

4

Coercion

Consider the following as a paradigm case of coercion:

Terry is targeted by an armed thief. The thief tells Terry to hand over his wallet. Terry refuses to cooperate. The thief then tells Terry that he will shoot him unless he cooperates. Terry agrees to hand over his wallet.

In accordance with the standard account of autonomy in bioethics, I assume that it is uncontroversial to claim that Terry is not autonomous with respect to his decision to hand over his wallet. Yet, despite the strength of our intuitions in this regard, coercion raises a number of puzzles, two of which in particular shall be my focus in this chapter. First, in what sense does coercion undermine voluntariness, and second, can offers as well as threats be coercive?

I shall begin this chapter by explaining my motivation for exploring these particular questions in more detail, before going on to explore two prominent streams in the literature on coercion that seeks to address these questions. I shall go on to argue for an account of coercion that accommodates the possibility of some coercive offers, and suggest that the answers to the two questions are interrelated. That is, I shall suggest that the way in which we should distinguish threats from offers (in responding to the second question) should draw on considerations that are closely linked to one's understanding of the way in which coercion undermines the interpersonal sense of voluntariness that I introduced in the previous chapter.

First though, since coercion is a term that is subject to a number of different interpretations in bioethical contexts, it is prudent to begin by delimiting the understanding that I shall invoke in my discussion. Although coercion is most often associated with the use of threats, some theorists adopt a broader understanding of the concept, according to which physically forcing an agent to do something (or to have something done to them) against their will can also amount to coercion.¹ For example, the term 'coerced treatment' is sometimes used in psychiatry to refer to treatment that has been forced in this sense. However, this use of the term conflates an important moral distinction between interventions in which an agent is forced to act, and interventions in which an agent is *acted upon*, potentially in the absence of their assent and even despite their dissent.²

¹ For instance, Wood claims that an agent can be said to be coerced when they do not choose to perform some action. Wood, 'Coercion, Manipulation, Exploitation', 21.

² See Bayles, 'A Concept of Coercion'; Lamond, 'Coercion, Threats, and the Puzzle of Blackmail'.

In order to avoid this conflation, I shall understand coercion to refer to a form of interference that necessarily involves the use of conditional proposals that render certain options ineligible for rational choice.³ However, unlike the standard account of autonomy's approach to understanding coercion I shall *not* assume at the outset that the kinds of conditional proposals involved in coercion must, *by necessity*, be threats.⁴

1. Two Questions Facing an Adequate Account of Coercion in Bioethics

With this in mind, I can now turn to the two broadly interrelated questions with which I shall be concerned in this chapter. The first is why we should understand coercion to amount to a form of controlling interference of the sort that undermines voluntariness *at all*.

Coercion represents something of a problem for authenticity-based approaches to decisional autonomy because, although we want to be able to say that victims of coercion do not decide autonomously, these individuals will often meet the authenticity conditions set out by the accounts I have so far considered in this book. As Irving Thalberg notes, most persons who are subjected to coercion '... would, at the time and later, give second-order endorsement to their cautious [and compliant] motives'.⁵ With regards to the theory that I have defended in the previous chapters, the motivating desire that an individual forms following a coercive threat could very plausibly be one that they rationally endorse with a preference that coheres with their character system. For instance, in the above example, Terry may endorse his desire to hand over his wallet, because the content of his motivating desire can be understood to include the outcome of staying alive, and he understands this to be a more valuable outcome than safeguarding his money. We may also note that victims of coercion also normally have adequate understanding of their situation.

Explaining how and why coercion undermines the voluntariness of the victim's decision is a considerable challenge for any authenticity-based conception of autonomy. However, in view of my claims in the previous chapter, it is perhaps a particularly acute challenge for the view that I am developing here. Recall that in the previous chapter, I claimed that deception and manipulation are forms of controlling influence that can undermine decisional autonomy without being intentional. This raises an important question for my theory because it appears that the same is not true of coercion; coercion must be intentional if it is to undermine decisional autonomy of the sort that undergirds the doctrine of informed consent. To see why, consider the following two cases:

Alan is told by his doctor that if he refuses to start taking painkillers that will cause him to suffer from alopecia as a side-effect, he will start to experience significant pain as a result of an underlying medical condition. Although Alan agrees to take the pills, he would strongly prefer to be in a situation in which his

³ Feinberg, *The Moral Limits of the Criminal Law*, 191.

⁴ Beauchamp and Childress, *Principles of Biomedical Ethics*, 138.

⁵ Thalberg, 'Hierarchical Analyses of Unfree Action', 126.

not experiencing pain was not also attended by the consequence of losing his hair.

Bernie is in fine health, but someone threatens to cause him similarly significant pain if he does not agree to take a drug that causes alopecia. Although Bernie agrees to take the pills, he would strongly prefer to be in a situation in which he does not experience pain, and where his not experiencing pain is not also attended by the consequence of losing his hair.

It seems that Alan can be autonomous with respect to his decision to take the drug, despite the limited choices available to him: Valid consent must be possible in the light of limited treatment options.⁶ On the other hand, it is highly counter-intuitive to claim that Bernie is autonomous with respect to his decision to take the drug; like Terry, it seems that Bernie has plausibly been subjected to controlling influence of the sort that might be said to undermine the voluntariness of his decision. However, the only difference between the two cases is that Bernie's available options have been influenced by an intentional agent, whilst Alan's have been engineered by forces of hazard. The question is why should this matter for coercion, but not for deception or manipulation?

