PHILOSOPHY RESEARCH ARCHIVES
VOL. X, MARCH, 1985

11. ANOTHER LOOK AT MORAL BLACKMAIL

LAWRENCE ALEXANDER
UNIVERSITY OF SAN DIEGO

ABSTRACT. In this paper I describe cases of moral blackmail as cases where A is told by B that if A does not commit an otherwise immoral act, B will commit an immoral act of equal or greater gravity. I describe cases of moral dilemma as cases where A must commit an otherwise immoral act to avert a natural disaster of equal or greater gravity. I then argue that cases of moral blackmail are structurally identical to cases of moral dilemma in all respects but one: In cases of moral blackmail, A is predicting the free actions of a moral agent (B), whereas in cases of moral dilemma, A is predicting natural events. I conclude that cases of moral blackmail are more problematic than otherwise similar cases of moral dilemma for this reason alone.

The scene is a familiar one to readers of moral philosophy, one made famous by Bernard Williams.[1] You are visiting a Latin American country whose government has oppressed various Indian tribes. As you enter a remote village, you encounter government troops and their twenty Indian captives. The commander of the troops is delighted to see you, and he informs you that the Indians, whom he was about to execute, will be spared in your honor if only you would execute one of them yourself. By your standards, none of the Indians deserves execution. You are convinced, however, that if you decline the "honor," the commander will indeed execute all the Indians. What are you morally permitted to do? What are you morally required to do?

Williams' problem is the problem of moral blackmail. A is told by B that if A does not commit what would otherwise be an immoral act--say, kill innocent person C--B will commit an immoral act of equal or greater gravity--say, kill C and D. There are two variants of moral blackmail that are worth distinguishing, which I shall call Moral Blackmail One (MB1) and Moral Blackmail Two (MB2).

MB1: A must inflict an otherwise morally unjustifiable harm on C, or B will unjustifiably inflict the same or greater harm on C and D (says B to A).

MB2: A must inflict an otherwise morally unjustifiable harm on C, or B will unjustifiably inflict the same or greater harm on D and E (says B to A).

The difference between MB1 and MB2 is that in MB1, A's victim, C, will also suffer equal or greater harm if A resists the blackmail and B carries out his threat, but in MB2, C will escape harm if A resists.

Now in order to analyze what A is permitted or required to do in MB1 and MB2, we must consider the situation in which the moral blackmailer B--Williams' commander of the government troops--is replaced by some threatening natural phenomenon. Suppose, for example, on your Latin American visit, you spy a huge boulder hurtling down a mountainside toward an encampment of twenty Indians, all of whom appear destined to be crushed if you do nothing. You notice, however, that if you were to shoot one of the Indians, causing him to fall, the boulder would upon hitting him skip over the other Indians. Let us call this a case of moral dilemma (MD), and let us structure the two variants of moral dilemma to parallel our two variants of moral blackmail.

MD1: A must inflict an otherwise morally unjustifiable harm on C, or nature will inflict the same or greater harm on C and D (so A believes).

MD2: A must inflict an otherwise morally unjustifiable harm on C, or nature will inflict the same or greater harm on D and E (so A believes).

One further distinction is necessary before we can understand the significance of moral blackmail, a distinction of moral frameworks. I shall call strong libertarianism (SL) the polar moral position that (1) rejects any enforceable duty to aid another, no matter how little sacrifice the aid requires and no matter how dire the circumstances of the one to be aided, and (2) rejects the permissibility of harming one person without his permission in order to aid another, no matter how small the harm to be inflicted, and no matter how great the harm to be averted. (1) and (2) are integrally related, if not strictly entailed by each other. We are not permitted by SL to touch C's brow without his consent in order to save D and E from death (per (2)), because C has no enforceable duty to allow his brow to be touched for that purpose (per (1)).

At the pole opposite SL is strong egalitarianism (SE). Under SE, everyone has an enforceable duty to sacrifice and submit to harm in order to avert a greater harm to others, or perhaps an equal harm to a greater number of others. All of us may have to compensate the one harmed if possible--which amounts to spreading the harm so that it falls equally upon the one harmed, the one saved, and all other persons. Thus, though we may and indeed must touch C's brow to save D from a worse fate, all of us, including D, must compensate C so that he ends up no worse off (but no better off) than the rest of us.

SL and SE are extreme moral positions. Any number of moral positions intermediate between them can be defined. One might, for example, vary SE by deciding that numbers don't count when we can choose between saving one and saving many from the same harm.[2] Or we may vary SE by setting a limit on the harm we can inflict on anyone, no

matter how great the harm to be averted by inflicting harm, and no matter how many persons face that harm.

