a friend and forgets his agreeing to tend the fire before which his invalidated mother sits for heat in the freezing weather. The young man is guilty and remorseful. Why, we may wonder, was he grateful to his father for placing him for a period of time in a shed without heat? The answer seems clear. It diminished the young boy's guilt, diminished it in a way that it would not have been were the father merely to have said, "You did something dreadful; I know you feel bad; don't let it happen again!" The young boy's guilt and remorse were painful; but because they were not deprivations imposed because of wrongdoing, they could not serve to reestablish what had been upset in the relations between parents and child.

What I have described is familiar. What needs emphasizing is that this parental practice of punishing is a complex communication to the child. It aids the child in learning what as a moral person it must know, that some things are not permitted, that some wrongs are more serious than others, that it is sometimes responsible for doing wrong and sometimes not, and that its degree of blameworthiness is not always the same. Further, the child's response to wrongdoing by feeling guilt, its willingness to accept some deprivation, and its commitment to acting differently in the future, all play an indispensable role in its restoring relationships it has damaged, relationships with others and with itself. The claim, then, is that this practice is, in fact, a significant contributing factor in one's development as a moral person.

Now, what more acceptably motivates a parent when it punishes its child than the desire to achieve a goal such as I have described? It would be perverse if the parent were generally to punish primarily from motives of retributive justice or optimal utility for the family. These ends are secondary to, though with retributive ends, to some extent essential to, the child's acquiring the characteristics of a moral person. This much may seem plausible but also quite beside the point. The topic is, after all, punishment in the adult world and there are significant differences between adults and children that may carry fatal implications for a paternalistic theory. I do not believe this is so, but before moving on I want to note a phenomenon that may cast doubt upon the legitimacy of the parental practice itself.

Parents sometimes, when imposing some deprivation upon their children, say, "I'm only doing this for your own good!" There is, I think, something offensive about this. Does it affect the legitimacy of parental concern primarily for the child's moral development in inflicting punishment?

The answer I think is clearly "no," for the offensiveness of those words is not limited to situations in which punishment is imposed. Giving some unpleasant medicine or compelling the child to eat some distasteful but allegedly nourishing food, if accompanied by a statement that it is for the child's own good, is equally offensive. The words are customarily uttered in response to some sign of resistance, of some anger, and what they neglect to address is the child's unhappiness. They rather defend the parents before the child, making the child feel guilty because of its failure to be grateful for the good done it. And so imposed upon the child is the burden of getting what it does not want, the burden of checking its understandable anger because of this, and, finally, the burden of having to be grateful for getting what it does not want and, if not grateful, then guilty. It is not the motive of promoting the child's good that is suspect in these cases; it is communicating to the child what one's motive is, with its distressing consequences for the child, and with the still more serious problem, perhaps, that the parent's own guilt is unconsciously sought to be transferred to the child.

IV

One can acknowledge the place of punishment in the moral development of children and acknowledge, too, that it must to some degree be imposed to further this development and wonder what all this has to do with legal punishment of adults. For the law as a means of social control presupposes that the individuals to whom it applies are already responsible persons, responsible both in the sense of having the capacity to govern their actions through an understanding of the meaning of the norms addressed to them and responsible in the sense that they possess a knowledge of and an attachment to the values embodied in the society's laws. There is, nevertheless, a place for punishment in society analogous to its role in the family. I shall briefly sketch what this is.
Through promulgation of laws, through provision of sanctions for their violation, and through the general imposition of sanctions in the event of violation, each citizen learns what is regarded as impermissible by society, the degree of seriousness to be attached to wrongdoing of different kinds, and the particular significance—especially when the punishment is in its severity and character linked to the offense—of the evil underlying offenses. Punishment is a forceful reminder of the evil that is done to others and oneself. Were it not present, or were it imposed in circumstances markedly at odds with criteria for its imposition during the process of moral development, only confusion would result. Brandeis, in a quite different context, observed: “Our government is the potent, the omnipresent teacher. For good or for ill, it teaches the whole people by its example.” My point is that law plays an indispensable role in our knowing what for society is good and evil. Failure to punish serious wrongdoing, punishment of wrongdoing in circumstances where fault is absent, would serve only to baffle our moral understanding and threaten what is so often already precarious.