Aristotle recognized the nature of the theoretical challenge that coercion raises in his discussion of voluntariness, when he suggests that our actions following such threats are 'mixed' with regards to voluntariness. He writes:

If a tyrant, for example, had one's parents and children in his power and ordered one to do something shameful, on the condition that one's doing it would save them, while one's not doing it would result in their death – there is some dispute about whether they are involuntary or voluntary. The same sort of thing happens also in the case of people throwing cargo overboard in storms at sea. Without qualification, no one jettisons cargo voluntarily; but for his own safety and that of others any sensible person will do it. Such actions, then, are mixed, though they seem more like voluntary ones, because at the time they are done they are worthy of choice, and the end of an action depends on the circumstances. So both voluntariness and involuntariness are to be ascribed at the time of the action.⁷

Despite claiming that these kinds of action are 'mixed' with regards to voluntariness, Aristotle himself eventually concludes that coerced actions should in fact be understood to be voluntary, in so far as the moving principle of the coerced agent's action lies in the agent herself.⁸ Notice though that Aristotle here does not draw a moral distinction between coercion manufactured by an intentional agent (i.e. the

⁶ A number of authors have raised this sort of counterexample to accounts of coercion discussed in the bioethics literature. See Wertheimer, 'Voluntary Consent'; Richards, *The Ethics of Transplants*; Pugh, 'Coercion and the Neurocorrective Offer'.

⁷ Aristotle, *Nicomachean Ethics*, 1110a 4–12.

⁸ *Ibid.*, 11110b 5–7. Notably, in her discussion, Meyer suggests that Aristotle believes such acts are mixed with regards to voluntariness because they are voluntary with respect to the sense pertaining to ignorance, but involuntary with respect to the sense pertaining to choice. See Meyer, 'Aristotle on the Voluntary', 141–2. However, this interpretation seems to miss out on the nuance of Aristotle's discussion quoted above, since the mixed nature of voluntaries here makes no reference to understanding; instead, the belief that coerced acts could be voluntary is grounded by the claim that no one would perform the act in question without the qualifications provided by the context of the choice.

tyrant example) and coercion manufactured by forces of hazard (i.e. the cargo example). Similarly, one could adopt an Aristotelian approach and simply claim that both Alan and Bernie make voluntary decisions in the above example, in so far as the moving principle of action in both cases is in some sense ‘internal’ to the agent. However, such an interpretation would run contrary to the widely held view in bioethics that coercion does undermine the voluntariness of the victim’s decision to comply.⁹ As such, I suggest that we instead need an account of why this is so, which is compatible both with an authenticity-based understanding of autonomy, and the claim that neither deception nor manipulation need be intentional to undermine decisional autonomy.

The second question I shall consider in this chapter is whether offers as well as threats can be coercive. For some, this question might seem misguided—it has been claimed that offers are just categorically different from threats for reasons that I shall consider below. Moreover, the view that only threats can be coercive is commonly assumed in bioethics; it is endorsed by advocates of the standard view in their discussion of coercion,¹⁰ and it is also enshrined in the Belmont Report, which states that ‘coercion occurs when an overt threat of harm is intentionally presented by one person to another to obtain compliance’.¹¹ However, although this position has been highly influential we should be careful not to simply assume that it is true, particularly when invoking the concept of coercion in the context of contemporary bioethics. The reason that we should take particular care in this context is that coercion-related concerns are commonly raised not with respect to the use of explicit threats, but rather by the use of incentives.

To further illustrate this, consider the following three cases in which bioethicists have argued that offers can be coercive. To be clear, the charge of coercion is not the only moral criticism that has been raised in the debates surrounding the issues outlined below. However, this moral criticism is particularly significant. Given the salient value that we attribute to personal autonomy in contemporary bioethics, establishing that a practice amounts to controlling interference that invalidates consent is a particularly powerful moral criticism.

(i) *Paying Research Subjects*

In 2006, eight healthy subjects were offered £2000 to participate in a phase I trial of a novel monoclonal antibody agent, TGN1412. As part of the informed consent form, participants were told that they were free to leave the trial at any time without giving a reason. However, they were also told that if they chose to withdraw and exercise their right not to give a reason, or if they were required to leave the study for non-compliance, then they would forfeit their entitlement to payment.¹² Having

⁹ Interestingly, drawing on some comments in Arthur Caplan’s *Am I My Brother’s Keeper?* Wertheimer notes (but does not endorse) the possibility that one could adopt the complete opposite view and claim that examples like Alan and Bernie above suggests that terminally ill patients cannot provide valid consent to treatment. Wertheimer, ‘Voluntary Consent’.

¹⁰ Beauchamp and Childress, *Principles of Biomedical Ethics*, 138. See also Nelson et al., ‘The Concept of Voluntary Consent’, 7.

¹¹ The Belmont Report.

¹² ‘TGN1412 Trial Consent Form’.

previously tested the drug in animals, researchers gave six of the subjects in this first in-human trial 1/500th of the highest dose used in animal testing (the remaining subjects received a placebo). However, the six who received a dose of the drug quickly developed serious adverse reactions to the experimental agent, as result of a cytokine storm.¹³

Although regulatory reports on the trial suggested that it fulfilled all ethical requirements for clinical research (including the criteria of social value, scientific validity, favourable risk/benefit ration, informed consent, independent review, and respect for participants),¹⁴ some bioethicists raised concerns about the quality of the participants' consent, with one commentator arguing that the language used in the consent process was 'very coercive'.¹⁵ This sparked a number of commentaries on the trial discussing whether the payment offered for participation in the trial amounted to coercion.¹⁶

(ii) Reduced Sentences for Sexual Offenders Who Agree to Undergo Chemical Castration

Chemical castration refers to the use of an anti-libidinal agent to significantly reduce the recipient's sex drive. Although the anti-libidinal effects of the drugs that are typically used for this purpose are generally reversible, there are concerns about the long-term safety of chemical castration.¹⁷ Despite both this and the scarcity of robust data about the effectiveness of chemical castration in preventing recidivism,¹⁸ the procedure is sometimes performed on convicted sex offenders. Whilst chemical castration is compulsory for sex offenders in some jurisdictions, in others, sex offenders who would otherwise face a long prison sentence may be offered a significantly reduced sentence on the condition that they agree to undergo chemical castration.¹⁹ Supporters of this 'offer' model argue that this benefits the offenders by giving them the chance to avoid a long prison sentence, whilst also respecting their autonomy by leaving the choice in the offender's hands. However, opponents have objected that the offer is inherently coercive. Here is a typical example:

The convicted rapist is faced with two options – a lengthy prison sentence or even death on the one hand and ... castration on the other. Freedom of choice is impossible because the convict's

¹³ Suntharalingam et al., 'Cytokine Storm in a Phase 1 Trial of the Anti-CD28 Monoclonal Antibody TGN1412'.