Let us now look again at MD1 and MD2, this time from within the moral frameworks of SL and SE. With respect to MD2, SL and SE provide clear answers. SL forbids us to inflict harm on C to save D and E; SE requires us to inflict harm on C to save D and E (if we ourselves will not suffer greater harm). With respect to MD1, SE again requires us to inflict harm on C. SL, however, does not *require* us to do anything, although it may *permit* us to inflict harm on C if C will suffer the same harm at about the same time. (The strongest version of SL would not permit us to inflict harm on C unless C would suffer the same harm at the same time no matter what we did.[3])

Our moral frameworks then help us deal with the moral dilemmas.[4] The question now becomes, will our moral frameworks handle the parallel cases of moral blackmail in exactly the same way? In other words, will the solutions SL and SE provide for MD1 and MD2 be the same solutions they will provide for MB1 and MB2? The only factor that distinguishes MB1 and MB2 from MD1 and MD2 is that in the cases of moral blackmail, the threat of harm comes from a moral agent, whereas in the cases of moral dilemma, the threat of harm is not such an agent. Therefore, unless moral agency makes a difference, moral blackmail should be treated as merely an instance of moral dilemma and subject to the same moral assessment.[5] I would argue that what is significant and interesting about moral blackmail is precisely this element of human agency that distinguishes moral blackmail from moral dilemma.

To show this, I ask the reader to consider what it is that seems troubling about Bernard Williams' hypothetical. Is it that we are being asked to kill? Surely, however, no matter how much we might wish that no had to die, it is better that only one die rather than all twenty.

Is what is troubling that we are not totally certain that killing one will save twenty, or certain that not killing one will doom the rest? Surely lack of certainly is troubling when the stakes are so high; but lack of certainty is present in many grave moral dilemmas. Killing in self-defense can never be said for certain to have averted death or great bodily harm. And lack of certainty cuts both ways. Your attempt to kill the Indian may be thwarted by some *deus ex machina* that saves your victim and all the other Indians as well.

Williams suggests that what is troubling is the interruption of your projects, the life you have chosen to lead. Even strong libertarians, however, face interruptions of their projects when the projects collide with the moral rights of others.[6]

Perhaps what is troubling is that killing the Indian may threaten our protagonist with moral corruption, a weakening of his moral fiber. In what sense? Surely not in the sense that the killing is morally wrong. That would beg the question at issue. Perhaps in the sense that killing the Indian weakens the moral disposition to avoid killing. Of course, the fact than an act weakens a moral dispositions *is* a reason that counts against undertaking the act; it is not, however, a conclusive reason. Given the crude nature of dispositions, their insensitivity to nuances among acts of a general type, no realizable set of moral dispositions will ever fit entirely harmoniously with the set of morally optimal

acts, even taking into account the effects of acts on dispositions. If killing the Indian is otherwise morally optimal, it may remain so even when account is taken of the effect of the killing on the actor's and others' moral dispositions.

We may be troubled by moral blackmail because we hold strong libertarian views, given which our protagonist has no enforceable duty to kill the Indian and indeed, under MB2 at least, has no right to kill the Indian. But, as I have said, moral blackmail and moral dilemma are no different with respect to their implications for libertarians.

We are left, then, with only one factor that explains why cases of moral blackmail are more troubling than parallel cases of moral dilemma. In cases of moral blackmail, we are predicting the actions of a moral agent, and we find such predictions to be problematic bases for harming third parties, even when we are as certain of them as we are of our predictions of natural events. Moral agents retain the power to falsify our predictions in a way that natural events cannot. No matter how certain we are that a moral agent will act in a particular way--no matter how much our evidence supports our prediction--the moral agent's freedom makes our prediction different in kind from a prediction of natural events, even if the prediction of natural events is statistically less certain than the prediction of the moral agent's actions.

As evidence that we are more troubled by inflicting harm on the basis of predictions about moral agents than we are troubled by inflicting harm on the basis of predictions about natural events, consider the following cases. In Case 1, employer says to job applicant, "Your heredity suggests a 50% chance of contracting genetic disease X. Because you are only marginally better qualified than other applicants, we don't judge it to be sound business practice to hire you, given the chance of disability." In Case 2, employer says to job applicant, "People with your characteristics (say, married females) *choose* 97% of the time to leave this type of job within five years (say, in order to accommodate their husbands' careers). Because you are only marginally better qualified than other applicants, we don't judge it to be sound business practice to hire you, given the risk of your leaving soon." Case 2 bothers us in a way Case 1 does not, even though the predictions in both cases are well-founded, and the employer's appraisal of the expected economic return is unimpeachable. In Case 2, harm is being inflicted on a moral agent because of a prediction about her future choices, a prediction she has the power to falsify (although she will not in fact do so 97% of the time). We are bothered in Case 2 because the job applicant is treated like a natural object, not like a being with freedom.[7]

