

This article was downloaded by: [University of St Andrews]

On: 04 April 2013, At: 06:16

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH,
UK



Australasian Journal of Philosophy

Publication details, including instructions for authors
and subscription information:

<http://www.tandfonline.com/loi/rajp20>

Why Coercion is Wrong When It's Wrong

Benjamin Sachs ^a

^a New York University

Version of record first published: 22 Dec 2011.

To cite this article: Benjamin Sachs (2013): Why Coercion is Wrong When It's Wrong,
Australasian Journal of Philosophy, 91:1, 63-82

To link to this article: <http://dx.doi.org/10.1080/00048402.2011.646280>

PLEASE SCROLL DOWN FOR ARTICLE

For full terms and conditions of use, see: <http://www.tandfonline.com/page/terms-and-conditions>

esp. Part II. Intellectual property and access and license types, § 11. (c) Open Access Content

The use of Taylor & Francis Open articles and Taylor & Francis Open Select articles for commercial purposes is strictly prohibited.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

WHY COERCION IS WRONG WHEN IT'S WRONG

Benjamin Sachs

It is usually thought that wrongful acts of threat-involving coercion are wrong because they involve a violation of the freedom or autonomy of the targets of those acts. I argue here that this cannot possibly be right, and that in fact the wrongness of wrongful coercion has nothing at all to do with the effect such actions have on their targets. This negative thesis is supported by pointing out that what we say about the ethics of threatening (and thus the ethics of coercion) constrains what we can say about the ethics of warning and offering. Importantly, our favoured explanation of the wrongness of certain kinds of threatening should not commit us to condemning as wrong parallel cases of warning and offering. My positive project is to show how this can be done. I defend the claim that wrongful coercion is nothing more than the issuing of a conditional threat to do wrong, and that an agent's issuing of a conditional threat to do wrong is wrong because it constitutes motivation for that agent to adopt the announced intention to do wrong. The idea of explaining the wrongness of wrongful coercion in this way has gone unnoticed because we have thus far been mistaken about what a threat is. In this essay I present my moral analysis of coercion only after presenting a careful descriptive analysis of threats. On my view, it is essential to a threat that the announced intention is one that the agent does not possess before announcing it. This analysis makes it possible to elucidate the descriptive differences between threats, warnings and offers, which sets up the later project of elucidating the moral differences between them.

Keywords: coercion, threat, warning, offer, intention

1. Introduction

At the broadest level there are at most two kinds of coercion: the kind that involves the use of force and the kind that involves the use of a conditional threat.¹ In this essay I'm concerned solely with the latter kind. My goal is to present a novel theory about why this kind of coercion (hereafter referred to simply as 'coercion') is wrong when it's wrong. But, before proceeding, three points about methodology are in order.

First, the fact that I aim to explain why coercion is wrong *when it's wrong* should not be taken as a sign that I believe coercion is sometimes not wrong. Rather, it represents an attempt to remain neutral on the question. My plan is to identify a class of actions that uncontroversially counts as coercion, and

¹I say 'at most' since some people reject the idea that the use of force counts as coercion. Lamond [1996] took this stance, though he later rescinded it [2010]. The point needn't be argued here.

is uncontroversially wrong, and explain why those actions are wrong. I will then argue that an explanation of why those actions are wrong constitutes an explanation of why coercion is wrong when it's wrong.

Second, one might be tempted to object straight off that I am begging important questions by assuming that only threats can constitute coercion and hence implicitly denying that people can be coerced by offers or their circumstances and that offers and circumstances can be coercive. Actually, however, I am not denying this. I am denying that an offer can constitute *coercion* and that one's circumstances can constitute *coercion*. 'Coercion' is the name of a type of action. 'Coercive' describes an effect that some actions and circumstances have on people, who are in turn 'coerced' by those actions and circumstances. I do not insist that only coercion is coercive. Now one might worry that if other things besides coercion can be coercive, then it makes little sense to confine one's efforts to explaining why coercion is wrong when it's wrong. One should, rather, try to explain why it's wrong, when it's wrong, to do things that are coercive.

But I reject this strategy, and the reason for this rejection is my third point: coerciveness is a red herring. Coercion, when it is coercive, is wrong for a reason having nothing to do with its being coercive. Coerciveness, whatever it is, is some effect that an action or circumstance can have on an individual. Yet I deny that acts of coercion are wrong in virtue of the effect they have on the coerced individual. I mention this view now not because I expect the reader to accept it; on the contrary, trying to establish it will occupy a large part of this essay. Rather, I simply want to explain why I focus my attention on coercion instead of coerciveness.

Thus, my aim here is to explain why coercion is wrong when it's wrong without appealing to the effects acts of coercion have on the coerced individual. My initial focus will be a class of actions that, it is generally agreed, is wrong and counts as coercion: conditional threats to do wrong. So I am out to explain the principle

(P1) Conditionally threatening to do wrong is, *per se, pro tanto* wrong.

The '*per se*' qualifier is important here. In explaining the wrongfulness of conditional threats to do wrong, we must avoid appealing to features that only some such threats possess; otherwise we will not have explained the wrongfulness of the entire set of such threats.

Furthermore, I am out to explain the principles

(P2) Unconditionally threatening to do wrong is, *per se, pro tanto* wrong.

and

(P3) Neither conditional warnings about one's own future wrongdoing nor conditional offers to do wrong are, *per se, pro tanto* wrong.²

²By 'conditional offers to do wrong', I mean an announcer's offer to do something that would wrong the target if and only if the target did not accept the offer. This definition might seem to render paradoxical the idea of conditionally offering to do wrong. When an agent issues an offer to ϕ he usually means to say

Just what P2 and P3 have to do with P1 will be made clear later. Briefly, however, P2 and P3 are relevant because there are no extant explanations of P1 that are consistent with P2 and P3. This constitutes a serious objection to the extant explanations of P1 and constitutes a large part of the reason why I reject those explanations. Of course, it is possible that we cannot have our cake and eat it too. Perhaps in order to explain one of the principles we have to give up another one. I take it, however, that these principles are intuitively plausible enough to make it worth our while to try to hang on to all of them.³

Before P1–3 can be explained, however, we need to distinguish conditional threats from unconditional threats, conditional warnings about one's own intentions, and conditional offers. This is the task I undertake in §2. All four kinds of utterance are intention-announcements, so it requires a bit of work to distinguish them from each other. My way of distinguishing these four kinds of utterance is, I will argue, defensible independent of the fact that it makes it easier to explain P1–3. My strategy will be to first defend what I will call the Non-Adoption Account, a descriptive account of conditional threats. The reason why we should accept the Non-Adoption Account, I shall argue, is that doing so is the only way to differentiate conditional threats from conditional warnings while explaining how it is possible for people to issue conditional threats. (Importantly, I will not attempt to explain how this is possible *for rational people*.⁴ I am going to try to explain how this is possible for people given the actual facts of human psychology, ignoring all questions about the extent to which this psychology is rational. This is the sensible approach for someone who is interested, as I am, in exploring the ethics of threat issuing. We should want to know how threats work when used by actual people.)