¹⁴ For an overview of such principles, see World Medical Association Declaration of Helsinki.

¹⁵ Evans, 'Paraxel Misled Subjects Sickened in London Study, Ethicists Say'.

¹⁶ Wertheimer and Miller, 'Payment for Research Participation'; Emanuel and Miller, 'Money and Distorted Ethical Judgments about Research'; Schonfeld et al., 'Money Matters'.

¹⁷ Garcia and Thibaut, 'Current Concepts in the Pharmacotherapy of Paraphilias'.

¹⁸ Thibaut et al., 'The World Federation of Societies of Biological Psychiatry (WFSBP) Guidelines for the Biological Treatment of Paraphilias'.

¹⁹ Forsberg and Douglas, 'Anti-Libidinal Interventions in Sex Offenders'; Douglas et al., 'Coercion, Incarceration, and Chemical Castration'. For discussion of the neurocorrective offer and coercion, see Pugh, 'Coercion and the Neurocorrective Offer'.

loss of liberty constitutes a deprivation of such a magnitude that he cannot choose freely and voluntarily, but he is forced to give consent to an alternative he would not otherwise have chosen.²⁰

(iii) *Markets for Organs*

In light of the global shortage of kidneys for transplantation, a number of bioethicists have suggested that we ought to permit a regulated market in which individuals are free to sell one of their kidneys.²¹ Effectively, on such a model, individuals would be offered a sum of money in return for their kidney. A common objection to such a market is that it would coerce the poorest members of the global community into selling their kidney. Michael Jaycox captures this objection as follows:

From the perspective of the poorest and most vulnerable persons in this world, the market of human organs is not something in which all persons choose to participate as equally free and self-determining individuals. For the poorest and most vulnerable members of the world community effectively have no or little choice but to participate in this market as vendors.²²

In order to ascertain whether the charge of coercion may appropriately be raised in these three examples, we need to have a sound understanding of the nature of coercion, how it undermines voluntariness, and how offers differ from threats. In short, we need to provide an answer to the two questions that I have raised in this section; however, as I shall now explain, our responses to these questions are likely to be importantly interrelated, since the way in which one might distinguish threats from offers (in responding to the second question) might plausibly draw on considerations that are closely linked to one's understanding of how coercion undermines voluntariness.

With these framing remarks in mind, I shall proceed as follows. In the next section, I shall introduce a highly influential 'content-based' view of the distinction between threats and offers, and the normative and non-normative accounts of coercion that are grounded by this view. Having outlined problems facing normative accounts, I shall, in section 3 consider non-normative accounts in more detail. Although I shall suggest that these accounts get closer to the truth about coercion, I shall argue that they offer inadequate accounts of why coercion undermines voluntariness. The reason for this is that they do not adequately emphasize the importance of interpersonal voluntariness in this context. In section 4, I shall argue that the significance of this sense of voluntariness can be better captured if we adopt a structural rather than content-based understanding of the distinction between threats and offers. This understanding accommodates the possibility of some coercive offers. In the final section, I shall consider the practical implications of my approach for the cases delineated above.

²⁰ Green, 'Depo-Provera, Castration, and the Probation of Rape Offenders'. More recently, Wood has defended an account of coercion according to which an agent is coerced to do something when they have no acceptable alternative. Wood, 'Coercion, Manipulation, Exploitation'.

²¹ Richards, *The Ethics of Transplants*; Taylor, *Stakes and Kidneys*; Wilkinson, *Bodies for Sale*.

²² Jaycox, 'Coercion, Autonomy, and the Preferential Option for the Poor in the Ethics of Organ Transplantation'. See also Rippon, 'Imposing Options on People in Poverty'; Annas, 'Life, Liberty, and the Pursuit of Organ Sales'.

2. The Content-Based View of Threats: Normative and Non-Normative Accounts of Coercion

On one prominent approach to coercion, one that is endorsed by the standard account of autonomy in bioethics, it is claimed that coercion necessarily involves the use of threats. Of course, the appeal of this understanding of coercion depends in large part on our understanding of threats. Although I shall refine this terminology over the course of this section, we may say, at this stage roughly, that a proposal constitutes a threat if it announces a conditional intention to make the recipient worse off if she does not perform some action that the coercer wants her to perform.

The first thing to acknowledge with regards to the ‘threat-based’ understanding of coercion is that even if *only* threats can be coercive, clearly not *all* threats are. For instance, suppose the thief from my previous example had instead threatened to verbally insult Terry if he refused to hand over his wallet. It does not seem plausible to describe this as an instance of coercion. Accordingly, advocates of the threat-based understanding of coercion must offer an account of why some threats are coercive and some are not.

I shall not be concerned primarily with such conditions here; my interest in this section is rather to explore the way in which the threat-based understanding of coercion seeks to distinguish threats and offers, and the implications of this view for our understanding of the implications for voluntariness in each case. Nonetheless it will be useful to have a broad understanding of some commonly accepted necessary conditions of coercive threats for my discussion below. Broadly, we may say that in order for P’s threat to have coerced Q into doing A, it must, *inter alia*, be the case that:

- (1) Prior to the threat being made, Q believes he has a decisive reason to not A.
 - (2) Q has sufficient reason to believe that P’s threat is credible.
 - (3) The fact that the non-performance of A will lead to the consequence that P threatens to bring about if Q does not A is the operative reason in Q’s decision to A; the consequences render the non-performance of A for Q ‘ineligible’ for rational choice.²³
- And
- (4) Q does A.