Whether we *should* be bothered more by Case 2 than Case 1--whether we should be less disposed to inflict harm on the basis of predictions about moral agents--I cannot resolve here. Even if well-settled social practices count as evidence of what is a correct moral stance, our practices are inconclusive on the point in question. We condone self-defense, even where the aggressor is a responsible moral agent and the defender has acted on the basis of a prediction, though it is unclear whether we regard such self-defense as justified or merely as excused. On the other hand, we generally permit preventive detention of those who we predict will endanger others only when we believe they are sufficiently mentally impaired as to escape full legal and moral responsibility for their acts. My point here is the modest one that cases of moral

blackmail are more troubling than parallel cases of moral dilemma because they involve such predictions about moral agents, and for no other reason. My point is not that such predictions are in fact morally problematic.

Terrance McConnell recognizes the distinction between moral blackmail and moral dilemma; but *he* finds significant the fact that in cases of moral blackmail, persons are threatened with violations of their rights by the blackmailer, while in the cases of moral dilemma, no one has a right against natural catastrophes.[8]

I confess that I cannot make sense out of McConnell's distinction. It is true that the blackmailer is threatening to violate rights, whereas nature cannot violate rights. Why that should implicate an intervenor in the blackmailer's violation is for me a mystery, *unless it is because the blackmailer's presumed moral agency makes the prediction of his acts different in kind from the prediction of natural catastrophe.* Beyond this reason, which McConnell never mentions, threats to violate rights are equivalent to threats of natural catastrophe *from the intervenor's perspective.* The question is whether the intervenor violates rights in cases of moral blackmail that he does not violate in cases of moral dilemma. The fact that the blackmailer is violating rights is beside the point, except insofar as it makes the intervenor's prediction of harm different in kind from his prediction of natural events.

There is one possible implication--beyond the difference in type of prediction--of the fact that cases of moral blackmail, but not cases of moral dilemma, involve threats to rights. If the (would-be) intervenor refrains from intervening in a case of moral blackmail, and the threat is carried out, there will be someone besides the (would-be) intervenor--the blackmailer--to blame and perhaps to sue for the harm. If, on the other hand, the (would-be) intervenor refrains from intervening in a case of moral dilemma, nature cannot be held morally responsible for the resulting harm.

I submit that this distinction, unlike the distinction between types of prediction, is irrelevant to the analysis of the (would-be) intervenor's rights and duties. It is irrelevant because, although all cases of moral blackmail involve threatened violations of rights, *some cases of moral dilemma also involve threatened violations of rights.* This fact is obscured by portraying moral dilemmas as though they always involve purely natural catastrophes such as rock falls. Some cases of moral dilemma are cases in which a morally responsible human agent---one who *can* be held responsible for his or her violations of rights--has unleashed a threatening force that can be deflected from its course by the intervenor, though at some (lesser) cost to innocent persons. Thus, if a moral blackmailer threatens to push a boulder at our innocent Indians and then carries out his threat, the case becomes one of moral dilemma if the intervenor can save most of the Indians by killing a few. If the (would-be) intervenor does not intervene, the blackmailer can still be held morally responsible for the deaths. But the blackmailer's responsibility does not necessarily extinguish the responsibility of the (would-be) intervenor. As Robert Nozick has pointed out, moral responsibility for an event is not an item that, when divided among several persons, must never sum to greater than one.[9]

Thomas Hill, writing about cases of moral blackmail, writes:

> [T]he response we might expect from a morally sensitive person [faced with moral blackmail] . . . would typically include the following: (1) concern about whether one correctly estimated the necessity and effectiveness of doing the lesser evil; (2) regret that one did not relate to others in the ways one wants to; (3) regret that one could not live a pure and integrated life, expressive of one's deepest values; and (4) anger or resentment towards those whose ill will forced one to make the choice. Each of these, I suggest later, points to a moral consideration against thinking that one is always, or easily, justified in doing the lesser evil.[10]

Points (2) and (3) apply with as much force to cases of moral dilemma as to cases of moral blackmail. And point (4), while it applies only to moral blackmail, suggests nothing about the justifiability of the agent's compliance with the blackmailer's terms. That leaves only point (1), our faith in our prediction of the blackmailer's future choices, as the only possible basis for contrasting moral blackmail and moral dilemma.

Nancy Davis considers the possibility that moral blackmail, unlike moral dilemma, involves *using* the blackmailed agent, and that this fact justifies a moral distinction between the cases.[11] However, as she points out, if the blackmailer unleashes a potential natural catastrophe, thus creating a case of moral dilemma, the agent who intervenes and commits a lesser evil to avert the catastrophe has been "used" just as much as in the case of moral blackmail. Resistance to being used by immoral agents cannot distinguish moral blackmail from moral dilemma.

