

## On the Moral Objection to Coercion

### I. THREE INTUITIONS

Suppose Green threatens Brown: “Stay out of Malibu, or you’ll be sorry.” And Brown has every reason to believe he will indeed be sorry if he shows up in Malibu again. And so Brown stays out of Malibu. Most of us will think that, *prima facie*, Green has done something wrong. There is perhaps some background we could fill in that would make it permissible for Green to issue the threat. But in the absence of some special justification, this type of coercion is objectionable.

But what exactly has Green done to Brown, so far, that one might object to? Of course, it would be wrong for Green to carry out her threat and, say, beat up Brown. But Green hasn’t done that yet and, in fact, won’t do that, because Brown, we can suppose, will stay out of Malibu.

In the literature on the topic, we find an interesting divergence of emphasis. Different accounts focus attention on different aspects of

I owe thanks, first and foremost, to A. J. Julius, with whom I have been discussing ideas in this article since the first Obama administration. For helpful discussion or written comments on earlier drafts, I thank Scott Anderson, Arudra Burra, Eric Chwang, Jerry Dworkin, Kyla Ebels-Duggan, Guy Elgat, Barbara Herman, Louis-Philippe Hodgson, Rob Hughes, Niko Kolodny, Cristina Lafont, Arthur Ripstein, Nishi Shah, Seana Shiffrin, Sarah Stroud, Steve Sverdlik, Will Tadros, Julie Tannenbaum, Ekin Zeytinoglu, and Rachel Zuckert. Ancestors of this article were presented at the Rocky Mountain Ethics Congress at the University of Colorado, Boulder; the Northwestern University Society for the Theory of Ethics and Politics conference; a workshop on lying, coercion, and deception held at the Maguire Center for Ethics and Public Responsibility at SMU; and a workshop on recent trends in the philosophy of coercion, held at the University of British Columbia. Thanks to all those who were present at those events for their feedback. Finally, I am grateful to the Editors of *Philosophy & Public Affairs* for extremely valuable comments and criticisms.

the coercive interaction. We can distinguish three main approaches to explicating the moral objection to coercive threats.<sup>1</sup>

The first approach, probably the most familiar, focuses on the victim's will and the way in which it is interfered with or undermined by a credible threat (or even, in some cases, by an offer). On this approach, the core objection to this mode of influence is found in the idea that, when the recipient succumbs to a coercive threat, she does whatever she does against her will—her choice is in some way unfree or not fully autonomous. Understood in this way, coercion is a matter of the coercer's subverting, or circumventing, the victim's agency. He interferes with his victim's ability to make up her own mind about what to do, substituting his own will for hers.<sup>2</sup>

Although this perhaps seems the most natural—even obvious—line to take, there is a danger of falling into tautology. It is no doubt true that Brown's choice to stay out of Malibu is not a free or autonomous choice. But if our primary reason for thinking this is that Brown was coerced into staying away, then although this may be an important fact about our concepts of freedom and autonomy, we will have merely located the moral objection to coercion in the fact that it is coercive. Hopefully a more illuminating account is available. The question for this first approach, then, will be whether there is a suitably

1. Unless otherwise noted, my focus in this article will be on the act of threatening someone in order to influence his or her behavior. Paradigmatically, this will involve communicating to the person (the recipient) that unless he or she does or refrains from doing something, the issuer of the threat (the coercer) will visit some harm on the recipient or, more generally, bring about some consequence the recipient does not want. I will not undertake a full analysis of the concept of a threat, or of coercion. I will, however, restrict the terms *coercion* and *coercive threat* to those cases that are presumptively wrongful—that is, to those threats that attract a (perhaps overrideable) moral objection. This is purely to simplify the exposition, since much of my discussion will concern the question of what distinguishes perfectly permissible threats from objectionable ones.

2. For some examples of writers who endorse this broad approach, or analyses of coercion that suggest it, see Michael Blake, "Distributive Justice, State Coercion, and Autonomy," *Philosophy & Public Affairs* 30 (2002): 257–96; G. A. Cohen, "Capitalism, Freedom, and the Proletariat," in *The Idea of Freedom*, ed. Alan Ryan (New York: Oxford University Press, 1979); Harry Frankfurt, "Coercion and Moral Responsibility," in *The Importance of What We Care About* (Princeton, N.J.: Princeton University Press, 1988); Onora O'Neill, "Between Consenting Adults," in *Constructions of Reason* (Cambridge: Cambridge University Press, 1989), and "Which Are the Offers You Can't Refuse," in *Bounds of Justice* (Cambridge: Cambridge University Press, 2000); and Joseph Raz, *The Morality of Freedom* (New York: Oxford University Press, 1986).

independent conception of freedom or autonomy, such that the claim that coercion undermines freedom or autonomy is nontrivial, but still true.

A second approach places the emphasis not so much on interference with the coercer's capacity for autonomous choice, but on what it is that she is led to do in response to the threat. Here, the objection is to the fact that successful coercion brings a person to pursue some end or purpose she does not see any good, independent reason to pursue. When the mugger threatens to shoot her victim if he doesn't give up his money, for example, she is attempting to get him to do something he would otherwise—that is, but for the threat itself—see no good reason to do.<sup>3</sup> Wrongful coercion, on this view, is inconsistent with the recognition that our reasons for action should be coordinate with and constrained by the independent reasons other people have for their actions.

The first two approaches focus on the coercee's point of view as an agent, pointing to different respects in which it is allegedly distorted. The third approach starts from a quite different intuition about the morality of coercion. This is that, in a wide range of cases, the deontic status of a threat—whether it's permissible or impermissible—seems to depend on the deontic status of the act being threatened.<sup>4</sup> For example, barring unusual circumstances, there is nothing objectionable about a coach threatening her players with having to run laps if they are late to practice. And a prosecutor may, under normal circumstances, induce a defendant to plead guilty to one crime by threatening to indict her for other crimes the prosecutor has good reason to believe she is guilty of

3. For a recent development of this sort of account of the wrongfulness of coercion, see A. J. Julius, "The Possibility of Exchange," *Politics, Philosophy, and Economics* 12 (2013): 361–74, and *Reconstruction* (Princeton, N.J.: Princeton University Press, forthcoming). See also Gideon Yaffe, "Indoctrination, Coercion and Freedom of Will," *Philosophy and Phenomenological Research* 67 (2003): 335–56.

4. Other theorists who have noted this deontic connection include Mitchell Berman, "The Normative Functions of Coercion Claims," *Legal Theory* 8 (2002): 45–89; Vinit Haksar, "Coercive Proposals (Rawls and Gandhi)," *Political Theory* 4 (1976): 65–79; Japa Pallikkathayil, "The Possibility of Choice: Three Accounts of the Problem with Coercion," *Philosophers' Imprint* 11 (2011): 1–20; T. M. Scanlon, *Moral Dimensions: Permissibility, Meaning, and Blame* (Cambridge, Mass.: Harvard University Press, 2008); James Shaw, "The Morality of Blackmail," *Philosophy & Public Affairs* 40 (2012): 165–96; and Alan Wertheimer, *Coercion* (Princeton, N.J.: Princeton University Press, 1988).

as well.<sup>5</sup> In cases like these, where we have a permissible threat, it is, it seems, no accident that the course of action being threatened is itself permissible.

The challenge for theories that take this third approach is, first, to explain the apparent connection between the moral quality of the threat and the moral quality of what is being threatened. But they will also need to say something about cases of wrongful coercion that appear not to fit the general pattern—specifically, cases in which it seems wrong to threaten to do what one would otherwise be permitted to do.

The three approaches just outlined correspond to three different intuitions about the nature of a coercive threat. The first is that the choice to submit to the coercer's proposal is not a free or autonomous one. When a person succumbs to a threat, she is in some sense brought to act against her will. The second is that the coercer fails to show due consideration for the possibility that the coerced has good reasons for not behaving as the coercer wants. And the third is that it typically matters in some way whether the coercer has any right to do what she is threatening to do should the victim fail to comply with her demands.

A good theory of the wrong of coercion should provide some account of each of these intuitions. Existing views have trouble doing this, however. In particular, approaches to the topic that focus on the autonomy-undermining or incentive-creating aspects of coercive proposals have difficulty explaining why it should ever make a difference, in itself, that what is being threatened is morally prohibited. On the other hand, accounts that focus on the impermissibility of the threatened conduct have difficulty providing a satisfactory account of how the objection to coercion is related to the sense that the coerced's choices are constrained in the face of a credible threat.