These conditions essentially state that for P’s threat to be coercive, it must not only be successful in getting Q to A, it must also be the case that Q came to be motivated to do A because she believed P’s threat, and wanted to avoid the consequences that P has credibly threatened to bring about.²⁴

I do not mean to claim that these conditions are *sufficient* for a threat to be coercive; one might wish to add further conditions in order to distinguish coercive

²³ I have adopted the terminology of eligibility for rational choice from Feinberg’s useful discussion. See Feinberg, *The Moral Limits of the Criminal Law*, 192. However, this terminology (and this condition generally) is compatible with the various accounts of coercion that I shall consider here.

²⁴ Feinberg, *The Moral Limits of the Criminal Law*, 198; Raz, *The Morality of Freedom*, 149; Nozick, ‘Coercion’.

threats from warnings, for example.²⁵ Moreover, as I suggested above, the example of Alan and Bernie might be taken to suggest that we should incorporate a condition of intentionality here, according to which, the coercing party must intend to coerce their victim if their threat is to be truly coercive.²⁶ Nonetheless, such a condition has been contested, as I shall discuss. In the interests of brevity, and to focus my discussion, I shall simply accept the above four conditions as broadly accepted necessary conditions of coercive threats without further discussion.

With this in mind, why should we accept the claim that coercion necessarily involves the use of threats? Arguably, our intuitions may speak in favour of this view. In addition to the fact that paradigmatic cases of coercion (such as Terry above) pre-theoretically seem to involve threats, we also commonly get people to do things that they would otherwise not do by offering them inducements, and it seems clear that this does not usually involve coercion. If Terry had been offered a million pounds to hand over his wallet to an eccentric collector (rather than being threatened with death), it perhaps seems more plausible that he could have been autonomous with respect to his decision to hand over his wallet (assuming it contained far less than a million pounds).

However, one problem with relying on these sorts of intuitions to ground the threat-based understanding of autonomy is that it is not at all clear how we should philosophically distinguish threats and offers. As Scott Anderson points out in his introduction to the concept of coercion:

... offers may also be made with the same general intention as coercive threats: that is, to make some actions more attractive, others less so.²⁷

Anderson goes on to point out that both threats and offers share the same basic structure; in both cases, the proposing agent, P, claims that she will bring about consequence C if and only if (*iff*) the proposed-to agent, Q, does some action A.²⁸ I shall suggest below that Anderson's structural analysis here is subtly incorrect. However, the claim that threats and offers are structurally similar has led both defenders and opponents of the threat-based understanding of coercion to adopt what I shall call a 'content-based' view of the difference between the two types of proposal.

The content-based view claims that threats can be distinguished from offers by appealing to the difference in the nature of the consequences that the proposer announces a conditional intention to bring about in each case. Advocating such a view in his seminal account of coercion, Robert Nozick claims that threats announce a conditional intention to bring about a consequence that would make the recipient *worse off* in comparison to the baseline of the 'normally expected course of events'.²⁹ Recall the original example of Terry. On this account, the thief's proposal amounts to a threat because it announces a conditional intention to bring about a consequence that would make Terry worse off than he might normally expect to be if he refrained from handing over his wallet to a stranger. In contrast, on this approach, offers announce a conditional intention to bring about a consequence that would make the

²⁵ Nozick, 'Coercion', 453–8.

²⁶ *Ibid.*

²⁷ Anderson, 'Coercion'.

²⁸ *Ibid.*

²⁹ Nozick, 'Coercion', 447.

recipient *better off* in comparison to the baseline of the ‘normally expected course of events’.

The ambiguity of the concept of ‘the normally expected course of events’ has led to two different accounts of coercion grounded by this way of distinguishing threats and offers, one non-normative, one normative. The ambiguity of the concept is attributable to the fact that there are different norms of expectability that might govern our understanding of the baseline ‘normal expected course of events’.³⁰ More specifically, the ‘normal’ expected course of events can be understood in a merely descriptive sense, to refer to what one might reasonably expect to happen, given one’s prior knowledge concerning descriptive facts about the world. However, it is also possible to understand the baseline in a normatively laden sense, to refer to how the world *ought* to be. To illustrate this point, consider the following example from Nozick’s discussion:

Suppose that usually a slave owner beats his slave each morning, for no reason connected with the slave’s behaviour. Today he says to his slave, ‘Tomorrow I will not beat you if and only if you now do A.’³¹

In relation to a descriptive understanding of the normal course of events, the slave-owner’s proposal qualifies as an offer, since the slave-owner proposes to make the slave better off than he could normally expect to be (given his previous experience of being a slave of this owner) if he performs the task that the slave-owner asks him to perform.

This raises something of a problem, because it seems highly plausible to claim that the slave is nonetheless coerced in this case. Accordingly, theorists who claim that *only* threats can be coercive must argue that it is possible to recast the slave-owner’s proposal as a threat if they are to accommodate this intuition. One way in which it is possible to do this is to consider the proposal against a normative baseline that incorporates what the agent might *morally* expect:

Morally Expected Course of Events: The baseline comparison course of events that incorporates what is minimally morally required of others in their actions towards the recipient in the pre-proposal situation.

In comparison to this normative baseline, the slave-owner’s proposal can be understood as a threat, since the proposal does not make the slave better off in comparison to the course of events in which others meet their moral obligations towards the slave. After all, in this course of events, the agent would not be a slave who is beaten every day, but rather a free man.³² Other normative accounts of coercion parse the normative baseline in terms of circumstances in which the individual’s *rights* are respected, and in which others meet their obligations towards the individual.³³ Rights-based normative accounts also claim that we should *always* invoke a normative baseline when distinguishing coercive proposals from non-coercive proposals. In contrast, in his original theory, Nozick suggests that when the normative and

³⁰ Feinberg, *The Moral Limits of the Criminal Law*, 219–27.

³² *Ibid.*

³³ Wertheimer, *Coercion*, 217–21.