In what follows, I will use the word 'announcer' to refer to the individual making an intention-announcement. I will use 'target' to refer to the individual to whom the announcement is addressed.

2. What a Conditional Threat Is

2.1 Nozick's Account

To appreciate the difficulty in explaining how conditional threats are both possible and distinguishable from conditional warnings, consider the descriptive account of conditional threats found in Robert Nozick's classic

(among other things) that he will not ϕ if the offer is not accepted. But if the agent treats the target's non-acceptance as a bar to ϕ -ing, and ϕ -ing would be wrong only if it was done without the target's acceptance, then it would seem that the agent's announcement does not actually raise the prospect of his wronging the target. So he doesn't really offer to wrong the target; at best he offers to 'otherwise-wrong' the target. I accept all this. Henceforth when I refer to an 'offer to do wrong' I mean 'offer to do otherwise-wrong'.

³I do not assume that the three principles are unimpeachable in their current form. I do not take up objections to these principles, nor do I tinker with the principles in order to make them impervious to those objections. This is a worthwhile project, but it is not my project. I aim only to explain the core truth that these principles capture.

⁴Accordingly, I do not seek to make any contribution to philosophical thought about the rationality of threatening or to game-theoretic thought about the likely outcomes of various uses of threats by rational agents.

[1969] paper. By way of motivating his account, Nozick offers the example of a labour dispute [ibid.: 453]:

[A]n election is about to be held in a factory to determine whether the employees will be represented by a labor union. The owner of the factory announces to his employees that if the union wins the election, he will close his factory and go out of business.

Nozick assumes, implicitly, that the owner of the factory adopted the announced intention before announcing it. The important question, according to Nozick, is why he did so; it is the answer to this question that determines whether the announcement was a threat. As Nozick is aware, if the owner adopted the intention for self-interested reasons (because it'd be good business, for example), we would think that his announcement of that intention was a warning [ibid.: 456]. Nozick suggests that the announcement qualifies as a threat if the owner adopted the announced intention for a strategic reason [ibid.: 455]. If, that is, the factory owner adopted his conditional intention in order to be able to credibly threaten his employees and thereby alter their behaviour. It is problematic, however, for Nozick to portray the factory owner as adopting an intention for a strategic reason.

To see why, a distinction first needs to be made. At the broadest level, there are two kinds of reason for adopting an intention [Parfit 2001]. First, there are object-given reasons: reasons having to do with the object of the intention. An object-given reason to adopt an intention to go to California would be a reason having to do with going to California. More technically, an object-given reason to adopt an intention to go to California is any reason that is (or is grounded in) a property of that prospective action: going to California. Second, there are state-given reasons: reasons that are (or are grounded in) properties of the state of affairs of possessing the intention. It's been heavily debated whether it is rational to adopt an intention for a state-given reason (see, for instance, Hampshire and Hart [1958], Anscombe [1963], Kavka [1983], and Gauthier [1994]). What has gone unquestioned thus far is whether it is *possible* to adopt an intention for a state-given reason. This is an empirical question regarding which no evidence has been systematically collected. Importantly, it is for our purposes a question about the psychology of actual humans, and so we can make at least some progress on it through simple introspection. For my part, when I imagine attempting to adopt an intention for a state-given reason, it strikes me as impossible that I could succeed. I will assume henceforth that actual humans cannot, in fact, do this. This is crucial for Nozick's descriptive account of threats, since a strategic reason to adopt an intention is a state-given reason. And so what Nozick describes is, given my assumption, impossible.

2.2 *A New Account*

Kent Greenawalt has hinted at a solution to this problem. He offers a thought experiment featuring Tim and Vicki. Tim wants to make love with Vicki, and tells Vicki that if she refuses to make love with him he will tell

Vicki's parents, who are ardently opposed to abortion, that she recently had an abortion. Greenawalt, explaining why this announcement qualifies as a conditional threat (or a 'manipulative threat', in Greenawalt's parlance), says the following [1983–4: 1098; see also 1989: 94–5 and 66–7]:

If Tim had for some reason only been able to communicate his wish to make love but not the threat, his reasons for telling her parents would never come into existence. We might imagine, for example, that Tim and Vicki speak different languages, and that Tim finds it much easier to convey his desire for sexual relations than a threat to report a past behavior. If Tim did find himself unable to make Vicki understand the threat, he would see no point in carrying it out.

What Greenawalt seems to be saying is that before Tim is able to communicate the conditional intention to tell Vicki's parents of her abortion, he does not have that conditional intention.^{5,6} I submit that this is an essential feature of a conditional threat: it is the announcement of an intention that one has not adopted before announcing it. Call this the Non-Adoption Account of conditional threats.

The Non-Adoption Account certainly explains how it is possible to issue a conditional threat, in that it does not attribute any unusual capability to conditional threat-issuers. Yet it seems to introduce a new problem: how can a threat be effective in altering a target's behaviour? If the target knows that the announcer doesn't actually have the conditional intention that he announces, why would the target ever be moved to give in to the announcer's demand?

Greenawalt does not try to answer this question. But he could have pointed out that although the announcer does not have his announced conditional intention before announcing it, he might acquire that intention contemporaneous with or subsequent to his announcement. As Greenawalt [1983–4: 1098] notes, the announcer might have a desire to stick to his word, and this could motivate him to adopt his announced intention before the time comes to follow through (see Lamond [1996] for similar comments). This is basic human psychology, and consequently the target should be aware of it. This, then, explains how conditional threats can be effective.

Note that we need not assume that the motivation is always sufficient to induce the announcer to adopt the announced intention; this assumption is clearly false. Many conditional threats are not followed through. Fortunately, we can concede this and still explain why they are sometimes effective. While it is common knowledge that many conditional threats are not followed through, many people who find themselves the target of such a

⁵I am not sure whether Greenawalt understood himself as making this claim or appreciated the difference between this claim and Nozick's—the claim that in the case of a conditional threat the announced intention is adopted in order to make it possible to issue a credible threat. (For evidence of possible confusion on Greenawalt's part, see Greenawalt [1989: 255]).

⁶It has been suggested to me that although Tim has not already adopted the conditional intention before announcing it, he may have already *conditionally* adopted the intention before announcing it. In other words, he may already have a plan to adopt-it-if-he-announces-it. This makes no trouble for me. Although it is true that Tim already has a doubly conditional intention before he makes his announcement, it is enough for my purposes if we can say that the intention Tim announces—the singly conditional intention—is not an intention that he has already unconditionally adopted.

threat have no reliable way of anticipating whether the announcer will follow through. This uncertainty can explain why such targets often give in.