5. Compare the Supreme Court's decision in *Bordenkircher v. Hayes*, 434 U.S. 358 (1979) (cited in Wertheimer, *Coercion*, p. 133). The trial prosecutor offered Hayes a plea of five years in prison if he pled guilty to forgery and threatened that if Hayes rejected the plea deal, he would be indicted in addition under the Kentucky Habitual Criminal Act, which applied since Hayes had already been convicted of two prior felonies, and which carried a sentence of life in prison. The court held that the prosecutor's threat was constitutionally permissible on the grounds that it was permissible for him to seek the additional indictment.

In this article, I develop an interpretation of what happens in a coercive interaction that I believe helps to illuminate all three of the above intuitions. The core idea is that, in order to induce the coercee to do something she would otherwise be inclined not to do, the coercer imposes on the coercee the responsibility for ensuring that the coercer does what she in any case ought to do. Thus, for instance, when Green threatens Brown, she makes it clear to Brown that whether she beats him up depends on the particular choices he makes. In this sense, it's now up to Brown to ensure he's not assaulted by Green. Obviously, this is not a responsibility Brown should have to bear.

I will argue that this responsibility-shifting aspect of coercive threats is crucial to understanding why it often makes a difference that the coercer is threatening to do something that is itself impermissible. But it also helps make sense of why we should not in fact expect a strict correspondence between the (im)permissibility of a threat and the (im)permissibility of what is being threatened. This is because threatening to do what one has no right to do is not the only way of illegitimately foisting responsibility onto someone who should not have to bear it. Further, the idea that coercion involves an illicit transfer of responsibility from coercer to coercee can be used to explicate the intuitions motivating the first two approaches as well. This way of viewing coercion helps to clarify the sense that the coercer, on the one hand, fails to have adequate regard for what the coercee herself has reason to do and, on the other hand, fails to respect the coercee's right to freely decide for herself what to do under the circumstances. Or so I will argue.

The discussion will proceed as follows. I begin, in the next section, by motivating the intuition that the deontic status of a threat often depends on the deontic status of what's being threatened. In particular, I argue that there is a moral presumption against threatening to do what one would not be permitted to do. I then make use of this, in Section III, to criticize versions of the first two approaches sketched above. These approaches are unsatisfying insofar as they are ill suited to account for how the moral assessment of what one is threatening to do can be relevant to the moral assessment of threatening to do it. In Section IV, I consider two initial attempts to explain more directly the impermissibility of threatening to do something by

appeal to the impermissibility of actually doing it. But these are rejected for failing to capture adequately the sense in which the coercee's choice is constrained or manipulated by the threat. I go on, in Sections V and VI, to develop my own account of the relevant principle of deontic dependence in terms of the illicit transfer of responsibility effected by coercive proposals. In Section VII, I show how this account can be generalized to explain what is objectionable about certain kinds of threats to do what one would otherwise be permitted to do. Finally, in Section VIII, I will explain how the objection to the sort of illicit responsibility transfer involved in coercive threats is connected to their *coercive* character in particular—and thus to the sense in which such threats constitute a mode of influence especially disrespectful of persons as autonomous agents with ends and reasons of their own.

## II. DEONTIC DEPENDENCE

What is the connection between the morality of threats and the morality of what is being threatened? Let's begin with the fact that it is not always objectionable to threaten a person in order to get him or her to do something. The basketball coach who threatens her players with having to run laps if they're late for practice does nothing wrong. There is no moral objection to a potential car-buyer threatening to go elsewhere if the dealer doesn't throw in the extended warranty. A boss may threaten to fire her employee if he doesn't get his TPS reports in on time. And the employee may threaten to take another job if he doesn't get the promotion he feels he deserves.

Contrast these cases with cases of clearly *impermissible* coercion: "Your money or your life"; "I'll beat you up if I catch you in Malibu again"; "If you don't get off of my barstool, I'm going to pour beer all over your head."

If we think about these examples (and many others), it seems that in cases where we intuitively take the threats to be impermissible, normally the threatened conduct is itself impermissible. And conversely, where the threatened conduct is permissible, it often seems that making the threat is as well. On the strength of such cases, several philosophers have accepted a strict correspondence between the deontic

status of a threat and that of the conduct being threatened.<sup>6</sup> In other words, they accept the following principle.

*Strong Deontic Dependence:* If A threatens to do  $x$  unless B does  $y$ , and does this in order to induce B to do  $y$ , then, other things equal, A thereby wrongs B if and only if it would be independently wrong for A to do  $x$  if B does not do  $y$ .<sup>7</sup>

There are, however, serious challenges facing an approach that seeks to account for the wrong of coercive threats entirely in terms of the impermissibility of the act or omission being threatened. First, the general rule linking the moral valence of the threat to that of the threatened act appears to have some important exceptions. Consider the case of blackmail. Much philosophical and legal writing on blackmail has focused on the supposed “paradox” involved in thinking that it should be wrong (or illegal) to *threaten* to reveal information about a person, which it would be perfectly permissible (and legal) to reveal.<sup>8</sup> Or, to take a different sort of example, I might try to get you to go out with me by threatening to destroy the only copy of my book manuscript if you don’t (I know you won’t want me to ruin my career in this

6. See, for instance, Haksar, “Coercive Proposals”; Pallikkathayil, “The Possibility of Choice”; Scanlon, *Moral Dimensions*; and Shaw, “The Morality of Blackmail.”

7. Three points of clarification: first, I am interested in accounting for the characteristic, prima facie moral objection to coercion. This objection may be defeasible—hence the *ceteris paribus* clause. Second, the term *independently* is meant to rule out the possibility that the threatened conduct would be (im)permissible solely in virtue of the fact that it would constitute the carrying out of an (im)permissible threat (cf. Warren Quinn, “The Right to Threaten and the Right to Punish,” reprinted in Quinn, *Morality and Action* [Cambridge: Cambridge University Press, 1993]). Third, I will not here enter into debates about how exactly to distinguish between threats and offers. None of my arguments will rely on this distinction. To avoid concerns about this, we could rewrite the principle in a way that is neutral on this issue as follows: “If A communicates to B that A will do  $x$  if and only if B does  $y$ , and does this to induce B either to do or refrain from doing  $y$ , then, other things equal, A thereby wrongs B if and only if it would be independently wrong for A to do  $x$  if B does not do  $y$ , or for A to refrain from  $x$  if B does  $y$ .” But for ease of exposition, I will continue to focus on threats in particular, and rely on an intuitive sense of what constitutes a threat.

8. See Glanville Williams, “Blackmail,” *Criminal Law Review* 79 (1954); James Lindgren, “Unraveling the Paradox of Blackmail,” *Columbia Law Review* 84 (1984); Mitchel Berman, “Blackmail,” in *The Oxford Handbook of Philosophy of Criminal Law*, ed. John Deigh and David Dolinko (New York: Oxford University Press, 2011); Shaw, “The Morality of Blackmail.”

way).<sup>9</sup> Although it is not impermissible for me to destroy my manuscript if you don't go out with me—only very imprudent—it would nevertheless be wrong for me to make the threat.

Second, even setting aside such counterexamples, there is the challenge of explaining why it should matter to the morality of threatening a person that one would or would not be permitted to do what one is threatening to do. It can't, for instance, be that the moral objection to, say, threatening to assault someone is the same as the moral objection to actually assaulting him. Whatever we want to say about why exactly it would be wrong for Green to assault Brown, this will not itself explain why it would be wrong for Green to threaten to do so if he doesn't steer clear of Malibu. The natures of these acts and their effects are just too different. We need some account, then, of how the impermissibility of the one could help to explain the impermissibility of the other.

This second challenge is one I think we need to confront head-on if we are to have an adequate understanding of the morality of coercion. This is because there simply are threats whose moral status is at least partly explained by the moral status of the threatened conduct. We can see this if we consider cases where we hold fixed other relevant features of the proposal situation, such as the undesirability or harmfulness of the threatened consequence, on the one hand, and what the person is being directed to do, on the other.