³¹ Nozick, ‘Coercion’, 450.

non-normative baselines diverge, it should be up to the agent which baseline should be invoked.³⁴

This normative approach is also understood to provide a theoretical foundation for an account of why coercion undermines voluntariness. Due to the different content of threats and offers on this approach, Nozick suggests that a rational agent would be willing to move from their pre-proposal 'baseline' situation to the post-proposal choice situation. With respect to offers, they would be willing to do this because the nature of the consequence that the third party announces an intention to bring about is such that it would make the recipient better off than they would have been in comparison to the baseline of the normal expected course of events. The same cannot be said in the case of threats; after all, in the case of a threat, the recipient is in a position in which their rights will be violated if they fail to comply, or they will not be treated in the manner that morality demands.³⁵ Accordingly, the reason that coercive threats undermine voluntariness on this account is that they involve moving the recipient into a post-proposal choice situation unwillingly.

This normative account of coercion captures the Aristotelian insight that a coerced act is 'mixed' with regards to voluntariness. Even though the coerced agent may willingly comply with the coercer's demands, they have unwillingly been placed into the post-proposal choice situation. Nonetheless, there are significant problems with this understanding of coercion. Some of these problems pertain to the content-based view of the distinction between threats and offers that partly grounds the account. Naturally, these problems are also shared by other (non-normative) accounts of coercion that are also based on this distinction; I shall consider such accounts in the following sections. However, to conclude my discussion of normative accounts, I shall outline those problems that are uniquely faced by normative accounts of coercion.

First, if coercion only occurs when one agent proposes to violate another's rights, or if they fail to meet some other sort of moral obligation, it makes the definition of coercion, and more importantly its moral wrongness, parasitic on these prior moral wrongs.³⁶ Yet this is to divorce the moral wrongness of coercion from the implications that it has on the voluntariness of the target's choice. This is problematic because it seems that these implications are fundamental to our understanding of why coercion is wrong, particularly in a bioethical context. Indeed, the normative account of coercion seems to jar with our intuitions in this regard. For instance, consider the following case:

S is a justly imprisoned prisoner who would strongly prefer not to be imprisoned. The warden tells S that if he is caught attempting to escape, he will be liable to have his entertainment privileges withdrawn. Assume that S does not have any right to leave prison, and that prison authorities would not be violating S' rights if they withdrew certain entertainment privileges from him following an escape attempt. Suppose that these privileges are highly valuable for S in

³⁴ Nozick, 'Coercion', 451. Notably though, Nozick himself came to endorse a rights-based understanding of the normative baseline in later work. See Nozick, *Anarchy, State and Utopia*.

³⁵ Nozick, 'Coercion', 459–60.

³⁶ Zimmerman, 'Coercive Wage Offers', 123.

prison, and, because he believes that he cannot risk losing them, he decides not to try and escape.³⁷

On the normative account, S's deciding not to escape because of the warden's threat of withdrawing these privileges would not qualify as coerced, and we would have no reason to claim that his decision to not try and escape here was non-voluntary. After all, the warden has not threatened to violate S's rights in any way. Yet this seems implausible; given that S takes himself to have very strong reasons to escape, but even stronger reasons to avoid the threatened consequence, it is not clear why the mere fact that the latter consequence would not violate S's rights should have any bearing on the effect of these threatened consequences on the voluntariness of S's choice.

More broadly, it is not clear that the normative account can make sense of morally legitimate instances of coercion that plausibly serve to undermine voluntariness, despite the fact that they are morally legitimate. In addition to the above example, we might also note the quite common example of parents legitimately shaping their children's behaviour through the use of coercive threats. It seems that we might want to say that parents can coerce their children (perhaps by telling them to eat their vegetables or they will not get dessert) without necessarily threatening to treat them in ways that violate their rights, or failing to treat them as morality demands.

Another reason that the normative account of threats and offers is jarring is that it is somewhat unnatural to describe the slave-owner's proposal in Nozick's case as a threat rather than an offer. In section 4, I shall provide a structural explanation of the proposal in question that explains why the normative interpretation is counter-intuitive. Of course, the point that the normative account of coercion can jar with our common understanding of threats and offers is not, by itself, a knock-down objection. Philosophers commonly invoke conceptions of everyday terms that do not map neatly on to our common usage of those terms. However, this point at least speaks against the normative account, particularly if there is an understanding of coercion that incorporates a construal of threats and offers that is congruous with the common understanding of these terms, and which can accommodate morally legitimate coercion that nonetheless undermines voluntariness.

3. Non-Normative Approaches, Coercive Offers, and Interpersonal Voluntariness

As I explained in the previous section, non-normative accounts of coercion can also be grounded by a content-based understanding of the distinction between threats and offers. Unlike normative theories, non-normative accounts appeal to a *descriptive* baseline of the normal expected course of events in drawing this distinction. This allows non-normative approaches to classify the slave-owner's proposal in Nozick's example as an offer, rather than a threat.

I believe that non-normative theories reach the right conclusion in this regard, although I shall argue that they do so for the wrong reasons. Notwithstanding this point though, presuming that a plausible theory of coercion ought to claim that the

³⁷ This is a modified version of an example given by Olsaretti, 'Freedom, Force and Choice', 59.

slave-owner's proposal is coercive, then plausible non-normative accounts must deny that only threats can be coercive. They must allow for the conceptual possibility of coercive offers. Accordingly, a key challenge that such accounts face is explicating which offers should qualify as coercive. In this section, I shall delineate what we might term the 'preference-based' view of coercive offers.³⁸

Recall that on a non-normative approach, offers will announce a conditional intention to make the recipient better off than they would normally expect to be given descriptive facts about the world. According to the preference-based view, this feature of offers is significant because it means that recipients will typically prefer to receive offers. The reason for this is that they would strongly prefer to move from the normally expected course of events (in their pre-proposal situation) to the post-proposal situation (laid out in the terms of the offer).