Why should we accept that this really explains what is going on in some cases of effective conditional threatening? First, as already noted, this explanation appeals only to familiar psychological facts. Second, as a matter of consistency it seems we must accept it. When we want someone to do something we often ask that person to say that he is going to do it. We do this even in those cases in which the other person has already expressed a willingness to engage in the desired behaviour. For us, the intention-announcement itself seems to be important. Why would this be? We believe, I submit, that the person's motivation to adopt an intention to ψ will increase *because* he announces an intention to ψ . This belief explains our announcement-seeking behaviour with respect to behaviour that we desire just as it explains our announcement-aversion with respect to behaviour that we do not desire (e.g., our aversion to being threatened). The third reason to accept the Non-Adoption Account is that the truth of this account makes it possible to draw easy, intuitively appealing distinctions between conditional threats and the other kinds of intention-announcement, as I shall now illustrate kind-by-kind.

2.3 Distinguishing Conditional Threats from Unconditional Threats

An unconditional threat, unlike a conditional threat, is not (typically) a good vehicle for getting one's target to behave a certain way. It may seem, in fact, that issuing an unconditional threat would be pointless, and so one may wonder whether there really is such a thing. However, any temptation to draw this conclusion should vanish given the stipulation that we are concerned with actual people, not rational people. It may not be rational to issue an unconditional threat, but undoubtedly people do it. People really do say things like, 'One of these days, I'm going to make you sorry for what you did to me.' I suspect that such utterances are typically motivated by a desire to strike fear in others or a desire to manipulate oneself into adopting a particular intention.

The same argument that demonstrated the need for a Non-Adoption Account of conditional threats applies, *mutatis mutandis*, to unconditional threats. We need an analysis of unconditional threats that distinguishes them from unconditional warnings while not rendering it impossible to issue an unconditional threat. To the best of my knowledge, there is no way to do this without insisting that in cases of unconditional warning the announced intention is an intention the announcer has not adopted before announcing it. So the Non-Adoption Account applies to both conditional and unconditional threats. The only difference between the two, I submit, is whether the announced intention is conditional.

2.4 Distinguishing Conditional Threats from Conditional Offers

Next, consider conditional offers. David Zimmerman [1981] says that a conditional offer is a proposal such that its target prefers his post-proposal

situation to his pre-proposal situation, whereas a conditional threat is a proposal such that its target prefers his pre-proposal situation to his post-proposal situation. The 'situation' is constituted by the options the target has. Yet while a target's option set may be shaped in part by the announcer's intentions, it is never affected by what the announcer *says* his intentions are. Thus, threats don't alter situations in the relevant sense.

A different account is offered by Nozick, Alan Wertheimer and Harry Frankfurt, who distinguish conditional offers from conditional threats by saying that in the case of a conditional offer, the announced conditional intention is an intention to make the target better-off, whereas a conditional threat is the announcement of a conditional intention to make the target worse-off [Nozick 1969: 447–53; Wertheimer 1987: ch. 12; Frankfurt 1988: 28].⁷ This way of drawing the distinction, however, cannot allow for what I will call *ridiculously bad conditional offers*. These are conditional offers such that a) the thing that the announcer conditionally offers to do is objectively bad for the target, and b) the announcer and the target both know this. I don't have any persuasive examples of such offers; any example would appear bizarre, seeing as the very act-type is itself bizarre. But the point is simply that it is possible in principle for there to be ridiculously bad conditional offers; it is not a necessary truth about offers that they must not be ridiculously bad. (Remember, we are not making any assumptions about the rationality of human behaviour.)

This datum—that there can be ridiculously bad conditional offers—raises an interesting question: what's the difference between making a ridiculously bad conditional offer and making a conditional threat? The answer, I think, is clear: the target of a ridiculously bad conditional offer gets to accept or reject it,⁸ whereas the target of a conditional threat has no such power. When an announcer threatens, 'If you don't ϕ I will ψ ', he gives his target the power to cause or prevent the announcer's ψ -ing (by ϕ -ing or refraining from ϕ -ing) but does not give his target the power to cause or prevent the announcer's adoption of the conditional intention to ψ . The recipient of an offer, by contrast, is endowed with this power.

Surprisingly, then, sometimes when an announcer utters something of the form, 'If you don't ϕ I will ψ ', he does not actually assert that he has adopted or will adopt a conditional intention to ψ . Sometimes what he asserts is that if the target accepts then he (the announcer) will adopt a conditional intention to ψ . I maintain that conditional offers essentially have this propositional content. Further, I submit, conditional threats essentially do not.⁹ (When the mugger says, 'If you don't give me your wallet I'll shoot

⁷Similarly, Lamond [1996, 2010] suggests that a conditional threat is the announcement of a conditional intention to do something that the target wants not to occur or believes will make him worse off. The objection I am about to raise against the Nozick/Wertheimer/Frankfurt account applies to Lamond's as well.

⁸And thereby return to the *status quo*.

⁹This claim has been disputed. The most famous challenge to it is Nozick's slave case. In Nozick's words, 'Suppose that usually a slave owner beats his slave each morning, for no reason connected with the slave's behavior. Today he says to his slave, "Tomorrow I will not beat you if and only if you now do A"' [1969: 450]. Here the slave is (apparently) given the power to control whether the slave-owner adopts his announced conditional intention, and yet Nozick considers it reasonable to categorize the utterance as a threat. I suspect, however, that this willingness arises from the prior commitment, held by Nozick, that threats and offers are distinguished from each other according to the content of the announced intention. If one were not in the grip of this theory, it would strike one as obvious that the slave owner's utterance is an offer. Alternatively, it may stem from a desire to accommodate the claim that the slave-owner's offer is coercive. However, once we

you', he is not asserting that you need to accept *and* fail to give him your wallet in order to get shot.) Conditional offers and conditional threats are similar, then, in that the Non-Adoption Account applies to them. The difference is that in the case of a conditional offer, but not in the case of a conditional threat, the announcer asserts that some action on the target's part (an acceptance) is required to trigger his (the announcer's) adoption of the announced conditional intention.

2.5 Distinguishing Conditional Threats from Conditional Warnings about One's Own Future Intentions

There is something suspicious about conditional warnings about one's own future intentions. In issuing a conditional warning about one's own future wrongdoing one speaks of one's agency in an impersonal way, as if it were not under one's control. This may give rise to a suspicion that any talk of an announcer's issuing a conditional warning about his own intentions is at best just a euphemistic way of referring to the announcer's issuing of a threat.

This suspicion, however, would be misguided. The clearest real examples of warnings about one's own intentions are the 'I know myself' cases—instances in which an agent anticipates that some turn of events could trigger malevolent behaviour on his part and would like to avert the triggering event. Even the sub-class of warnings about one's own future wrongdoing contains genuine instances. These would be cases in which an individual can anticipate his own future wrongdoing, and not because he anticipates losing his moral agency but rather because he anticipates that his concern to do right might be trumped by something else he cares about. I have in mind, for instance, the protective parent who warns, 'If you lay a finger on my child ...'.

Greenawalt [1983–4: 1096] claims that, when one sincerely issues a warning about one's own intentions, one already has the announced intention. This feature is present in the 'I know myself' cases and, I submit, is an essential feature of warnings about one's own intentions. Thus, conditional warnings about one's own intentions and conditional threats differ in that the Non-Adoption Account applies only to the latter.