For example, take the case of the supervisor who threatens to fire her employee unless he gets his reports in by the end of the day. This seems morally innocent. By contrast, it would not be permissible for her to threaten to slash the tires on his car. This is so even if having his tires slashed is on the whole preferable and less harmful to the employee than losing his job. We may suppose he would find it much easier to resist the threat to slash his tires. And if, in addition, he is being asked to do the same thing in both cases—get his TPS reports in—then what seems to be doing the moral work here is the difference in the independent moral status of threatened conduct.

To take another example, suppose I tell you that I won't investigate the theft of your jewels unless you pay me \$1,000. This is permissible if I am a private investigator but not if I am a detective with the local

9. The example was suggested to me by an Associate Editor for *Philosophy & Public Affairs*.



police. The difference seems explained by the fact that if I am a police officer, then I have a duty to investigate, which I am proposing not to do unless you pay me. It's permissible for a private investigator to refuse to investigate a crime if she doesn't receive payment; it is not permissible for an officer of the state to do this.

As noted, however, there seem to be cases where it is wrong to threaten to do something, though one would have been permitted to do it had one not issued the threat. And hence there seem to be counterexamples to Strong Deontic Dependence. But there also appear to be cases in which the wrong of issuing a threat turns on the wrongness of what one is threatening to do. So, although we should be skeptical of Strong Deontic Dependence, we should nevertheless consider

*Weak Deontic Dependence:* If A threatens to do  $x$  unless B does  $y$ , and does this in order to induce B to do  $y$ , then, other things equal, A thereby wrongs B *if* it would be independently wrong for A to do  $x$  if B does not do  $y$ .

Weak Deontic Dependence holds that the wrongfulness of the threatened conduct is a sufficient, though not a necessary, condition of there being a moral presumption against using such a threat to influence a person's behavior.<sup>10</sup> This principle seems to be tracking a

10. Are there perhaps counterexamples to Weak Deontic Dependence? Consider the following example (suggested to me by an Associate Editor of *Philosophy & Public Affairs*): Suppose you know I'm tempted to cheat on my spouse and in order to get me not to, you threaten to cheat on your own spouse, whom I care a great deal about and do not want to see hurt. Let's suppose further that you're bluffing and have no intention of cheating on your spouse either way (perhaps you have even cleared the plan with her ahead of time). Is it really wrong for you to make this threat in order to discourage my infidelity? Assuming you present it as, and it is taken by me to be, a genuine threat—as opposed to a strategy to get me to think about my own actions in a new light, by focusing my attention on the kind of hurt and betrayal involved in infidelity—then what makes the difference is surely that you are trying to get me to do something I'm independently obligated not to do. Remember, though, that Weak Deontic Dependence identifies a condition establishing a presumption against, or prima facie objection to, certain threats. In this kind of case, the right thing to say, I think, is that *if* your threat is permissible, it is not because in this case there is no presumption against threatening to do what you are not permitted to do, but rather that your reasons for stopping me from doing what I have an obligation not to do are sufficient to overcome this presumption. This seems confirmed by the thought that such a threat should be viewed as a last resort—a strategy to be used only after you've tried other, less problematic ways of dissuading me from the affair.

genuine phenomenon, even if it's not the whole story about the morality of coercion. And this causes trouble for the first two approaches introduced in the previous section. For, as I argue below, neither of these approaches are well suited to account for the evident connection between the (im)permissibility of certain threats and the (im)permissibility of what is being threatened. But this is a phenomenon we need some account of. It's not self-explanatory.

### III. AUTONOMY AND INDEPENDENT REASONS

The kinds of examples that support Weak Deontic Dependence present problems for approaches to the morality of coercion that seek to locate the basic wrong of coercion either in the way it undermines the victim's autonomy or in the way it redirects the victim's activity away from what she has independent reason to do.

Consider the autonomy-based approach. It is not clear why the ability of a person to make a free or voluntary choice in the face of a threat should depend on the *moral* evaluation of what's being threatened. On the traditional approach to understanding coercion, what matters is not so much the moral content of the threat but the kind of influence or pressure coercive threats bring to bear on their targets. The latter, however, would seem to depend on psychological features of the coercee—what she cares about, what she is afraid of, what she is equipped to handle. And there seems to be no necessary connection between such things and the impermissibility of someone's bringing about the consequence that is being threatened.

To see this more clearly, we need to consider the form that this sort of autonomy-based view must take. It does seem natural to think that when a person is coerced into performing an action, there is a sense in which he has not really made the choice for himself to act in that way, or at any rate that his choice is not a free one. But if this is to provide a substantive account of the objectionable nature of coercion, we obviously cannot be relying on a conception that defines free or autonomous choice partly in terms of the condition that such choices are not coerced. Otherwise, to say that coercion is wrong because it undermines the victim's autonomy would be to say little more than that coercion is wrong because it's *coercive*.

An account that appeals to autonomy and related notions in order to explain why coercion is wrong will therefore need to begin with an independent conception of autonomy that is of some recognizable value or that otherwise commands respect. The problem is that conceptions of autonomy that do not have the absence of coercion built into them—and that instead center, for instance, on the structure of the agent's motivations—do not seem well placed to explain how the *deontic* status of what a person is threatening to do could itself determine whether the choice to comply with the threat was freely made.

Consider, for example, Harry Frankfurt's account of what makes a threat genuinely coercive.<sup>11</sup> Frankfurt thinks that one is coerced into doing something if two conditions are met. First, he claims that someone who is coerced "is *compelled* to do what he does. He has *no choice* but to do it."<sup>12</sup> This condition is met, Frankfurt claims, when the person who is coerced is unable to resist the desire to avoid the consequence being threatened and so, in that sense, has no alternative to doing what the coercer wants.<sup>13</sup> The second condition is that the coeree must prefer not to act on the desire that actually moves her.<sup>14</sup> He adds this second condition because he thinks that it's possible to be presented with an offer so attractive that one might not be able to refuse it. But he is hesitant to say that making such an attractive offer amounts to coercion.<sup>15</sup> Thus, on Frankfurt's account, if A succeeds in coercing B into doing *x* by threatening some consequence C if B does not do *x*, then B must (a) be moved by a desire to avoid C that she is unable to resist and (b) prefer not to be moved by that desire.

There seems to be no reason to think that there should be any correlation at all between the threats that are likely to undermine autonomy, as Frankfurt conceives of it, and the deontic status of what is being threatened. The moral quality of the consequence does not appear to

11. Frankfurt, "Coercion and Moral Responsibility."

12. *Ibid.*, p. 36.

13. *Ibid.*, p. 39: "If the victim's desire or motive to avoid the penalty with which he is threatened is—or is taken by him to be—so powerful that he cannot prevent it from leading him to submit to the threat, then he really has no alternative other than to submit."

14. *Ibid.*, p. 43: "It seems that a threat is only coercive, then, when the motive from which it causes its victim to act is a motive from which he would prefer not to act."

15. Or that one would therefore not be morally responsible for one's action. For a related discussion, see Frankfurt, "Alternate Possibilities and Moral Responsibility," in *The Importance of What We Care About*.

be relevant. Return to the contrasting pairs of cases considered in the previous section. In these cases, the moral valence of the proposals in question (the boss's threat to fire her employee vs. slash his tires, the cop's threat not to investigate vs. the private eye's) depends on the moral valence of what is being proposed. But it's hard to see how the impact or pressure on the victim's will differs among the cases. The impermissibility of the threatened conduct does not as such affect what the coerced is psychologically capable of.

The point holds for both of Frankfurt's two conditions. With respect to the first, one can, on the one hand, obviously have very strong, perhaps even irresistible, desires to avoid certain consequences (for example, being fired from one's job) that do not involve any moral wrongdoing whatsoever. On the other hand, there is no reason to think that, in general, agents are especially incapable of resisting desires to avoid harms or other consequences when those are *wrongfully* brought about. There doesn't seem, in other words, to be some *special* difficulty about making autonomous choices in the face of a person's threat to behave impermissibly.