However, an offer may nonetheless qualify as coercive on the preference-based view if two conditions are met. First, it must be the case that the recipient would *even more* strongly prefer to move from their actual pre-proposal situation to some alternative feasible pre-proposal situation. Second, the proposer must be *actively preventing* the recipient from being in this alternative feasible pre-proposal situation.³⁹ On this view, the slave-owner's offer is coercive because he can be understood as actively preventing the slave from being in a feasible alternative pre-proposal situation that he would strongly prefer to be in; namely one in which he is not a slave who receives daily beatings.

In the case of coercive offers, the preference-based view holds that an offer will be coercive if the offering party is frustrating the recipient's desire to be in some other feasible alternative pre-proposal situation. In turn, this feature of coercion is central to its prima facie moral wrongness on the preference-based view; coercion renders the victim unfree by frustrating their preference to be in some other pre-proposal situation. The preference view can further cash out the wrongness of frustrating an agent's desire in this way in both broadly utilitarian and broadly Kantian terms. On the utilitarian approach, rendering another unfree in this manner is wrong because desire frustration per se is prima facie wrong, whilst on the Kantian approach, frustrating another's desires is wrong because it fails to acknowledge the victim's full status as a rational being.⁴⁰

I shall return to the utilitarian and Kantian approaches to the wrongness of coercion on the non-normative preference-based account below. Prior to doing so though, we may notice that both the normative account of coercion and the non-normative preference-based account answer the first question outlined at the beginning of this chapter in broadly the same way. They both endorse the general picture that coercion undermines voluntariness because coercive proposals serve to frustrate a certain set of the victims' preferences (namely, to be in a different set of

³⁸ Zimmerman, 'Coercive Wage Offers'. Notice that Zimmerman's use of the term 'preferences' differs from Ekstrom's understanding (which I have been invoking), since Zimmerman does not claim that preferences presuppose the existence of higher order mental states in the way that Ekstrom does.

³⁹ Zimmerman, 'Coercive Wage Offers', 132.

⁴⁰ *Ibid.*, 129.

circumstances).⁴¹ They simply disagree on their understanding of how we ought to define and correspondingly distinguish threats and offers.

There are three reasons for thinking that the non-normative approach is at least more plausible than the normative approach. First, on the former approach, the wrongness of coercion is not parasitic on other normative violations. Second, it is more compatible with our intuitions regarding the effect of morally legitimate threats on voluntariness; the prisoner's decision not to escape in the example in the previous section qualifies as having been coerced, even if the warden is not threatening to violate his rights in any way. Third, the non-normative account classifies the slave-owner's proposal in Nozick's case as an offer rather than a threat in a manner that corresponds with natural use of the term 'offer'.

However, I believe that that we ought to reject both the normative and non-normative approaches that I have so far outlined, due to their reliance on the content-based approach to the distinction between threats and offers. Their reliance on this view leads them to obscure a feature of coercion that is indispensable to an adequate understanding of its implications for voluntariness, namely the significance of intention.

To see why, recall the examples of Alan and Bernie from section 1. I suggested that it is plausible to claim that Alan can consent to treatment but that Bernie has been coerced in a manner that undermines the voluntariness of his decision. If this intuition is correct, it cannot simply be the case that coercion undermines voluntariness *just* because it frustrates a preference, or because it involves the recipient moving unwillingly from one choice circumstance to another less preferable one. The reason for this is that this occurs in *both* Alan and Bernie's cases.

It might be argued that I am being slightly uncharitable in suggesting that these cases raise problems for the normative and non-normative preference-based approaches to coercion. After all, at least some advocates from both camps explicitly stress that coercion is an interpersonal, intentional phenomenon.⁴² Zimmerman, in particular stresses this by stipulating a prevention condition in his account of coercive offers. Recall that an offer will only qualify as coercive on his account if the proposer is *actively* preventing the recipient from being in a pre-proposal situation that they would strongly prefer to be in.

Whilst this is true, by cashing out the wrongness of coercion by appealing to the fact that it involves the frustration of preferences, these accounts obscure the role that intentional agency plays in the wrongness of coercion. This failure is clearest in the utilitarian approach to cashing out the wrongness of the frustration of preferences that coercion involves. According to the utilitarian interpretation, coercion is wrong because the frustration of desires per se is prima facie wrong. Yet, even if this is true,

⁴¹ Zimmerman phrases the moral wrongness of coercion in terms of its effects on freedom rather than voluntariness. However, since Nozick (and others) believe that the effects of coercion on freedom have implications for voluntariness, and in so far as it seems plausible to claim that a desideratum of a theory of coercion is that it is able to capture the thought that coercion undermines voluntariness, I shall interpret Zimmerman's claims about the effects of coercion on freedom to have implications for our understanding of its implications for voluntariness.

⁴² Nozick, *Anarchy, State and Utopia*, 272.

it is not clear why the fact that *the coercer*, rather than forces of hazard, is frustrating the individual's preferences is material to the wrongness of coercion on this approach. The problem with the utilitarian approach then is that it cannot account for the materiality of having one's desires frustrated by another agent to the way in which coercion undermines voluntariness.

In so far as intentional agency is a prerequisite of expressing an attitude of disrespect, the Kantian approach can claim that forces of hazard do not disrespect an individual's rational agency when they diminish her freedom. In this sense the Kantian interpretation is on stronger ground than the utilitarian approach to understanding the wrongness of coercion. I shall argue that the Kantian approach fails for other reasons that I shall detail in the next section. Here though, we need only acknowledge the point that the failure of the utilitarian approach in this regard suggests that in explicating the implications of coercion for decisional autonomy, we should focus less on the frustration of preferences it involves, and instead on its interpersonal features. That is, in order to adequately capture the effect of coercion on voluntariness in a way that is congruous with our intuitions in the Alan and Bernie cases, we have to focus not so much on the fact that the victim's preferences are frustrated, but rather on the fact that they are frustrated *by another agent*.

This echoes my discussion of the interpersonal sense of voluntariness that I highlighted at the end of the previous chapter. Coercion does not undermine voluntariness just because it frustrates the victim's preference; rather it does so because it subjugates the victim's will to the coercer's, violating the victim's freedom from domination.