If this is true, then there should be no fool-proof way to discern whether one has been conditionally threatened or conditionally warned. If we reflect on what it is like to be a target of a malevolent intention-announcement, I think we will find that this is true. Specifically, we should take note of the familiar fact that it is often difficult to know how to respond to one. (Consider, for instance, the situation faced by Nozick's factory workers after their boss's announcement that if they unionize he will close the factory.)¹⁰

realize that other things besides coercion can be coercive, there is no pressure to label the slave-owner's announcement a threat (and hence coercion) on account of its coerciveness.

¹⁰Since it is often difficult to know whether one has been threatened, it will also often be difficult to know whether one has been wrongfully coerced.

This concludes my descriptive analysis of the four kinds of intention-announcement. I have identified one or two necessary features of each kind. For the purposes of this paper, that is as good as offering a complete conceptual analysis of each kind. My goal here is to explain P1–3, and to do that we simply need to be able to correctly categorize any intention-announcement with which we are confronted. The work I have done in this section allows us to do that.

3. Explaining P1: Conditionally Threatening to Do Wrong Is, *Per Se*, Wrong¹¹

3.1 *The Inadequacy of the Extant Explanations of P1*

I want to concede, straight off, that there is a good extant explanation of why conditionally threatening to do wrong is sometimes wrong: conditionally threatening to wrong a person can, for example, scare, humiliate, or disgrace that person. However, this is not an explanation of P1. An explanation of P1 would tell us why conditionally threatening to do wrong is wrong *per se*, not merely on those occasions when it has certain effects.¹² It is this kind of explanation, I will argue here, that has not yet been offered. And it will not do to simply point out that it is wrong to adopt a conditional intention to do wrong.¹³ This is an answer to a different question. Our question is why it might be wrong to announce such an intention. *That* question presents something of a conundrum: how can the announcement of a malevolent intention add anything to the wrongness of adopting it and following through on it? Only an answer to this question will count as an explanation of P1.

The extant answers to this question can be divided into two kinds. The first kind of answer homes in on the effects that such an announcement has on its target's circumstances. The second kind focuses on the effects that such an announcement has on the target's mental states.

Let's start with the first kind of answer. What objectionable effect might a conditional threat to do wrong have on the target's circumstances? I am aware of just one answer to this question: threats to do wrong are wrong because they *wrongly* restrict the target's options [Fowler 1982; Zimmerman 1981; Pallikkathayil 2011].¹⁴ Suppose the announcer says, 'If you don't ϕ I will ψ ', where the announcer's ψ -ing would wrong the target. Whereas before the utterance, the target had the option of not- ϕ -ing and avoiding the announcer's ψ -ing, the target no longer has this option [Nozick 1969;

¹¹For the rest of the paper I drop the cumbersome '*pro tanto*' qualifier. It should be understood as implicit in all my future uses of 'wrong'.

¹²Thanks to Japa Pallikkathayil and Alan Wertheimer for pushing me to clarify my stance on this.

¹³Grant Lamond [2002: 49] seems to believe that pointing this out amounts to explaining P1. He says, '... what makes all cases of [coercion by threat] stand in need of a defence is the fact that they involve the plan purposely to bring about something unwelcome to another'. See also Haksar [1976], who believes coercion is wrong for a second reason as well—namely, it places wrongful pressure on the coerced party. For an explanation of the problem with this proposal, see note 20.

¹⁴Pallikkathayil holds that this wrong gives rise to an additional wrong. For a sceptical view about the moral importance of the structure of one's options, see Wertheimer [1987, ch. 11].

Feinberg 1986: 191–2; Raz 1986: 150]. Since the threatener’s ψ -ing would, *ex hypothesi*, wrong the target, the threat robs the target of an option to which he is entitled. But this won’t do as an explanation of P1. As I mentioned in §2.4, the announcing of an intention cannot change the target’s available options.

So we turn to the second kind of answer. What objectionable effect might a conditional threat to do wrong have on the target’s mental states? The common answer to this question is that the announcement can constitute an invasion or undermining of the target’s freedom or autonomy. In order for this suggestion to get us anywhere, we are going to need an analysis of the undermining of freedom/autonomy that does not reduce it to some alteration of the target’s circumstances. Just such analyses have been offered by Harry Frankfurt and Joseph Raz. Frankfurt [1988: 44–5] maintains that an individual’s freedom or autonomy is undermined when he is caused to behave in accordance with a motive by which he would prefer to not be moved, while Raz [1986: 148–57, 410] contends that autonomy is invaded when A subjects B to A’s will.¹⁵ But these analyses of the undermining of freedom of autonomy cannot be used as the basis of an explanation of P1.¹⁶ The resulting explanation would explain both too little and too much.

It would explain too little because conditional threats influence the target’s motivation or will only when they are successful, so the objection would apply only to successful conditional threats. Yet it seems that the wrongness of issuing a conditional threat has nothing to do with whether it turns out to be successful (we don’t think that people who bravely resist conditional threats of wrongdoing have not been wronged). Hence it would not explain why conditionally threatening to do wrong is wrong *per se*, and thus it wouldn’t qualify as an explanation of P1.^{17,18}

The purported explanation of P1 would explain too much in two ways. First, conditional threats can influence the target’s motivation or will regardless of whether the threatened action would be wrong. So this objection to conditionally threatening to do wrong would extend to conditionally threatening to do not-wrong. This runs counter to the widespread intuition that some successful conditional threats are entirely unobjectionable [Greenawalt 1989: 98], such as the threats that parents use against their children to ensure their good behaviour. Second, as we learned in §2, there need be no difference in propositional content between conditional threats on the one hand and conditional warnings about one’s own intentions and

¹⁵For Raz, this is just one way of invading someone’s autonomy. The other is by restricting a person’s options in a significant way. But we have already seen why this cannot be the basis of an explanation of P1.

¹⁶For Frankfurt and Raz, they weren’t supposed to.

¹⁷It has been suggested to me that the mere attempt to undermine someone’s autonomy or freedom might be morally objectionable. But it strikes me as implausible to suggest that very many threat-issuers are attempting to do such a thing to their targets—at least they are not attempting it in the sense of having an intention to do it (which is the morally relevant sense of ‘attempting’, I take it). Whether the motive that he forces his target to act upon is one that the target finds distasteful is totally unimportant to the announcer (unless he’s sadistic). And whether the target’s will becomes subject to the agent’s is, again, unimportant. The important thing to the announcer is just that the target acts a certain way.