The problem of moral blackmail, therefore, as distinct from the problem of moral dilemma, lies solely in the relevance, if any, of predictions of moral agents' actions. If the distinction between predictions of moral agents' actions and predictions of natural events is morally irrelevant, then cases of moral blackmail are just cases of moral dilemma, no more, no less. Several writers have correctly identified moral blackmail as a significant moral issue. Unfortunately, no one has both identified and dealt with the precise element in moral blackmail that makes it significant.[12]

ENDNOTES

[1] J. Smart and B. Williams, *Utilitarianism: For and Against* 98-99 (Cambridge University Press: 1973).

[2] See, e.g., Taurek, "Should the Numbers Count?," *Philosophy and Public Affairs* 6 (1977): 293-316.

[3] Terrance McConnell argues that moral blackmail should be governed by the following principle:

> An agent in a situation of moral blackmail may do something evil in order to prevent the blackmailer from doing a greater evil just in case either (1) the agent's action does not irreparably violate the rights of some nonconsenting, inno-

> cent person, and if an innocent person's rights are violated he is to be compensated, or (2) if a nonconsenting, innocent person's rights are irreparably violated, then the same (or equally strong) rights of the same person would have been violated even if the agent did not comply with the blackmailer's demand.

McConnell, "Moral Blackmail," *Ethics* 91 (1981): 544, 551. McConnell's principle, as a principle for handling moral dilemmas, is close to SL, though unlike SL, McConnell would allow infliction of *reparable* harm.

[4] Dan Brock has argued that because under SL the potential victim's consent would make killing one of them permissible (though not required), and because the victims will rationally consent to any proposal that increases their chances of survival, the case for SL over SE is undermined. That is so because failure of someone to consent to a strategy like SE that maximizes the number of lives saved can only be accounted for by that person's knowledge of his own relative chances under SL and SE, knowledge that should be considered morally irrelevant. Brock, "Moral Prohibitions and Consent," in M. Bradie and M. Brand, eds., *Action and Responsibility* 111 (Bowling Green State University: 1980).

[5] Both Dan Brock and Nancy Davis argue that moral blackmail and moral dilemma should be regarded the same way from the standpoint of the agents faced with them. Brock, *Ibid.*, at 113, 120-21n.6; Nancy Davis, "The Priority of Avoiding Harm," in B. Steinbock, ed., *Killing and Letting Die* 202-08 (Prentice-Hall, Inc: 1980).

[6] See Hollis, "Jim and the Indians," *Analysis* 43 (1983): 36, 38-39; Piper, "A Distinction Without a Difference," in P. French, *et al*, eds., *Midwest Studies in Philosophy VII* 403, 427-28 (University of Minnesota Press: 1982).

[7] See Underwood, "Law and the Crystal Ball: Predicting Behavior with Statistical Inference and Individualized Judgment," *Yale L.J.* 88 (197): 1408. See also Alexander, "Retributivism and the Inadvertent Punishment of the Innocent," *Law and Philosophy* 2 (1983): 233, 244-45.

[8] McConnell, *supra* note 3, at 560-62.

[9] R. Nozick, *Anarchy, State, and Utopia* 130 (Basic Books: 1974). See also Hollis, *supra* note 6, at 38.

[10] Hill, "Moral Purity and the Lesser Evil," *The Monist* 66 (1983): 213, 222.

[11] Davis, *supra* note 5, at 202-07.

[12] Alan Donagan, at 207-09 in *The Theory of Morality* (The University of Chicago Press: 1977), can be read as suggesting that the significance of Bernard Williams' hypothetical lies not in its implications for consequentialist moral theories, but rather in its implications for treating predictions of human actions like predictions of natural events. Unfortunately, Donagan goes no further along this, the correct line of analysis, than to assert the absence of a moral duty to save innocents from threatened wrongdoing.

Nancy Davis considers but rejects the possibility that moral blackmail is morally distinguishable from moral dilemma because the former involves a prediction of an agent's free choice. Davis, *supra* note 5, at 207–08. Davis maintains that we can predict what will happen with equal warrant in both kinds of cases. And she concludes that beyond its effect on our confidence in our prediction of what will happen, the fact that moral blackmail involves a prediction of how a free agent will act is morally irrelevant.

If Davis is correct, then she solved the problem of moral blackmail by reducing it to the problem of moral dilemma. But is she correct? That is, is the relevance of the fact that the prediction in a case of moral blackmail is a prediction of a free choice exhausted by its effect on our confidence in the prediction?