With respect to the second condition, it's hard to see why one would, in general, prefer not to be moved by a desire to avoid the consequences of others' wrongdoing. If you threaten my life in order to get me to fork over my money, I'll fork it over because I want to live. This is a desire I fully, wholeheartedly identify with. There is a sense in which it's true that I don't really want to give you my money. This is because I don't want to be in this situation; I wish you hadn't threatened me in the first place. But this is different from saying that, in responding to a threat, I do not endorse the motive on which I act *in* my situation as it actually is—that I would resist acting on this desire, if only I could.<sup>16</sup>

The basic point I'm making is not limited to Frankfurtian accounts of autonomy. The same could be said of accounts that characterize autonomous choice in terms of, say, rational control, or reason responsiveness. The problem arises so long as we focus on the impact a threat has on the psychology of its target. The fact that there are cases in which what appears to make the difference between a permissible and

16. For a related discussion of Frankfurt's account, as well as further objections to the autonomy-based approach, see Pallikkathayil, "The Possibility of Choice."

an impermissible threat is the deontic status of the threatened conduct remains unexplained on the autonomy-based approach.

In fact, the difficulty autonomy-based accounts have with these cases seems to be an instance of a more general problem with this approach. The approach aims to specify the wrong of coercion by an appeal to some sense in which the coerced's choice is defective—nonautonomous or inauthentic or non-reason-responsive, or whatever. However, though coercive threats may often involve distorting influences such as fear, they primarily work by altering the reasons the recipient has for pursuing certain options over others. But responding to changes in one's situation that make certain alternatives more attractive or reasonable is just what acting rationally and autonomously normally involves. The fact that the situational change that affects one's reasons is due to a *threat* is not in this respect relevant. To be sure, from the point of view of the victim, the overall situation is hardly desirable. But people frequently make free, rational decisions in less than ideal circumstances. It is very difficult to supply an understanding of what it is for a choice to be *unfree* that (a) is sufficiently independent of viewing the choice as *coerced*, and (b) makes it plausible that choices made in response to what are intuitively coercive threats are unfree in that sense, given that such threats typically work by introducing reasons to which the recipient is expected to respond rationally.<sup>17</sup>

This problem for autonomy-based accounts leads naturally to the second approach outlined in Section I. Perhaps the relevant objection is precisely that the coerced is being given reason to do something she would not otherwise have, or take herself to have, good reason to do. For instance, in a recent article, A. J. Julius has defended an account of coercion as violating what he calls the Independence Principle.

*Independence Principle:* I should not (do *y*, intend by *y*-ing to bring it about that you do *x*, and fail to believe with warrant that, for some

17. This is not to say that the notion of autonomy is irrelevant to the morality of coercion. It's very plausible that it is part of our normal understanding of autonomy that it constitutively requires the absence of wrongful coercion. On this view, rather than being based on a prior understanding of autonomy, an account of wrongful coercion will be part of the theory of autonomy, helping us to elaborate its nature and moral significance.

reasons R independent of me, my *y*-ing facilitates your [doing *x* because you take R as giving you sufficient reason to *x*].<sup>18</sup>

If I threaten you, I generally intend to get you to do something without believing that you will thereby come to act on the basis of some reason you have that is independent of my threat. This may either be because there is no such independent reason (you don't have reason to give me your money) or because, even if I believe there is an independent reason for you to do what I'm trying to get you to do (as in paternalistic coercion), I will be leading you to do it *in order to avoid the threatened consequence*, and not for the sake of that independent reason.

It's clear that this approach will have trouble accounting for examples of permissible threats.<sup>19</sup> It is often (if not always) permissible to threaten to do what one otherwise has a right to do. But a threat, if it is not completely superfluous, would seem to influence the recipient in ways other than by bringing her to act on reasons that apply to her independently of that threat. If a coach threatens to make her players run extra laps if they are late to practice, this is presumably because she is not confident they are sufficiently motivated by the independent reasons favoring punctuality. But it hardly seems that, in introducing this new incentive to arrive by the start of practice, the coach thereby wrongs the members of her team. More generally, because the intentional introduction of any new reason or incentive to act is rendered suspect by the Independence Principle, the account fails to track the phenomenon underlying Weak Deontic Dependence. There is no reason to think that the moral status of carrying out a threat will be the thing that determines whether or not it introduces new reasons for action beyond those that already obtain. The idea that there is something objectionable in the fact that threats alter their targets' reasons

18. Julius, "The Possibility of Exchange," p. 363.

19. There is also a question about how consistently to account for permissible *offers* with this principle. Much of Julius's article is concerned to show that most intuitively permissible offers will not in fact violate the Independence Principle. In brief, a (legitimate) offer can be understood as a proposal that the parties perform a set of acts, which they collectively have reason to perform independently of the proposal itself. Permissible threats are not amenable to such an analysis, however, since issuing a threat seems to depend for its point on the belief that the recipient is not sufficiently motivated to act on reasons that apply independently of the threat itself.

for action doesn't give us the resources we need to make the distinctions we want to make.

#### IV. READINESS TO DO WRONG AND THE NARROWING OF OPTIONS

I turn now to consider some ways of accounting for the connection between the wrongfulness of an act and the (prima facie) wrongfulness of threatening to perform that act in order to induce someone to behave in a certain way. We can begin with a proposal made a number of years ago by Vinit Haksar. Haksar writes:

Since violation of moral duty is wrong, so is the readiness to violate our moral duty wrong; those who use coercive threat are normally ready to violate a moral duty, should that be necessary [sic].<sup>20</sup>

This simple explanation is too simple, however. One problem is that it does not account for bluffs. Haksar himself does not think this is a serious objection, since he thinks we can account for the wrong of bluffs by appeal to the fact that they are deceptive. But this does not seem to me an adequate response. Imagine two scenarios in which Gray tells Black that she will break his finger unless he tells her what she wants to know (assume she has no right to the information). In the first, Gray means it; she intends to break Black's finger if he doesn't talk. In the second case, she is bluffing—though we can assume Black does not know this and believes what she says. In both cases, Black gives up the information in response to the threat. Haksar's view seems to imply that the nature of the wrong committed by Gray is fundamentally different in the two scenarios. But this is implausible. Black's complaint against Gray seems the same in both cases. In particular, it does not seem that in the second scenario, Black's primary complaint is that he was *lied* to.

There is a more fundamental problem with Haksar's explanation. It may be that the readiness to commit a wrong is itself wrong.<sup>21</sup> And it may be that this is one aspect of what is wrong with typical instances of threatening a person. But this can't be the whole story. Gray might

20. Haksar, "Coercive Proposals," p. 71.

21. Compare Thomas Nagel's discussion of the wrong of intending evil in *The View from Nowhere* (Oxford: Oxford University Press, 1986), pp. 181–82.

be ready to commit a wrong against Black, without ever communicating this to Black. Or, Gray might make an unconditional threat—telling Black simply, “I’m going to break your finger, the first chance I get.” And yet, there seems to be something different about the way in which Gray wrongs Black when she makes a *coercive* threat as a way of influencing Black’s behavior. This comes out clearly if we imagine that what Gray is threatening to do is violate the rights of some third party, Blue. Gray’s readiness to wrong Blue may already be morally problematic, but it’s hard to see how it could be a wrong against anyone other than Blue. And yet, when Gray coerces Black into doing something by threatening to harm Blue, it seems that Gray wrongs Black in so doing.

More generally, it seems clear that what explains the wrong of a coercive threat in particular must have something to do with the manipulative character of such a threat—the fact that it aims to get someone to do something she wouldn’t otherwise do—which depends on its conditional form. One way to put the point is that Haksar’s proposal may help account for our third intuition concerning deontic dependence, but it provides no help with the first two, which concern the way coercion constrains a person’s choices. It does not help us to see why a conditional threat, designed to influence the recipient’s behavior, should be in any distinctive way objectionable.

Let’s turn, then, to a more sophisticated theory. Japa Pallikkathayil has recently argued that when a coercer is threatening to do something impermissible, this fact plays an important role in explaining how the threat undermines any possibility of the victim’s consenting to the resulting transaction.<sup>22</sup> And this—acting in a way that undermines the very possibility of consent—is in turn said to violate the Kantian injunction never to treat a person merely as a means to one’s ends.<sup>23</sup>

Pallikkathayil’s claim is that the recipient of a coercive proposal has her options constrained in such a way that she is unable genuinely to

22. Pallikkathayil, “The Possibility of Choice.” In fact, it’s not just consent that’s undermined. Pallikkathayil holds that coercion interferes with the possibility of exercising one’s normative powers generally.