¹⁸Another point about this explanation of P1 is that it could not be extended to unconditional threats; unconditional threats are not designed to influence the target’s motivation or will. So this objection to conditional threats to do wrong would leave us at square one in attempting to explain P2. Of course, we cannot rule out the possibility that P1 and P2 have distinct explanations.

conditional offers on the other; the difference can be all in the head of the announcer. Thus, it stands to reason that if conditional threats can alter their target's mental states in the ways Frankfurt and Raz describe, conditional warnings about one's own intentions and conditional offers can as well.¹⁹ So if we say that threats to do wrong are wrong in virtue of the effect they have on their target's mental states, then we have no choice but to say that conditional warnings about one's own wrongdoing and conditional offers to do wrong are objectionable in the same way. Thus, we render ourselves unable to explain P3.²⁰

Let me briefly summarize what I have argued in this sub-section. We are looking for an explanation of P1, and we found that conditional threats to do wrong can have effects on the mental states of the target, but none of these effects is unique to conditional threats. In particular, they are shared by conditional warnings of one's own future intentions and conditional offers. Furthermore, they are *not* shared by unconditional threats. So this way of explaining P1 puts P2 and P3 in jeopardy. The other strategy for explaining P1 is to point out the effects that a conditional threat to do wrong has on its target's circumstances. But it turns out that conditional threats have no such effects at all. The lesson to be learned from this sub-section, then, is that we cannot appeal to the effects of threats to do wrong on the target's mental states or circumstances in order to explain their wrongfulness. This being the case, then *a fortiori* we cannot appeal to the coerciveness of coercion in order to explain its wrongfulness, since the coerciveness of an action, whatever it is, surely is some kind of effect that the action has on the its target's mental states or circumstances.

3.2 A New Explanation of P1

Given the inadequacy of these explanations of P1, it would be worth our while to ask whether there are any explanations of P1 that naturally follow from the Non-Adoption Account. One would hope that we could look to Nozick or Greenawalt for an answer to that question, since their work on conditional threats formed the basis for the Non-Adoption Account.

¹⁹This is confirmed by ordinary experience. Conditional warnings and conditional offers can cause us to act from unsavoury motives (in the case of conditional warnings—fear; in the case of offers—greed or desperation). They can also be used to influence a target to act in the way the announcer would have him act, which strikes me as the best gloss on what it means to make one person's will subject to another's.

²⁰Mitchell Berman identifies the wrong of coercion (and by extension, conditional threatening, I take it) as having to do with the fact that coercion puts pressure on the coerced party. This appears to be another appeal to the effects that conditional threat issuing can have on the mental states of one's target, but it is not vulnerable to the objection I just levelled against other versions of that appeal. This is because Berman specifies that coercion places '*wrongful* pressure on [the target's] freedom to choose otherwise' [2002: 53]. Greenawalt [1989: 97–8, 104] claims, similarly, that one reason why some conditional threats may be properly criminalized is that they place improper pressure on the target.

The insertion of the '*wrongful*' or '*improper*' qualifier makes it possible to avoid some of the problems that confronted the Frankfurt/Raz-inspired explanation of P1. In particular, it does not commit one to saying that warning-issuers and offer-issuers do wrong when they put pressure on their targets. So we can say that inflicting pressure by announcing a conditional intention to do wrong is only sometimes wrong. And this would be because the pressure itself—the very same pressure—is only sometimes wrongful or improper. But why would *that* be? Certainly it's not a brute fact. Unfortunately, no explanation of it jumps out; certainly neither Berman nor Greenawalt has attempted to explain it. So this explanation of P1 does not explain what is to be explained.

Unfortunately, neither Nozick nor Greenawalt—nor Alan Wertheimer, whose descriptive account of conditional threats [1987: 96–9] largely follows the Nozick/Greenawalt account—have ever tried to answer this moral question about conditional threats.

First, let's distinguish two kinds of case. The Non-Adoption Account tells us that when an announcer issues a conditional threat, he does not, before making his announcement, have the intention that he announces. But there are two ways of lacking an intention to ψ : one either has an intention to not- ψ (or an intention to φ , where φ -ing is incompatible with ψ -ing and one knows this) or no intention at all with respect to ψ .

Take the first case—the case of bluffing. A bluffing threat is a lie. Thus, it is vulnerable to whatever objection applies to lying and wrong in any context in which lying is wrong. In those cases in which the bluff is used successfully to extract something from the target, we can say furthermore that the announcer has defrauded the target.

The second case is more complex. To ground an objection to this kind of conditional threat, we need two premises. First, a general moral principle: if it would be wrong for an agent to ψ at t_2 , then it would be wrong for that agent to give himself at t_1 significant motivation to adopt an intention to ψ -at- t_2 .

The second premise is that an individual's threatening to ψ significantly motivates him to adopt an intention to ψ . According to the Non-Adoption Account, when a person conditionally threatens to ψ he does not, before making his announcement, conditionally intend to ψ . And we have seen how issuing a conditional threat can motivate a person to adopt his announced intention (§2.2).

These two premises yield the conclusion that if it would be wrong of an announcer to ψ at t_2 , then it would be wrong of that announcer to issue a conditional threat at t_1 to ψ -at- t_2 . This is just a restatement of P1.

Of course for a certain kind of person, this objection would not apply. Premise two says that the issuing of a threat is a significant source of motivation to do what one threatens to do. There could be people for whom this is not true—people who are left more or less cold by all the usual considerations of self-respect, strategy, and character that motivate most people to keep their word. Whether this amounts to much of an objection to my argument here depends, I think, on whether we think it's wrong for such people to conditionally threaten to do wrong. (Remember, I am trying only to explain why coercion is wrong *when it's wrong*.) This is something we almost never have occasion to ponder, since it's nearly impossible to identify people who are left cold in this way. (They have strong reasons to conceal this fact about themselves.) Indeed, I find it difficult to even paint a mental picture of such a person. If she doesn't care about self-respect, strategy or character, what *does* she care about? Is she a sociopath? A nihilist? Neither? Is she a moral agent? I don't know what to say.²¹

²¹I have been considering the possibility that premise two is false in the case of psychologically atypical people. Another possibility is that it is false in atypical circumstances (even for psychologically normal people). I was alerted to this point by an anonymous referee, who offered the following thought experiment: Mobster Tony Soprano tells his daughter, Meadow, that he will kill her boyfriend unless she breaks up with him. The fact of the matter is that Tony has never even considered killing Meadow's boyfriend. And as soon as Tony hears himself saying this, he is immediately horrified by the idea of actually doing it. Contemplating

Another worry about the second premise is that it is internally inconsistent. How can it be that in cases of threatening the announcer does not describe himself as having any particular intention *and* that the issuing of the threat motivates the announcer to keep his word? If the announcer has not asserted anything about himself, what word is there for him to keep?