However, the appeal to freedom from domination in the context of coercion raises a different problem, at least in view of the arguments I made in the previous chapter. There, I suggested that forces of hazard *can* manipulate or deceive an agent in ways that undermine their decisional autonomy. The question then is why intentional third-party influence should matter so much more in the case of coercion than it does in the case of manipulation—is there any relevant sense in which manipulation differs from coercion with respect to its implications for decisional autonomy?⁴³

Gideon Yaffe has offered one account that denies this sort of difference, but which purports to explain why intentional third-party influence in particular might matter for freedom. Yaffe argues that intentional influence can undermine freedom in a way that forces of hazard do not, because those who exert intentional influence will track the compliance of their victims.⁴⁴ If an initial threat fails, an intentional coercer is likely to increase the severity of their threats until their victim agrees to comply. This matters for Yaffe, because tracking compliance in this manner involves a greater restriction of the target's freedom, and forces of hazard will not typically track compliance in this way.⁴⁵

⁴³ Those who are sceptical of a plausible answer to this question have used this as motivation for accounts of autonomy that claim that only intentional interpersonal interference can undermine autonomy. Bublitz and Merkel, 'Autonomy and Authenticity of Enhanced Personality Traits'.

⁴⁴ Yaffe, 'Indoctrination, Coercion and Freedom of Will', 355–6.

⁴⁵ Yaffe acknowledges that there could be cases in which forces of hazard could track compliance, and that his conclusions would apply to such cases. *Ibid.*, 354.

I believe that Yaffe is correct to highlight this feature to show how some forms of interpersonal influence infringe upon the victim's freedom in a manner that forces of hazard will typically not. However, I do not believe that this feature of interpersonal influence is *necessary* for undermining decisional autonomy. To see why, suppose the coercer in Bernie's case issued her initial threat, but privately resolved not to threaten anything worse if Bernie failed to comply with the issued threat. On Yaffe's analysis, the sort of coercion involved in this permutation of the case would be functionally equivalent to the restriction of choice by forces of hazard in Alan's case.⁴⁶ In contrast, it seems plausible to claim that Bernie may not be autonomous with respect to his decision, even if his coercer would hypothetically not track compliance in the way that Yaffe describes. To claim otherwise would be to adopt a highly revisionist understanding of the relationship between coercion and autonomy in bioethics.

However, discussion of Yaffe's position points towards a solution to the puzzle at hand. What actually seems to be important about a coercer's tracking compliance is that they are reducing another agent's freedom with the intention of affecting their victim's practical decision-making in a manner that will serve the coercer's *own purposes*. In contrast, when our freedoms are reduced by hazard, there is no rhyme or reason to it; as such, even though we might say that an agent's freedom can be diminished by hazard, it does not involve a *subjugation* of the coerced agent's will.

But why should this matter so much in the case of coercion? To begin to explain why, it is useful to compare how coercion and intentional manipulation violate the victim's freedom from domination. Compare the case of Terry with Simon. Suppose that prior to the thief's intervention, both Terry and Simon would harbour the following evaluative ranking of their available conjunctive options:

Best = Option A: (Don't hand over wallet) + (My money is safe).

Worst = Option B: (Hand over wallet) + (My money is stolen).

Psychological manipulators might succeed in getting Simon to hand over his wallet by causing him to change his evaluative ranking of the available options (so that he now prefers option B to option A) without rationally engaging with his evaluative beliefs.⁴⁷ Let us suppose that this new preference ranking does not cohere with Simon's other core preferences and acceptances, those that led him to develop the initial preference ranking outlined above. Simon, nonetheless unreflectively acts on the manipulated preference, despite its conflict with his other preferences and acceptances; the preference is not responsive to his beliefs about the good.

My arguments in the previous chapter suggest that the violation of Simon's freedom from domination here is contingent on the fact that Simon lacks reflective autonomy with respect to this manipulated preference. Imagine now another thief targets someone with a different character system, Rupert. Rupert is somewhat repelled by his own wealth, and thinks that his life would be easier if he gave his money away; he acquired his wealth through complicity with things that he took to be immoral, and it deeply conflicts with his overall self-conception, and his positive

⁴⁶ *Ibid.*, 354.

⁴⁷ Yaffe notes that indoctrination can evince a new pattern of taking facts to be reasons for acting in particular ways amongst its victims. *Ibid.*, 342.

evaluation of helping others. Although he does not know what to do with his money, when faced with a would-be thief, he does not feel inclined to hand over his wallet.

Suppose that Rupert is similarly manipulated through hypnosis into changing his motivating desire. Unlike Simon though, Rupert comes to reflect on his changed motivation following his hypnosis, and comes to rationally endorse this ranking as being coherent with other core preferences and acceptances that have been in conflict with his initial preference not to hand over his wallet (for example, his belief that his money is morally tainted, that the thief is desperate, and his preference to help those in need). In Rupert's case, the account I have outlined in the previous chapters suggests that this psychological intervention does not violate Rupert's reflective autonomy; despite its causal history, his motivating desire is sensitive to his other pre-existing preferences and acceptances, in particular, his beliefs about the good. In view of my arguments in the previous chapter, this manipulation also does not violate his freedom from domination, insofar as it is Rupert *himself* who comes to rationally assimilate this manipulated preference into his own coherent self-conception.

In contrast to the cases of Simon and Rupert, when the thief *threatens* rather than manipulates Terry, his initial preference ranking is left intact; he still prefers option A to option B. However, the thief's coercive threat serves to take away option A from Terry's available choice set, and replaces it with option C which, for Terry, is less good than B:

Option C: (Don't hand over wallet) + (My money is safe) + (I am killed).