23. See Immanuel Kant, *Groundwork of the Metaphysics of Morals*, trans. Mary Gregor (Cambridge: Cambridge University Press, 1996). For defenses of the “possible consent” interpretation of Kant’s Formula of Humanity, see Onora O’Neill, “Between Consenting Adults,” in *Constructions of Reason* (Cambridge: Cambridge University Press, 1989); and Christine Korsgaard, “Kant’s Formula of Humanity,” in *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996).



consent to the interaction. For a putative act of consent to be valid, one must have had other options available. But obviously, it's not enough for one merely to have had *some* possibility of withholding consent. The fact that one could have refused to hand over the money to the mugger, and taken a bullet instead, is not the sort of alternative that is relevant to the validity of one's consent. On some views, the problem here is that one's options are not sufficiently attractive. For Pallikkathayil, though, the problem is that one is being denied an alternative one is morally entitled to have:

[The mugger] impermissibly takes control of an option that the victim is entitled to have—to keep both her money and her life. Since the victim does not have access to the options she is entitled to with respect to her money, the act of handing the money over to the mugger cannot be an exercise of her discretionary authority with respect to her money.<sup>24</sup>

Here, we can see why it matters for Pallikkathayil that what is being threatened is itself impermissible. One is normally entitled, for instance, to have available the option of *keeping one's money without being murdered*. By contrast, one is not necessarily entitled to have the option of *showing up to work late without being fired*.<sup>25</sup> The idea is that one does not have access to the options one is entitled to if one's options are limited by another's (conditional) intention to behave impermissibly.

There are thus two parts to the view. First, for a person to be in a position to give valid consent (or exercise some other normative power), she must have access to at least those alternatives to which she is morally entitled. And second, by threatening to do something impermissible, the coercer deprives the coercee of some option(s) she is entitled to. I have doubts about both parts.

First, it does not seem that merely lacking an option that one is entitled to have is sufficient to vitiate consent. Suppose your appraiser has promised you that she'll come inspect your property on a certain day and that you've arranged your schedule accordingly. She subsequently

24. Pallikkathayil, "The Possibility of Choice," p. 13.

25. *Ibid.*, p. 12.

informs you that the date has become inconvenient for her and she's canceling—she will have to do the appraisal a day earlier. Although you may be entitled to have her do the inspection on the agreed-upon date, the fact that this is no longer an option does not mean that you are not in a position to consent to the appraiser's entering your property a day early. Or again, although I may owe it to my students to hold regular office hours, and wrong them if I don't (in the absence of some good excuse), this doesn't by itself imply that, when they set up special appointments to meet with me, they are not bound by these arrangements.<sup>26</sup> One may wrong a person in failing to make available an option to which he is entitled, but this alone doesn't undermine the possibility of his genuinely consenting to related transactions. To explain why coercion vitiates consent, it's not enough to point out that it eliminates an option to which the coercee is entitled.<sup>27</sup>

Second, it's not clear that threatening something impermissible always constitutes the elimination of an option the coercee is entitled to have available to her. We need to be clear about the sense in which threats narrow their recipients' options. When Gray threatens Black—telling him she'll break his finger if he doesn't tell her what she wants to know—it's not that refusing to give up the information is no longer an option for Black. Rather, it's that Black no longer has the option of keeping the information to himself without having his finger broken. One way to think of this is that the threat forecloses a conjunctive

26. Pallikkathayil does consider a possible restriction of her thesis to options that are “deliberatively significant” (“Possibility of Choice,” p. 14). Would this restriction allow us to sidestep these counterexamples? I don't think so. The only sense in which, for instance, the option of having the appraiser come on the day she originally agreed to come is not deliberately significant is that it's no longer available. But this is not—and cannot be—what Pallikkathayil means by the term. Rather, an option lacks deliberative significance for a person if its removal “does not affect the reasons she takes herself to have” (p. 13). But in this sense, the option of having the appraiser come on the day she said she'd come is deliberately significant. We can easily imagine that no longer having that option available does affect the reasons you take yourself to have for allowing her to do the appraisal a day earlier. Still, lacking this option is not enough to vitiate your consent.

27. I suspect (though I will not argue here) that the fact that coercion tends to vitiate consent is explained more directly by the fact that it is an impermissible means of getting someone to agree to something. For this view, see Wertheimer, *Coercion*. Of course, if that's right, then we cannot appeal to the fact that coercion vitiates consent to explain what is wrong with it as a means of getting someone to do something.

option: Black no longer has the choice to keep quiet *and* keep his finger intact.<sup>28</sup> The problem is that the target of an impermissible threat will not necessarily be entitled to the conjunctive options eliminated by that threat. This is because one is entitled to a conjunctive option only if one is entitled to both conjuncts. Sometimes this will be the case, as when the mugger removes the conjunctive option (YOUR MONEY and YOUR LIFE). But often it won't be. Consider the option (SITTING AT THE BAR and NOT HAVING BEER POURED ON YOUR HEAD). I might eliminate this option by threatening to pour beer on your head if you sit down. Or I might eliminate it by taking the last available barstool myself. You have a claim against me not to limit your options in the first way, but not the second. Yet in both cases I foreclose (by different means) the very same conjunctive option. (And in the case of the threat, this is the *only* option I foreclose—you are still free to sit at the bar, and you are free to keep your head dry, just not both.) Thus, we can't explain what's wrong with making the threat by appealing to the fact that you are entitled to that option (since it is not a fact).

In response to this difficulty, we might think about the option removed by the threat differently. Even if you're not entitled to a seat at the bar, you are entitled, *if* you sit at the bar, to not have beer poured on your head.<sup>29</sup> On this interpretation, the entitlement contravened by the threat is an entitlement to a conditional option—like the option to purchase a yacht if one can afford the asking price.

The fact that, by threatening to do something impermissible, the coercer forecloses a conditional option or possibility to which the coercee is entitled does not, however, seem to go to the heart of what is objectionable about coercion. To see this, notice that it does not distinguish between coercive threats and unconditional proposals to behave impermissibly. If I credibly threaten to pour beer over your head—not as a means of influencing your behavior, but just because I'm annoyed with you—then I deny you indefinitely many conditional options to which you're entitled. You no longer have the options: to not have beer poured over your head if you sit at the bar, to not have beer poured over your head if you stand, and so on. Although we might allow that

28. See Saba Bazargan, "Moral Coercion," *Philosophers' Imprint* 14 (2014): 1–18.

29. I'm grateful to an Associate Editor for suggesting this interpretation of the relevant entitlement.

the elimination of such possibilities is wrong, given that you are entitled to them, evidently it is not essentially coercive. As with Haksar's proposal, something is left out. Coercion is not just option-theft. When one is subjected to a coercive threat, there is something distinctly problematic about being forced to do something in order to avoid the threatened consequence. This is presumably why Pallikkathayil wants to say not only that a coercive proposal wrongly deprives the coercee of certain options, but also that, because of this, the coercee is not in a position to consent to her part in the coercer's plans. But since, as I've argued, this latter claim is also questionable, we need to look beyond the fact that coercive proposals narrow their recipients' options.

#### V. ACCOUNTING FOR WEAK DEONTIC DEPENDENCE: DELIBERATIVE SECURITY

If coercing a person by threatening to do something impermissible objectionably constrains that person's choice, it must do so in a sense that goes beyond the fact that the threat eliminates certain options. What, then, is the relevant sense of constrained choice? It's not just that the coercee confronts a restricted set of options. Instead, I suggest, it's that, in deliberating about the alternatives that are open to her, she has to take into account the fact that opting for certain of them will likely lead another to treat her in ways that are morally prohibited. It's natural to think, however, that in deliberating about what to do, one should not have to consider the possibility that others will behave impermissibly as a result of the decision one makes.

To explain this, let me first introduce a definition.

*Deliberative Security:* A person has deliberative security with respect to a particular option P if she can, given the circumstances, rationally proceed on the assumption that whether others act permissibly does *not* depend on whether she chooses in favor of P.

For example, if, in thinking about whether he should take a trip to Malibu, Brown has to take into consideration the likely fact that Green will beat him up if he does take the trip, then Brown lacks deliberative security with respect to the option of going to Malibu.

We can use this notion to make sense of Weak Deontic Dependence. Recall that Weak Deontic Dependence holds that a threat of the form

“unless you do *y*, I’ll do *x*” is prima facie wrongful if the issuer of the threat is not morally permitted to do *x*. What I want to suggest now is that such a threat is prima facie wrongful in virtue of the way it undermines the recipient’s deliberative security with respect to some option otherwise open to her (i.e., refraining from *y*).<sup>30</sup> Call this the *Deliberative Security Thesis*.