By way of response, there is nothing in the Non-Adoption Account that precludes us from saying that in issuing a threat an announcer asserts that he has the intention he announces; the theory simply says that he doesn't actually have that intention (before announcing it). Indeed, it would make sense for the threat-issuer to make such an assertion, since threat-issuers should want to be perceived as warning-issuers. I am not sure whether to endorse this response, since it has the somewhat counterintuitive implication that all threats are insincere. Fortunately, there is a more surefire response available: while threat-issuers might not portray themselves as *currently* possessing any particular intention, they certainly do portray their future selves—the selves that will exist when the time comes to act or not act—as having their announced intention. This utterance is an assertion, and thus it puts the announcer in a position to keep or break his word.²²

Thus, I think the second premise can be vindicated. As to the first premise, the only worry I can see being raised is that motivating oneself to adopt an intention to do wrong is too indirectly related to wrongdoing itself to be objectionable. The indirectness here is twofold. First, gaining motivation to adopt an intention to ψ does not guarantee that one will adopt an intention to ψ . But we do think that motivating others to adopt wrongful intentions can be wrong, even though the mere provision of motivation is insufficient to guarantee such adoption. We call it incitement, and I see no reason why self-incitement would be less objectionable than other-incitement. Second, adopting an intention to ψ does not guarantee ψ -ing. This is true—intentions can always be rethought—but it does make ψ -ing more likely. As Raz [1975] points out, adopting an intention to ψ necessarily involves terminating, at least temporarily, one's deliberations about whether to ψ . In this way, adopting an intention to ψ makes ψ -ing the default course of action, and thereby makes it more likely that one will ψ .²³

the act for the first time has such a profound effect on Tony that it settles for him decisively the question of whether he will under any circumstances commit the murder.

Given my descriptive account of conditional threats, Tony has issued a conditional threat. So according to premise two of the argument of this section, Tony's announcement should cause him to become significantly motivated to adopt the announced intention. Plausibly, however, it doesn't. Consequently, the argument I have just given for the wrongness of conditionally threatening to do wrong doesn't explain why what Tony did was wrong.

I concede all this, but maintain that a different argument for the wrongness of conditionally threatening to do wrong can be offered to deal with cases of this kind. By 'of this kind', I mean cases in which, owing to circumstances that the threat-issuer cannot have foreseen, his issuing of the threat has the effect of making it less likely that he will adopt the announced intention. In cases like this, we should say that the announcer does wrong in virtue of acting negligently with respect to the possibility of becoming motivated to adopt the wrongful intention. Remember, the fact that announcing an intention that one doesn't yet have serves as motivation to adopt it is basic human psychology. So one can be expected to anticipate what will happen as a result of one's threat-issuing.

²²Thanks to Collin O'Neil for discussing this and the previous objection with me.

²³Thomas Pink [2008: 153–4] goes so far as to say that adopting an intention to ψ causes one to ψ .

Since the argument is valid and its premises appear to be true, it seems we must accept the conclusion—P1.²⁴

3.3 Conditional Threats and Coercion

My goal in this essay is to explain why coercion is wrong when it's wrong. Since coercion involves the issuing of a conditional threat, I have sought to explain why issuing a certain kind of conditional threat—a conditional threat to do wrong—is wrong. But does that suffice to explain why coercion is wrong when it's wrong? That depends on whether some other element of coercion is wrongful. Suppose, for the sake of argument, that there are cases in which coercion involves a threat to do something that's not wrong. In these cases coercion is not wrong. So we may continue to focus on cases of coercion that involve a conditional threat to do wrong. Our only remaining questions, then, are these: When an act of coercion includes a conditional threat to do wrong, does it include anything more than that? And if so, are any of those other things ever wrongful?

There are three things that can happen when a person issues a conditional threat to do wrong. (In what follows I use the proposition, 'If you don't ϕ I will ψ ', as a template for such a threat.) The first possibility is that the target stands his ground and the announcer doesn't ψ . In other words, after the conditional threat is issued nothing happens. In this case, coercion involves nothing more than the issuing of a conditional threat.

In the second kind of case, the target stands his ground and the announcer does ψ . Does his ψ -ing count as part of the act of coercion? Surely not. The target put an end to the coercion when he stood his ground.²⁵ Now the announcer is just ψ -ing, nothing more.²⁶ So in this case, too, coercion involves nothing more than the issuing of a conditional threat.

The third kind is the case in which the target gives in—he ϕ s. Is the coercion ongoing while the target ϕ s, or did the coercion end with the issuing of the conditional threat? I'm not sure how to answer this question, but I'll grant for the sake of argument the claim that coercion can be temporally extended such that it includes the period during which the target ϕ s. But even if coercion is temporally extended, it need contain only one

²⁴Even if we cannot identify a premise to reject, perhaps it's reasonable to withhold our acceptance of P1 if it has severely counterintuitive implications. I have assumed thus far that the conclusion is one we already accept and simply need to explain, but perhaps this is not the case. There is one case, the case of nuclear deterrence, which is often thought to present a challenge to P1 (as noted by Scott Anderson [2011: §3.2]). In these cases, Country A threatens to deploy nuclear weapons against Country B if Country B deploys nuclear weapons against Country A. Of course, nuclear deterrence is just one member of a class of cases in which one threatens wrongdoing in order to deter wrongdoing. Another case—this one by way of Alan Wertheimer [1987: 102]—is saying to a burglar, 'I'll shoot you if you take my television'. Many people, and I must admit I'm one of them, are tempted to say that it is morally permissible to issue such threats. So P1 might be in need of some tinkering. But this truly would be mere tinkering. Such counterexamples arise only in degenerate contexts. In the central contexts P1 is true. This assertion represents the sum total of my allegiance to P1, and the sum total of what I have sought to explain.

²⁵Some theorists believe that if the target stands his ground, then there is no coercion, but rather attempted coercion. (For a list of them, see Anderson [2011: §2.1.2].) I can happily accept this claim; it does not cast doubt on the point I am about to make regarding the wrongfulness of coercion.

²⁶Of course, the ψ -ing itself might be coercion, but it is not coercion-by-threat, which I stipulated in the introduction is my only concern here.

action on the part of the announcer: conditionally threatening the target. Aside from conditionally threatening, the announcer can be entirely passive. So if coercion is a temporally extended action, and it need only involve conditional threatening, then the conditional threatening must itself be a temporally extended action. In other words, it must be that while the target φ s the announcer continues to press his threat in virtue of not withdrawing it. To me, this seems sensible; for one thing it would make sense of our inclination to say that if the target gives in and φ s, he φ s ‘under threat’. So even in the case in which the target gives in, there is only one action on the part of the announcer: the conditional threatening.

We may conclude from this investigation that wrongful coercion always involves the issuing of a conditional threat to do wrong and never involves anything more than that. So it turns out that an explanation of P1 is an explanation of why coercion is wrong when it's wrong.

I contend that the explanation of P1 that I gave in §3.2 is the best one available. First, it relies only on an ordinary moral principle (premise 1 of my argument)²⁷ and everyday observations about human psychology (which constitute the case for premise 2 of my argument). Furthermore, the other posited explanations have serious flaws (§3.1)—most importantly that they put pressure on us to reject P2 and P3. What remains to be shown is that my explanation lacks those same flaws.

4. Explaining P2: Unconditionally Threatening to Do Wrong Is, *Per Se*, Wrong

Very little has been written on the morality of unconditional threatening, and what has been written is less than satisfying. Anderson [2011: §3.2] claims that ‘making a conditional threat is often a more troublesome and ethically problematic act than making a plain, *unconditional* threat’. This, I think, runs contrary to common sense. As far as I can tell, the usual belief is that threatening to do wrong is wrong whether the threat is conditional or unconditional, and that any variation in the degree of wrongness arises from variation in how wrong the threatened act would be.