Further, our punitive responses guide the moral passions as they come into play with respect to interests protected by the law. Punishment, among other things, permits purgation of guilt and ideally restoration of damaged relationships. Punishment, then, communicates what is wrong and in being imposed both rights the wrong and serves, as well, as a reminder of the evil done to others and to oneself in the doing of what is wrong.

Now in addition to making out that punishment may reasonably be thought to play its part, even with adults, in promoting the good of one’s moral personality, the paternalist has to have some argument for this as a morally permissible way of proceeding. The paternalist is, I believe, on firm ground here. The guilty wrongdoer is not viewed as damned by his wrongful conduct to a life forever divorced from others. He is viewed as a responsible being, responsible for having done wrong and possessing the capacity for recognizing the wrongfulness of his conduct. Further, the evil—as Socrates long ago pointed out—that he has done himself by his wrongdoing is a moral evil greater than he has done others. His soul is in jeopardy as his victim’s is not. What could possibly justify an unconcern with this evil if the person is one of us and, if we sense, rightly I believe, that there but for the grace of God go we? In considering, for example, why we might wish to have a society of laws, of laws associated with sanctions for their violation, of laws that are in fact enforced against others and ourselves, it would be rational, indeed it would be, I think, among the most persuasive of considerations for establishing such a social practice, that it would promote our own good as moral persons. Thinking of ourselves as potential, and thinking of ourselves as actual wrongdoers, and appreciating the connection of punishment with one’s attachment to the good, to one’s status as a moral person, and to the possibility it provides of closure and resumption of relationships, would we not select such a system, if for no other reason, than that it would promote our own good?

V

We have now to consider certain objections to the theory. First, does it fail to respect one as an autonomous being? The answer is that it does not. One’s choices are throughout respected, and it is one’s status as a moral person that is sought to be affirmed. But is there not something offensively demeaning in instituting punishment for such a reason? More demeaning, one might ask, than addressing the wrongdoer’s sense of fear to which others appeal in their theories of punishment? More demeaning than an indifference to the moral status of the person but totally committed to retributive justice? I am not convinced that this is so either. On the theory I propose one is throughout responded to as a moral person.

But does not a paternalistic theory lead to two unacceptable extremes with respect to punishment, the first that we should always warn before punishing, and wait to see the effects of our warning, the other that we should continue punishing until we achieve the desired effects? The answers here can be brief. First, the announcement of the norm and the provision for punishment in the event of its violation is itself the warning and to allow a person to disobey and threaten that next time there will be punishment is to issue not one but two warnings. Second, the practice of punishment, given the paternalistic goals I have described, cannot permit open-ended punishments, repeated punishments or punishments that are excessively severe. For, first, the goal is not repentance at all costs, if that has meaning, but repentance freely arrived at and not merely a disposition toward conformity with the norms; secondly, the punishment provided for wrongdoing must reflect judgments of the seriousness of the wrong done; such punishment cannot focus on some end state of the person and disregard the potential for moral confusion that would arise from repeated or excessive punishment.
Another criticism might go as follows: "You have ruled out conditioning a person, even with their consent, so that they might not be disposed to do evil in the future. But surely, while perhaps an unjustifiable practice without consent, it is acceptable with it, for it provides a person what they freely choose and delivers them from an affliction that promotes evil." Two points need to be made here. First, the theory would not preclude freely chosen forms of conditioning, surgery and the like in those circumstances in which it is acknowledged that the person is not, with respect to the conduct involved, an autonomous agent. There is nothing wrong, for example, in a person choosing surgery to remove a tumor that is causally related to outbursts of violence over which the person has no control. The class of person, then, whose choice would be accorded respect is made up of those we should be disposed to excuse from criminal liability. Second, the theory would regard as morally unacceptable a response, conditioning or otherwise, that had as its goal, not just aversion to doing wrong, but obliteration of one's capacity to choose to do so. What must be aimed at is that the afflicted become autonomous not automatons. There must be freedom to disobey, for the moral price is too high that is paid in purchasing immunity from temptation and guaranteed conformity.