The case for accepting the Deliberative Security Thesis is clearest, I think, when what’s being threatened is a violation of the coercee’s rights. Think, for example, of rights to bodily integrity or rights to personal property. On many accounts of such rights, their significance lies at least partly in supporting the right bearer’s autonomy—she is entitled to be the one, by and large, who decides what happens with her body and possessions. Arguably, then, part of the point of having such rights is that they provide some security that one can rely on in making plans for the future. If this is right, then the entitlement to expect that one’s rights will be respected cannot itself be contingent on any (otherwise permissible) decisions that one makes. Moreover, even if, say, one’s right to bodily integrity is not itself violated, if one is deprived of the ability to (rationally) rely, in one’s deliberation and planning, on the expectation that this right will be respected, one loses something of significant value. This supports the claim that one wrongs a person if, by issuing a threat, one undermines the deliberative security provided by that person’s rights.

Consideration of these autonomy rights might suggest that the best way of explaining the Deliberative Security Thesis is by postulating an independent *right* to deliberative security.<sup>31</sup> We would thus be able to

30. Compare Shaw: “If one is permitted to deliberate about performing an action A at all, one is entitled to deliberate as if one could perform A free of any sanction on its performance that would constitute a wrong” (“The Morality of Blackmail,” p. 187). The claim I am making is similar. However, Shaw’s formulation raises a number of questions. Does his principle imply that one is entitled to deliberate as if one could perform A free of any impermissible sanction *even in the face of good evidence* that such a sanction will be attached to A? If so, it cannot plausibly be construed as a rational entitlement. Is it then a *moral* entitlement to deliberate in this way in the face of credible evidence to the contrary? This, too, seems doubtful. It’s not clear what value there would be in a moral entitlement to deliberate irrationally. The principle seems best read as a moral entitlement to the conditions under which it would be rational to deliberate as if one could perform A free of any impermissible sanction—an entitlement, roughly, to what I have called deliberative security.

31. Roughly along the lines endorsed by Shaw. See previous note.

explain why coercive threats to do what is impermissible are themselves wrong: they violate their target's right to deliberative security.

There is a problem with this suggestion, however. A credible threat would certainly undermine such a deliberative right. But so, it seems, would a warning by a third party. Suppose you know that Ralph, who has anger issues, is extremely sensitive about the size of his nose and that if I so much as mention it, he's likely to punch me in mine. In telling me this, you undermine my deliberative security with respect to the option of, for instance, asking Ralph whether he has to have his sunglasses specially made. But you do not coerce me into keeping my mouth shut, nor do you wrong me in the way that Ralph would wrong me if he were to threaten to break my nose should I call attention to his. You do not wrong me at all. So, if we are to understand coercion as objectionable in virtue of its constituting the violation of some deliberative right, like the right to deliberative security, we will need to explain why a threat to do something wrong infringes this right, while a warning that someone is likely to do something wrong does not. I do not know how to do this. I'm therefore skeptical of the attempt to account for the objection to coercion by appeal to a prior theory of deliberative rights.

This does not mean, however, that we should abandon the Deliberative Security Thesis. It may not rest on a prior *right* to deliberative security, but we may be able to supply an alternative basis for the thesis. This is the task I turn to now.

#### VI. A WRONGFUL TRANSFER OF RESPONSIBILITY

I've argued for the need for some explanation of why it should be wrong, in particular, to threaten to do what it would be impermissible for one to do. We've seen that it will not work to claim that such threats eliminate certain options—options the coercee is entitled to have available. Rather, what seems peculiarly objectionable is the way in which they impact the victim's deliberation concerning the options that are open to her. This has led us to the Deliberative Security Thesis and the question of its basis.

I now want to suggest that we understand undermining a person's deliberative security as, in a sense, involving an illicit extension or transfer of responsibility. Suppose, for instance, that a mugger

threatens to shoot you if you don't give her your wallet. It is now, in a sense, up to you to decide whether to do one thing and be shot, or do something else and walk away unharmed. Your life depends on handing over your wallet, and the choice is yours. From your point of view, then, it is as though you are the one responsible for ensuring that you are not intentionally shot by this person. Whether she acts, in this respect, as she is morally required to act has thus become your responsibility. She has, in effect, disavowed it. But it is obviously a responsibility that properly falls to the mugger herself, not to you. In disavowing her responsibility in this way, as a means of influencing your deliberation and choice, the mugger thus wrongs you, even if she never pulls the trigger.

To elaborate: there are two principal claims being made here. First, I claim that the mugger, in making her threat, denies or disavows responsibility for a certain aspect of her conduct. What does this mean? Discussions of moral responsibility often focus on a backward-looking sense of responsibility—the sense in which one may be responsible for something that has happened or that one has done. I am not claiming that coercers necessarily disavow their responsibility for what they do in this sense. The mugger might, for all I say here, fully acknowledge that if she shoots you, she will be responsible for having done that—hence, legitimately subject to blame and punishment on account of her actions. Nor am I claiming that the mugger must be failing to recognize that she has the moral duty not to shoot. Rather, my claim is that her threat communicates a refusal to take responsibility for what she does in the sense that she purports to be unwilling to ensure (or even to attempt in good faith to ensure) that her conduct conforms to her duty.

Normally, functioning, mature adults are morally responsible in a forward-looking sense for making sure they act only in ways they are morally permitted to act. To say that one is morally responsible for what one does in this sense is to say more than that one is expected to behave only in ways that are morally permissible—since one might succeed in this, as it were, by accident. In addition, one is expected, for instance, to be sensitive to various moral risks that may accompany certain opportunities (think of the need for a judge or politician to be on the lookout for possible sources of bias or conflicts of interest), to

attend to changes in one's circumstances that may be morally relevant—for example, by calling on one to abandon, for the time being, a moral rule or policy that is suited to more normal conditions—and so on.<sup>32</sup> More generally, one is to recognize that it is part of one's task, in conducting one's life, to take proper account of the moral standards that apply to one and to do what one can to ensure that one's behavior does not run counter to these. It is this sort of responsibility the mugger abandons when she proposes that what will determine whether or not she unleashes violence on you depends on your decision about whether to hand over your wallet.

My second principal claim is that, as a direct result of this refusal of responsibility on the part of the coercer, the victim's responsibility is thereby expanded. One might object at this point that it cannot really become *your* responsibility to ensure this mugger does not shoot you. Of course I admit that, in one sense, her responsibility not to harm you does not transfer to you. The mugger is still herself responsible for whether or not she shoots you; threatening you could not change this. But it does not follow that your responsibility with respect to your own rights has not been extended—albeit illegitimately—to cover what properly belongs in this other agent's sphere of responsibility. This may be easier to see if we modify the current example slightly. Imagine that instead of the mugger threatening your life in order to get you to hand over your wallet, she is threatening the life of an innocent third party. Your situation, then, is that you have been handed by the mugger a means of determining whether she shoots this other person. You cannot reasonably or responsibly come to a decision concerning what to do with your money without taking this into account. Your decision about whether to give this person your cash is at the same time a decision about whether this person will murder another human being, and you are responsible for it as such.

Note that, on this side of the interaction, your responsibility is implicated not just in the forward-looking sense, but also in the backward-looking sense. There is no reason to think you would be entirely blameless if the bystander were shot as a result of your refusal

32. For a helpful discussion of the notion of forward-looking responsibility along these lines, see Henry Richardson, "Institutionally Divided Moral Responsibility," *Social Philosophy and Policy* 16 (1999): 218–49.



to hand over your money. Of course, it's the mugger who's primarily to blame. Nevertheless, in refusing to give in, it may be perfectly reasonable to criticize you for acting with callous disregard for the other's life and well-being.<sup>33</sup>

All of this is true of the original case, as well, in which it is your own life the mugger is threatening. You would, for instance, be open to criticism for jeopardizing your welfare, and perhaps those who depend on you, were you to refuse the mugger's demand.<sup>34</sup> Again, however, the crucial thing is not that you might be forced to accept blame for something you would not otherwise have to. The relevance of considering this extension of backward-looking responsibility is that it reflects the fact that simply in having your life threatened by the mugger, you acquire a genuine and normatively relevant forward-looking responsibility for what the mugger goes on to do to you. Of course, you should not be in the position of having to bear this responsibility. But this doesn't mean that you don't in fact bear it.