T. M. Scanlon [2008: 76] seems to agree that unconditional threats (or what he calls ‘nonincentive threats’) are less morally objectionable than conditional threats (‘incentive threats’). As to why unconditional threats should be objectionable at all, here is what Scanlon has to say [loc. cit.]:

When nonincentive threats are wrong, this impermissibility is mostly explained by the impermissibility of inflicting the penalty. . . . The impermissibility of a nonincentive threat can be explained on the ground that if it is wrong to do something to a person, then it is wrong to give him good grounds for fearing that you are going to do it.

²⁷Actually, it relies on another ordinary moral principle: that it is wrong to lie. This takes care of the case of bluffing threats.

Why would it be wrong to give a person good grounds for fearing that you will wrong him? Presumably because fear is an unpleasant mental state and so giving someone good grounds for fear is likely to cause harm to that person.

Scanlon is right, I think, to postulate that issuing an unconditional threat can be wrong in virtue of how it gives the target good grounds for fear. Yet Scanlon has not succeeded in explaining P2, because one can threaten to wrong a person without giving that person good grounds for fear. Perhaps the threatened person is unusually stoic. Why, then, should it be wrong to issue the threat? Of course, Scanlon might respond that what constitutes good grounds for a fear-response does not vary based on who the person is. But in the morally relevant sense, it does. Analogously, what constitutes teasing somebody too much, morally speaking, depends on just how thick-skinned the particular person is, not how thick-skinned most people are.

If we accept what I said earlier, that the Non-Adoption Account is true and that the only difference between conditional and unconditional threats is the conditionality (or lack thereof) of the announced intention, then we can say that unconditional threats to do wrong are objectionable for the same reason as conditional threats to do wrong: the issuing of a bluffing unconditional threat is a lie, and the issuing of a non-bluffing unconditional threat to do wrong motivates the announcer to do wrong.

5. Explaining P3: Neither Conditional Warnings about One's Own Future Wrongdoing Nor Conditional Offers to Do Wrong Are, *Per Se*, Wrong

5.1 The Inadequacy of the Extant Explanations of P3

As I said at the outset, threats, conditional offers, and conditional warnings about one's own intentions are all intention-announcements. Consequently, one might worry about being able to identify a moral difference among them. Scanlon, in fact, uses this worry to motivate his discussion of the impermissibility of threats. Unfortunately, the account that he settles on does not manage to explain why there should be a difference in permissibility between threats, on the one hand, and conditional warnings and conditional offers, on the other.

Scanlon's account has two parts; corresponding to the two objections he levels against threatening. The second objection [ibid.: 80–7] focuses on the wrongness of adopting and acting upon certain intentions. Consequently it fails to address the central conundrum about threats identified earlier: it tells us nothing about why the announcement itself might be objectionable. The first objection, however, focuses on the announcement itself, and is therefore worth a closer look.

Here, Scanlon argues that the announcement can change the meaning of some of the target's possible future actions in a way that the target has good reasons to object to. Scanlon offers the example of the employer who threatens to fire his female employee unless she has sex with him. Whereas before the announcement the employee faced the simple choice of having sex

or not having sex with her boss, she now has the choice of not having sex or having *coerced* sex with her boss. The boss's issuing of the threat has changed the meaning of the employee's possible future decision to have sex with him. More generally, when one threatens, 'If you don't ϕ I will ψ ', one turns the target's option of ϕ -ing into the option of being pushed around [ibid.: 75–80].

This account, unfortunately, doesn't do enough to explain what's distinctive about threats. As Scanlon acknowledges, sometimes the issuing of a conditional offer alters the meaning of the recipient's options in ways that the recipient has reason to object to. This happens, for instance, when one is offered money to engage in an activity that one would have done out of charity [ibid.: 79–80]. Moreover, Scanlon fails to notice that his account doesn't manage to fully distinguish threats from conditional warnings. Just as one can be pushed around by a conditional threat, so can one be pushed around by a conditional warning. Some taunts have this effect.

5.2 A New Explanation of P3

If we accept my argument in §2.5 about the difference between threats and conditional warnings, it becomes easy to understand how conditional warnings about one's own future wrongdoing escape the moral objection to which threats to do wrong are vulnerable. Premise 2 in my moral argument against non-bluffing conditional threats to do wrong says that the issuing of such a threat motivates the announcer to adopt the announced intention. This is not true of conditional warnings. When one issues a conditional warning about one's own future wrongdoing, one already has the wrongful conditional intention.²⁸ Thus, conditional warnings escape the moral objection.²⁹

To explain why conditional offers to do wrong escape the moral objection to which some non-bluffing conditional threats are subject, we need only recall the point I made earlier: offering, like warning, does not motivate the announcer to adopt his announced intention. As explained in §2.4, in making the announcement the offerer puts the matter of his intention-adoption in the target's hands. The target can prevent the announcer from adopting the announced conditional intention by simply not accepting the offer. Thus, the only thing that can motivate the announcer to adopt the

²⁸It has been suggested to me that it is bizarre that an important moral difference, the difference between its being permissible or impermissible to announce an intention to do wrong, could turn on a trivial empirical difference: whether the adoption of the announced intention preceded the announcement of it, even by an instant. The point I have been at pains to make, however, is that the given empirical difference is not empirically trivial at all. That slight difference in timing makes all the difference as to whether one's intention-announcement can significantly increase the likelihood of one's adopting that intention.

²⁹When discussing threats, I said that people are motivated to keep their word, and therefore by making such an intention-announcement the announcer thereby gains motivation to adopt his announced intention. But if this claim is true, then it holds for conditional warnings as well, although in the case of conditional warnings what one gains is motivation to *maintain* the intention to do what one announces one will do. The important question, morally speaking, is: when that happens does one become more likely to do what one already intends to do? The answer is no. If an announcer has a malevolent conditional intention to ψ then he becomes *less* likely to ψ when he announces this conditional intention to his target. This is because the making of the announcement informs the target of something he (the target) can do to deter the announcer's ψ -ing. (Thanks to Alan Wertheimer for pointing this out to me.)

announced conditional intention, when the announcement is a conditional offer, is *the target's own action* of accepting the conditional offer.

5.3 An Objection

One might worry that in this section I have gone too easy on conditional warnings and conditional offers and that, because of this, I have left us no way of objecting to wrongful behaviour that involves only a conditional warning or a conditional offer. Extortion, at least some instances of it, fits this description. To illustrate, consider the following case. X and Y strike an agreement whereby X will purchase Y's car for \$5000. However, after receiving X's \$5000 deposit, Y announces that the price has just gone up. Now it will take an additional \$1000 to secure the title to the car. Y assures X that if he refuses then his initial \$5000 will not be refunded.

Let's assume that before announcing that the price had gone up Y had already adopted the conditional intention to keep X's initial \$5000 if X does not hand over another \$1000. This assumption means that on my account Y's announcement of this conditional intention is a warning, not a threat. Yet it seems that he has acted wrongfully. How can we account for that, given what I have said about conditional warnings?³⁰ What we can (and should) say is that Y has done wrong in virtue of having adopted a conditional intention to keep X's initial \$5000 if X does not hand over another \$1000. Adopting this intention is wrongful because carrying out the intention would involve breaking a promise.