The most troubling objections to the theory are, I think, these: First, it cannot account for the accepted disposition to punish those who are already, as it were, awakened and repentant. And, second, even more seriously, it cannot account for the disposition to punish those who know what the values of society are but who are indifferent to or opposed to them. Someone, for example, may feel inclined to say: "Look—most serious crimes—and your theory surely most neatly fits such crimes not petty offenses—are committed by individuals who are perfectly aware of what they are doing and perfectly aware that society's values are being flouted. These individuals are not going to be instructed about evil or brought to any moral realization about themselves by punishment. Surely, you can't be serious about repentance when considering them, and they certainly do not care a jot about paying off any debt because they do not feel any guilt over what they have done. Your theory fails so to match reality as to be just one more tedious example of a philosopher spinning out fantastic yarns without any genuine relevance to reality." What can be said in response to these points?

As to the first, I would claim that the guilty and repentant wrongdoers are naturally disposed to accept the appropriateness of the punishment provided, both because this will evidence to them and to others the genuineness of their feelings and because the punishment rights the wrong, brings about closure and restores relationships that have been damaged. The experience of guilt and remorse, the avowal of repentance do not by themselves achieve this. A general practice of pardoning persons who claimed that they were repentant would destroy the principal means of reestablishing one's membership in the community.

Now for the second major objection. A response here requires that attention be paid to certain general features of the theory that has been put forward. The theory is, of course, not intended as a description of any actual practice of legal punishment or even as realistically workable in a society such as ours. Things are in such a state that it is not. What is proposed is a moral theory of punishment and, as such, it includes at least two conditions that may be only marginally congruent with our social world. The first is that the norms addressed to persons are generally just and that the society is to some substantial extent one in which those who are liable to punishment have roughly equal opportunities to conform to those just norms. The second condition is equally important. The theory presupposes that there is a general commitment among persons to whom the norms apply to the values underlying them. If these two conditions are not met, we do not have what I understand as a practice of punishment for which any moral justification can be forthcoming.

At this point it may be thought, "fair enough, but then what is the point of the whole exercise?" My response is this: First, the theory is not without applicability to significant segments of our society. Second, it has value, for it provides an important perspective upon actual practices; it throws into relief our society's failures to realize the conditions I have stipulated. And, finally, it assists us in sensitive and intelligent forbearance from putting our moral imprimatur upon practices which the paternalistic model would find unacceptable. Excessively lengthy prison terms and the inhumane conditions under which they are served, for example, can be effectively criticized with a clear conception of the good defined by the paternalistic theory. The theory may serve as a guide in our attempts to adjust present practices so

---

that they more closely accord with moral dictates, to work for precisely that society in which the paternalistic conception provides not just the ring of moral truth but descriptive truth as well.

VI

I want now to shift attention to the issue of restrictions on punishment. The proposed paternalistic theory limits punishment, I believe, in a way that accords more closely with our moral intuitions than a number of alternative theories. First, it follows from the theory that any class of persons incapable of appreciating the significance of the norms addressed to them cannot justifiably be punished. Absence of a free and knowing departure from the norm makes pointless imposition of punishment. Second, it also follows that excuses must be recognized and that mitigating factors be taken into account, including as an excuse, of course, reasonable ignorance or mistake of law.

Perhaps most significantly, a paternalistic orientation implies a position that matches our moral intuitions more closely than other theories on the issue of what kinds of punishment may be inflicted. Punishments that are aimed at degrading or brutalizing a person are not conducive to moral awakening but only to bitterness and resentment. But there is also, I believe, another paternalistic route to limitations upon certain modes of punishment, a limitation that follows from the conception of the moral good.

The wrongdoer has, as we all do, a basic right to be free. How, we may wonder, are we able to justify our imposing our will upon him and limiting his freedom? One answer is that by wrongful conduct he has forfeited his right to freedom. The wrongdoer is in no position to complain if he meets with a response that is similar to what has been visited by him upon another. Such a theory of forfeiture places great weight upon an individual's choice. It holds that rights are forfeitable, waivable and relinquishable—just so long as the choice involved is informed and free. A person might forfeit his right to life by murdering; a person might relinquish his right to be free by selling himself into slavery. The paternalistic position that I have proposed holds otherwise. It implies that there is a non-waivable, non-forfeitable, non-relinquishable right—the right to one's status as a moral being, a right that is implied in one's being a possessor of any rights at all.