If this account is right, then the victim of coercion takes on a responsibility he should not have to bear. And although there is no normative sense in which the coercer is thereby relieved of her responsibility to govern her actions in accordance with her moral obligations, it is no mere coincidence that the extension of the victim's responsibility is accomplished through the coercer's effective disavowal of her own responsibility. That is, via her threat, the coercer forces her victim to take responsibility for how, ultimately, he is treated by her (whether she harms him, for instance). And the coercer does this precisely by refusing herself to take responsibility in the way that she should for how she treats her victim (or, in some cases, a third party). The threat is effective only because the coercer successfully communicates to the coerced (a) her renunciation of her responsibility to make her decision about what to do in a way that depends on what she is morally permitted to do, and (b) that the coerced is in a unique position to take on this responsibility in place of her. We might say, therefore, in light of the direct connection between this renunciation of responsibility, on the one hand,

33. And even in cases where you would not be to blame, because refusing to comply with the threat is the right thing to do, you may be expected to answer for your decision as a decision to allow the coercer to carry out her threat.

34. Which is not to say that it would be appropriate, all things considered, for others to actually criticize you.

and the subsequent imperative for the coerced to take up the responsibility the coercer has renounced, on the other hand, that coercion effects a kind of “transfer” of responsibility. And that is the problem. It is not a morally valid transfer. It is in some ways analogous to a transfer of one person’s physical property into another’s possession that occurs by way of fraud or theft. Such a transfer has no legal or moral reality. But its empirical effects are real enough. The practical possession and control of the object now belong to someone who has no right to them. Coercion effects a similar sort of mismatch between the practical reality concerning who must take responsibility, say, for ensuring the victim’s rights are not violated, and the moral reality as to whom that responsibility properly belongs.

This sense in which the responsibility is transferred—refused, on the one hand, imposed, on the other—also helps to distinguish what happens when one issues a threat, from what happens when one issues a warning that someone else is likely to behave impermissibly. For in that case, the imperative that the recipient take responsibility for avoiding or preventing another’s wrongful conduct does not stem from a refusal, on the part of the one who warns, to accept responsibility for what he should.

This completes my account of what is distinctively objectionable about coercing someone by threatening to do what one has no right to do. We can explain Weak Deontic Dependence by appeal to the way in which such threats undermine their targets’ deliberative security with respect to options otherwise open to them. Other things equal, this is wrong when and because it illegitimately saddles the victim with the responsibility to ensure the coercer acts as he or she is morally required to.

#### VII. BEYOND DEONTIC DEPENDENCE

In this section, I turn to other forms of wrongful coercion—that is, forms that do not necessarily involve threatening to do what is independently impermissible.

So far, in explaining the moral presumption against threatening to do what one has no right to do, I have relied on the idea that, as moral agents, we are each responsible (in a forward-looking sense) for ensuring we act only in ways that are permissible. It is normally wrong to

offload this responsibility onto someone else as a way of influencing that person's behavior. We can think of the relevant responsibility as an executive moral responsibility. Whether the mugger, for example, is permitted to do what she is threatening to do is not really in question. The question is only whether she will act in accordance with her duty.

But there are other types of responsibility that may be implicated when one threatens another. There are thus other ways in which one might wrongly shift the burden of taking responsibility for one's conduct onto the person being threatened. And not all of these require that what one is threatening to do must itself be impermissible.<sup>35</sup>

First, it is sometimes possible to coerce another by threatening to act in *self*-destructive ways. Such threats can be just as objectionable, qua coercion, as threats to harm others. I might, for instance, threaten to set fire to the only copy of my book manuscript unless you agree to go on a date with me. This appears objectionably coercive. Yet, presumably, I have a moral right to destroy my manuscript if I so choose. I do not, therefore, undermine your deliberative security, as I have defined that term. Nevertheless, I have illegitimately transferred to you an executive responsibility that properly belongs to me. The only difference is that, in this case, I am renouncing my responsibility to ensure my actions conform to requirements of prudence, rather than morality. The result, however, is potentially just as problematic, for you are now in the position of having to take responsibility for ensuring that I do not act in an obviously self-destructive manner. And this is not a responsibility you should be required to bear as part of your decision about whether to go out with me.

Second, consider the alleged paradox of blackmail. There appear to be cases in which one is barred, morally, from threatening to reveal information it would be otherwise permissible for one to reveal.<sup>36</sup> For example, Red might be permitted to tell Gold's spouse about her adulterous affair. Even where this is so, however, it does not mean that Red may threaten to expose the affair in order to induce Gold to pay him to keep quiet. It would be wrong for Red to use the information in this way to coerce Gold. Why should this be so?

35. This is why Strong Deontic Dependence is too strong. See Section II.

36. For discussion, see the references listed in note 8.

Here, I suggest, we can identify a different component of the basic responsibility to ensure that one's own actions conform to moral standards. The Deliberative Security Thesis concerned cases involving the illegitimate "transfer" of the responsibility to ensure one's actions are in line with what one is permitted to do. But beyond this executive responsibility, there is a separate question of the agent's conscientiousness. Where the permissibility of acting in a certain way depends on the balance of competing moral considerations—as opposed to the innocuousness or moral neutrality of the act in question—one has a kind of "judgmental" responsibility to discern and respond to those morally relevant considerations. That one's conduct is, in fact, permissible given the circumstances does not by itself show that one has discharged this responsibility. One might have decided to act as one does on quite irrelevant grounds, and without taking care to ensure that, say, the resulting harm or interference with others is justified in this case, despite the normal presumption against these things.

It is, I suggest, this judgmental responsibility that the blackmailer disavows in making his threat. Red is in the position of deciding whether to intervene in another's personal life in a way that is likely to cause considerable pain to both Gold and her spouse, as well as seriously disrupt their marriage. Perhaps this is justified given the facts of the situation and Red's relationship to the parties involved. But it's not the kind of thing that should be taken lightly. And yet, Red's attempt at blackmail means that the question of whether Red tells Gold's spouse about her affair is no longer to be decided by way of Red's reflection and judgment as to whether the damage it will cause is justified by the reasons in favor of revealing the information. Rather, it is to be determined by *Gold's* decision as to whether avoiding this damage is worth the sum of money Red is asking for. Though the nature of the responsibility is different, what we have here is the sort of illegitimate disavowal and extension of responsibility we found in the other cases of wrongful coercive threats.

One virtue of this account is that it can explain what might seem like a puzzling divergence between otherwise similar cases. It seems wrong for Red to threaten to expose Gold's affair in order to get Gold to pay

him. But it does not seem in the same way objectionable for Red to use the same threat to induce Gold to confess her infidelity herself.<sup>37</sup> If the objection to blackmail were simply an objection to the use of sensitive information to manipulate the vulnerable party, these two cases should appear equally problematic.

The account I've offered can help explain this difference. Where Red's demand is for Gold to fess up on her own, there is no indication that Red has abdicated his responsibility to ensure that, insofar as he plays a role in causing Gold marital strife, there is sufficient moral justification for doing so. The suggestion, implicit in Red's threat, that Gold's refusal to confess will be taken by Red as reason for going to Gold's spouse himself is not obviously inconsistent with the conscientiousness that is called for by the situation. This is in clear contrast to the case in which Red bases his behavior on Gold's decision to pay him or not.<sup>38</sup>

37. Of course, this *might* be objectionable. Perhaps it is simply not Red's place to intervene in Gold's marriage in this way. But if this is so, it's doubtful that Red would be justified in going straight to Gold's spouse with the information. This is not the case I am imagining.

38. Decisions about morally significant actions may raise questions of conscientiousness even where the acts themselves are supererogatory. Just because one is not required to perform some morally good deed does not mean that there can be no objection to one's grounds for refusing to do it, having considered the possibility. And this may affect our assessment of certain proposals regarding the performance of such acts. Consider Feinberg's example of a Lecherous Millionaire, who, in order to get a woman to agree to go to bed with him, offers to pay for an operation her child desperately needs and which she is unable to afford. Whether such a proposal actually amounts to coercion is controversial. But insofar as we are inclined to view it that way, I suggest it is because we are inclined to think that, even if LM is not obligated to pay for the operation, it would be objectionable for him to consider doing so and then regard it as a matter to be decided according to whether he is able to have sex with the child's mother. This would be objectionable even if he does not make this known to the mother in the form of a *quid pro quo*. But in making the proposal, he does seem to make the woman responsible for his decision in a way that raises a further objection. For arguably she should not have to see herself as accountable, in choosing not to have sex with this man, for thereby failing to get her child the needed operation. At any rate, to the extent that we agree that this is so, we should, in line with the framework I've offered, view LM's proposal as objectionably coercive. See Joel Feinberg, *Harm to Self* (New York: Oxford University Press, 1986), p. 229.