At the beginning of §3 I posed the question, 'How can the communication of a malevolent intention add anything to the wrongness of adopting it and following through on it?' Given what I have argued here, the following answer should appear plausible: it can't. The actual adoption of a malevolent intention is often subject to moral objection, whereas the announcement of it is objectionable only when it motivates the announcer to adopt it. So in cases in which the announced intention has already been adopted (i.e., cases of conditional warning), any moral objection is going to have to focus on the act of intention-adoption, as in extortion. In cases in which the announced intention has *not* already been adopted, any moral objection is going to have to focus on the act of intention-announcement. If announcing the intention is a lie (i.e., in cases of bluffing threats) or gives the announcer motivation to adopt the announced intention (i.e., in cases of non-bluffing threats), that objection may go through. If the intention-announcement is inert (i.e., in cases of conditional offering), the objection will not go through.

6. Conclusion

My goal here has been to show that once we gain a proper understanding of what a conditional threat is, the moral orthodoxy about coercion becomes,

³⁰I thank Oren Bar-Gill for alerting me to this objection.

as a whole, much easier to make sense of. When an announcer issues a conditional threat it is always the case, according to the theory I advanced here, that he did not possess the announced intention before making his announcement. With this account in hand, I tried to elucidate three aspects of the common-sense morality about threats and other intention-announcements. I turned first to the question of the moral permissibility of conditional threats, focusing on the cases in which the announced conditional intention is an intention to wrong the target. I argued that in those cases the issuing of the threat (when it is not a bluff) is itself wrongful because of the effect that action has on the announcer—specifically, motivating him to adopt the announced wrongful conditional intention. By focusing on the effect on the announcer, as opposed to the effect on the target's mental states or circumstances, I accounted for why the announcement of the intention, and not just the adoption of it, is morally objectionable. I was also able to make clear why unconditional threats are wrongful in the same way (i.e., because they have the same effect on the announcer) and conditional warnings of one's own future wrongdoing and conditional offers to do wrong are not (i.e., they don't). Finally, by showing that coercion, when it's wrong, is nothing more than the issuing of a conditional threat to do wrong, I was able to demonstrate that my explanation of the wrongness of conditionally threatening to do wrong is also an explanation of why coercion is wrong when it's wrong.

This may seem to amount to quite a demotion in moral importance for wrongful coercion. If I am right, the wrongness of coercion is not to be explained by reference to our most lofty or sacred moral values, such as freedom and autonomy, but rather by reference to the wrongness of motivating oneself to do wrong. I suspect it will seem to many that this explanation cannot account for the moral egregiousness of wrongful coercion. It seems that coercion, when wrong, is a particularly detestable kind of wrongdoing. For my part, I accept this appearance as veridical, and I would suggest that in order to accommodate it we should reconsider whether the injunction against motivating oneself to do wrong shouldn't merit a place in the pantheon of high moral ideals.³¹

New York University

Received: July 2011

Revised: October 2011

References

- Anderson, Scott 2011. Coercion, *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), ed. Edward N. Zalta, URL = <<http://plato.stanford.edu/entries/coercion/>>.
- Anscombe, G. E. M. 1963 (1966). *Intention*, second edition, Ithaca: Cornell University Press.
- Berman, Mitchell 2002. The Normative Functions of Coercion Claims, *Legal Theory* 8/1: 45–89.
- Feinberg, Joel 1986. *Harm to Self*, Oxford: Oxford University Press.
- Fowler Mark 1982. Coercion and Practical Reason, *Social Theory and Practice* 8/3: 329–55.

³¹I am extraordinarily grateful to Collin O'Neil for helping me, over the course of countless conversations, to work through the puzzles I grapple with here. This paper was presented before audiences at Fordham University and Bowling Green State University, and I would like to thank those audiences for their feedback. Finally, I am indebted to Oren Bar-Gill, Dale Jamieson, Matthew Liao, Japa Pallikkathayil, Alan Wertheimer and two anonymous referees for this journal for reading and commenting on previous drafts of the manuscript.

- Frankfurt, Harry 1988. *The Importance of What We Care About*, Cambridge: Cambridge University Press.
- Gauthier David 1994. Assure and Threaten, *Ethics* 104/4: 690–721.
- Greenawalt, Kent 1983–4. Criminal Coercion and Freedom of Speech, *Northwestern University Law Review* 78: 1081–124.
- Greenawalt, Kent 1989. *Speech, Crime, and the Uses of Language*, New York: Oxford University Press.
- Haksar, Vinit 1976. Coercive Proposals [Rawls and Gandhi], *Political Theory* 4/1: 65–79.
- Hampshire, Stuart and H. L. A. Hart 1958. Decision, Intention and Certainty, *Mind* 67/265: 1–12.
- Kavka, Gregory 1983. The Toxin Puzzle, *Analysis* 43/1: 33–6.
- Lamond, Grant 1996. Coercion, Threats, and the Puzzle of Blackmail, in *Harm and Culpability*, ed. A. P. Simester and A. T. H. Smith, Oxford: Clarendon Press: 215–38.
- Lamond, Grant 2002. The Coerciveness of Law, *Oxford Journal of Legal Studies* 20/1: 39–62.
- Lamond, Grant 2010. Coercion, in *A Companion to Philosophy of Law and Legal Theory*, second edition, ed. Dennis Patterson, Malden: Wiley-Blackwell: 642–53.
- Nozick, Robert 1969. Coercion, in *Philosophy, Science, and Method: Essays in Honor of Ernest Nagel*, ed. Sidney Morgenbesser, Patrick Suppes, and Morton White, New York: St. Martin's Press: 440–72.
- Pallikkathayil, Japa 2011. The Possibility of Choice: Three Accounts of the Problem with Coercion, *Philosophers' Imprint*. 11/16: 1–20.
- Parfit, Derek 2001. Rationality and Reasons, in *Exploring Practical Philosophy: From Action to Values*, ed. Dan Egonsson, Björn Petersson, Jonas Josefsson, and Toni Ronnow-Rasmussen, Aldershot: Ashgate Publishing: 17–39.
- Pink, Thomas 2008. Intentions and Two Models of Human Action, in *Reasons and Intentions*, ed. Bruno Vervbeek, Burlington: Ashgate Publishing Company: 153–80.
- Raz, Joseph 1975. Reasons for Action, Decisions and Norms, *Mind* 84/336: 481–99.
- Raz, Joseph 1986. *The Morality of Freedom*, Oxford: Clarendon Press.
- Scanlon, T. M. 2008. *Moral Dimensions: Permissibility, Meaning, Blame*, Cambridge, MA: Belknap Press.
- Wertheimer, Alan 1987. *Coercion*, Princeton: Princeton University Press.
- Zimmerman, David 1981. Coercive Wage Offers, *Philosophy and Public Affairs* 10/2: 121–45.