Such a view, when punishment is at issue, makes morally impermissible any response to a person, despite what that person has done, that would be inconsistent with this fundamental right, even though the person were unattached to it, indifferent to its moral value and eager to forfeit it. A retributivist might respond in kind to any wrong done. A social utilitarian might calculate the effects on people and society in doing so. A paternalist, attached to the good of the wrongdoer, would reject retributive justice and utility as the sole determinative criteria, and would propose a good to be realized that is independent of these values. Punishment will not be permitted that destroys in some substantial way one's character as an autonomous creature. Certain cruel punishments, then, may be ruled out, not merely because they are conducive to hardening the heart but, more importantly, because they destroy a good that can never rightly be destroyed. As I see it, this precludes, on moral grounds, punishment that may be like for like but which nevertheless violates one's humanity by either destroying one's life or destroying one's capacity for rejecting what is evil and again attaching oneself to the good.

Let me be more specific. Suppose that a sadist has cruelly destroyed another human being's capacity for thought while leaving the person alive. Is there a retributivist argument that would bar a like treatment for the sadist? I do not know of it. Certainly, the lex talionis would seem to sanction it. Is our inclination to forbear from treating the sadist in a manner that he has treated his victim derived exclusively, then, from social evils that we foresee might flow from such punishment? I do not find this persuasive. Our moral repugnance precedes such calculation and findings inconsistent with this repugnance would be rejected. Is it simply revulsion at the thought of oneself or one's agents deliberately perpetrating such acts? Is it a concern for our own good that motivates us? No doubt, this may play a role, but my conviction is that something else is involved. It is the ingredient to which the moral paternalist draws attention. The wrongdoer possesses something destroyed in another. The wrongdoer may desire to destroy it in himself as well, but that is not his moral prerogative. It is immune from moral transformations brought about by free choice.

VII

I would like, in conclusion, to make somewhat clearer what I am and am not claiming for the theory proposed in these pages and, further, to draw attention to two ironies connected with it.

I have claimed that to have as one's aim in punishing the good of the wrongdoer counts strongly
in favor of the moral legitimacy of punishing. I do not claim, of course, that this is the sole justification for punishment, though I do believe that what it seeks to promote is among the most important, if not the most important, of human goods. The practice of punishment is complex and any justification proposed as an exclusive one must, in my judgment, be met with skepticism, if not scorn. There is, too, as I earlier briefly noted, a significant logical overlapping of this theory with retributivism, though at a certain point, when one considers types of punishment, they diverge. A paternalistic theory, given the good as defined, would support principles that are familiar dictates of retributivism—that only the guilty may be punished, that the guilty must be, and that the punishment inflicted reflect the degree of guilt. Failure to comply with the demands of retributivism would preclude realization of the paternalist’s goal. I have also, however, suggested that retributivism needs supplementing if it is to meet our intuitions of what is morally permissible punishment. But, of course, this overlapping of justifications for punishment includes as well some form of utilitarianism, for if our goal is as I have defined it, and punishments are threatened and imposed, deterrent values are also furthered. I do not question the rich over-determination of goods promoted by the practice of punishment. I do urge that weight be given, and on the issue of restrictions on punishment, determinative weight, to paternalistic ends.

There are, finally, two ironies to which I wish to draw attention. The first is this. I have selected as the good to be realized by this paternalistic theory of punishment the very good to which philosophers often make appeal in their principled objections to paternalism with regard to specific prohibitions and requirements. Secondly, I have proposed a theory that justifies forceful intrusion into the lives of people. But it is also an atypical paternalistic theory, for it prohibits certain types of intrusion. I reach this conclusion because the good sought does not allow weight to be given to an individual’s free choice when the issue is relinquishment of one’s status as a moral being. The paternalistic aspect in this derives from the fact that there is a good for the person to which we are attached, though the person might not be, and which we continue to respect in disregard of the usual consequences of a person’s free choice. I would guess that something like these thoughts underlies the view that we possess some goods as gifts from God and that it is not within our moral prerogative to dispose of them. It is easy to suppose, but a mistake nevertheless, that because we may be favored by the gods that we are one of them.

*University of California, Los Angeles*

*Received January 15, 1981*