## VIII. RESPONSIBILITY TRANSFER AS COERCION

The foregoing presents an account of coercive threats as involving an illicit transfer of responsibility from the coercer to the coerced. More needs to be said, however, about how the moral objection to such threats relates to their coerciveness. As I remarked above (in discussing Haksar's view), the distinctive objection to a coercive threat is an objection to a particular mode of influencing another person's behavior. Yet there are ways in which one might illegitimately disavow one's responsibility for something, and thereby impose it on someone else, which do not have as their aim coercing the other person into doing one's bidding. I could, for example, declare my willingness to commit some crime on your say-so—perhaps as a way of demonstrating my loyalty—leaving it up to you to decide whether I do it. Whatever is objectionable about my behavior here looks to be quite different from the wrong involved in one person's coercing another.

In response, I will try to explain how the wrong involved in the shift of responsibility from coercer to coerced is not merely a feature of the transaction in addition to or alongside the fact that the victim is (if he submits) coerced by the threat. Rather, it's that the illicit shifting of responsibility turns a proposal designed to provide its recipient with an incentive for action into something objectionably coercive. To make this case will require showing how the responsibility-shifting feature of coercive proposals helps to account for the additional intuitions concerning the nature of such proposals laid out in Section I.

I have introduced the responsibility transfer account in response to the intuition that the fact that the coercer has no right to do what she is threatening to do is relevant to the wrong of making the threat. But does the account shed any light on the intuitions (a) that a person does not act freely or autonomously when he submits to a coercive threat, and (b) that seeking to influence a person's actions through the use of coercion shows inadequate regard for that person's reasons for wanting to pursue some alternative course of action?

Let's take (b) first. Julius, recall, proposes a blanket prohibition on deliberately creating reasons for a person to do something she would not otherwise have sufficient reason to do. This, I argued, is too strong. On the account I've offered, however, a coercive threat does not merely create a reason that would not otherwise exist, but creates a reason we

have independent grounds for saying *should not* exist. This is most obvious where the (conditional) intention expressed by the threat is an intention to do something the coercer has no right to do. By foisting on you the responsibility to ensure that I do what I am morally required to do, so as to get you to do my bidding, I provide you with a reason for action you should not have to take into account. But the same also goes for other types of wrongful threats (discussed in the previous section). That any such threat provides you with a reason you should not have for doing the particular thing I want you to do makes it an illegitimate means for overcoming the fact that you would not otherwise have sufficient reason to do as I wish.

This allows, then, for a further elaboration of the objection to coercion. Plausibly, for me to respect you as my moral equal requires me to acknowledge that my reasons for wanting you to act in a certain way do not, as such, take precedence over the reasons you have to behave otherwise. When I get you to act as I want you to act by introducing factors that should not be relevant to your decision—as I do when I impose a responsibility on you that properly belongs to me—this manifests a failure to acknowledge you as a person with reasons of your own, which sometimes bids you to act in ways I would rather you didn't.

Note that this way of elaborating the objection marks a difference between paradigmatic cases of coercion, where the coercer makes her proposal in order to get the coercee to act in some specific way, and cases where there is no such specific aim. Consider, for instance, the following well-known example of Bernard Williams's.<sup>39</sup> Pedro is threatening to kill twenty villagers unless Jim kills one of them, in which case Pedro will release the other nineteen. What distinguishes this case from those we have been focusing on is that Pedro is indifferent to what Jim decides to do. Although Jim may certainly feel coerced into killing someone, if that is what he concludes he must do—and Pedro certainly introduces a fact that should not be relevant to Jim's choice—because Pedro's aim is not to induce Jim to act in one way or the other, his threat does not involve the element of subordinating Jim's reasons for acting in a certain way to Pedro's desire that he act in some other way.

39. Bernard Williams, "Critique of Utilitarianism," in *Utilitarianism: For and Against*, ed. J. Smart and B. Williams (Cambridge: Cambridge University Press, 1973).

Insofar as we are inclined to think that Pedro does not wrong Jim in quite the way that is characteristic of standard cases of coercion, I suggest it is because Pedro does not shift the responsibility for the villagers' lives onto Jim specifically in order to cancel out the legitimate reasons Jim has to refrain from murder.<sup>40</sup>

Let's turn, then, to (a), the intuition that a coerced choice is not a free or autonomous choice. The use of coercion not only fails to respect the fact that the victim may have good reason to act contrary to the coercer's preference; it also, we tend to think, fails to respect the victim as a person who should be free to decide for herself whether to act as the coercer wants her to.

Coercion is an affront to a person's autonomy in that its use fails to respect that person's right to make up his own mind about whether to pursue a certain option, given his circumstances and the alternatives available to him. How so? Doesn't the coercer merely help to determine what the circumstances are, with regard to which her victim is left free to respond as he sees fit? The answer is that the coercer cannot reasonably view her announced intention as a further feature of the situation to which the coercee simply has to respond. Green, for example, cannot legitimately take Brown's situation to be one in which Brown is responsible for ensuring that Green doesn't beat him up. For Green cannot consistently hold this to be Brown's situation while at the same time accepting that it is her (Green's) responsibility not to assault Brown. For Green to insist that it is up to Brown to decide for himself whether to return to Malibu under the circumstances, a feature of which is that, if he does, Green will beat him up, simply amounts to an abdication of this responsibility on Green's part. This is not a description of Brown's circumstances Green can offer in good faith. Looked at from Green's point of view, then, her threat can only be seen as interfering with Brown's ability to respond to his circumstances on his own terms. And because it is, in this sense, an interference with Brown's will aimed at ensuring he acts as Green intends for him to act, we can put our objection to the threat by saying that it amounts to an effort on Green's part to subject Brown to her will.

The problem with the type of autonomy-based approach discussed in Section III—for example, Frankfurt's account—is that it takes up the

40. Thanks to an Associate Editor for *Philosophy & Public Affairs* for encouraging me to consider this case.



question of the coercee's autonomy as if from the coercee's own point of view. But from this point of view, the threat looks like a change in the coercee's situation to which she must now (freely) decide how to respond. Viewed from the coercer's perspective, however, we can't accept this characterization of the coercee as an autonomous agent responding to unfavorable circumstances. This is because to accept this characterization would be to go along with the coercer's refusal to take responsibility for his conduct. In morally assessing the situation, it therefore makes sense to draw a contrast between the coercer's intervention and the other factors the victim must take account of, and to regard the former as an interference with, rather than a parameter of, the autonomous exercise of her will.

It's clear that not just any instance of making one's own response contingent on what another chooses to do is in this way incompatible with viewing the other as autonomously responding to his circumstances—circumstances that include one's own conditional intentions. This is because making a certain response contingent on another's decision need not express an unwillingness to assume responsibility for one's conduct in the way one should. A coach may rightly say of her players that it is for them to decide when to get to practice, given that, if they are late, they will have to run laps. The fact that she is the one who will make them run the laps does not imply that in regarding this as a feature of her players' circumstances, which they ought to take into account, she has failed to acknowledge some responsibility that properly belongs to her. Hence, the threat of having to run laps is not objectionably coercive.

If that's right, then if we consider an obviously wrongful coercive threat, the fact that the coercer effectively transfers to the coercee his responsibility for ensuring that he (the coercer) behaves in conformity with the relevant normative standards provides the ground for our assessment that the coercee's choice to comply is not a free choice. It is central to what makes the threat coercive, in the sense that it raises a distinct moral objection. In issuing such a threat, the coercer takes advantage of the power he has to act in disregard of normative standards that apply to him in order to induce another to act in the service of the coercer's ends. In doing so, the coercer fails to respect his victim, both as someone who should be free to decide for herself whether to act as directed, and as someone who may well have legitimate reason for choosing not to do so.