This analysis elucidates the following point. As Simon and Rupert's cases suggest, when manipulation *does* violate freedom from domination (i.e. in Simon's case), it does so by virtue of the fact that the third-party intervention undermines the reflective autonomy of another. When attempted manipulation fails to undermine reflective autonomy (i.e. in Rupert's case) it fails to amount to domination of the target's will. The third-party influence is immaterial to the individual's sustaining the desire following rational evaluation. Although the intervention caused Rupert to initially form the desire to hand over his wallet, it was Rupert himself who decided to rationally sustain the desire in the light of his other preferences and acceptances. In so far as Rupert sustained the desire on the basis of this sort of reflection, he can aptly be construed as the driving force underlying his decision to hand over his wallet to the thief; in hypnotizing Rupert to have this desire, the thief merely planted a seed in fertile ground.⁴⁸

In contrast to manipulation, in coercion, the coercer must *rely* on their victim retaining their reflective autonomy if they are to dominate their victim's will. The control exerted in coercion does not function by undermining the victim's reflective autonomy; rather, it functions by influencing the practical parameters within which

⁴⁸ Such an interpretation does not commit me to claiming that the thief's intervention was therefore morally neutral. Even if it does not undermine Rupert's autonomy, we may still plausibly claim that the thief's intervention violated his right to mental integrity for example. For discussion, see Bublitz and Merkel, 'Crimes Against Minds'. See also McKenna, 'Responsibility and Globally Manipulated Agents'.

the victim's reflective autonomy will operate.⁴⁹ The coercer influences their victim to act in a certain way by taking away their most preferred conjunctive option from their choice set (option A above), and replacing it with an option (C) that is less preferable than the option that reflects the coercer's will (B). Accordingly, although coercion does not undermine reflective autonomy, it may still plausibly be understood to amount to a domination of its victim by virtue of this interference with the agent's practical freedoms, and the imposition on the agent's choice domain connoted by that interference.

This analysis is also compatible with the Aristotelian claim that coercion is 'mixed' with regards to voluntariness, without denying that we can draw a morally significant distinction between the cases of Alan and Bernie above (or Aristotle's own examples of the cargo and the tyrant). The reason that we believe that coercion is 'mixed' is that, on the one hand, we must concede that coerced agents would reflectively endorse their compliant motives; in fact, in order to formulate an effective threat of the sort that coercion requires, the coercing party must make an accurate assessment of the kinds of preferences and values that are central to their victim's character system. In this sense then, we can understand the victim as voluntarily performing the act, the non-performance of which would lead to consequences that he has strongly decisive reasons to avoid.

On the other hand though, in so far as it is a third party that is responsible for changing the parameters of the individual's choice situation in a manner that serves to further the coercer's own ends, coercion involves a subjugation of the victim's will, and violates his freedom from domination. Whilst this freedom pertains to a different sense of voluntariness from the sense related to reflective autonomy, it is nonetheless central to our assessment of ourselves as autonomous agents. Whilst in manipulation, the agent's reflective autonomy and their freedom from domination stand and fall together, in coercion they can be mixed in the way that Aristotle describes in his examples, and that I have explicated above.

I shall further support this point in the following section by offering a revised account of the distinction between threats and offers. To conclude this discussion though, one might object that my claims here contradict my assessment of the implications of global manipulation for autonomy in the previous chapter. Recall that I previously claimed that global manipulation does not undermine autonomy, but rather amounts to the creation of a psychologically 'new' person. I suggested that this new person could still be autonomous, despite the fact that her 'creation' was the result of another's agency. This might seem to be in tension with what I am claiming here, namely that the victim of coercion may lack decisional autonomy even though she meets the conditions of practical and theoretical rationality discussed in Chapter 2.

The key to reconciling these two contrasting claims about global manipulation and coercion is recognizing that in coercion, the central elements of the agent's character system are retained. We can sensibly say that the 'will' that is subjugated in coercion

⁴⁹ In a similar vein, Yaffe suggests that coercion involves the manipulation of what reasons its victim has. Yaffe, 'Indoctrination, Coercion and Freedom of Will', 340.

existed both prior to the threat and following it. In contrast, in global manipulation, the will that exists following manipulation is, *ex hypothesi*, distinct from the will that existed prior to it. Global manipulation thus does not subjugate an extant will in a manner that undermines the autonomy of the agent following the intervention; rather it involves replacing the extant will with another of the manipulator's own design. This is an egregious harm; but it is, I suggest, a harm concerning identity, rather than autonomy. Coercion, in contrast, is a harm concerning autonomy that relies on the maintenance of the elements of identity that global manipulation destroys.

4. A Structural Account and Coercive Offers Revisited

In order to adequately understand why coercion undermines autonomy on my approach, we cannot merely appeal to the fact that coercion frustrates a certain kind of preference; we must also stress the fact that coercion violates the victim's freedom from domination. I shall now argue that we can better capture this aspect of coercion by abandoning the content-based view of the distinction between threats and offers, which is adopted by both the normative and non-normative accounts analysed above, instead adopting a structural account of the distinction between threats and offers. As well as capturing the moral significance of freedom from domination and interpersonal voluntariness, such an account is better placed to respond to Joel Feinberg's famous lecherous millionaire coercive offer example (outlined below), and the putative examples of coercive offers in bioethics that I delineated in section 1.

As I mentioned above, one reason that theorists have been attracted to the content-based view of the distinction between threats and offers is that these proposals seem to share the same sort of structure. Broadly, it might be claimed that it is true of both threats and offers that the proposing agent, P, claims that *she* will bring about consequences C if and only if the proposed-to agent, Q, does some action A.⁵⁰ If it were the case that offers and threats both had this precise structure, then it seems that the only way in which we could distinguish the proposals would be to appeal to the content of C, in the ways discussed above.

However, contrary to initial appearances, this analysis of the shared structure of threats and offers overlooks an important structural difference between the two types of proposal. In the case of offers, the proposer will bring about a certain consequence C that otherwise would not have occurred if and only if (*iff*) the proposed-to party *complies* with the proposer's demand to do A. In the case of offers, C is normally something that the proposer believes will make the recipient of the proposal better off; however, this is not a necessary feature, as the content-based view implies.⁵¹ In contrast, in the case of threats, the proposer will bring about a certain consequence C that would not have otherwise occurred *iff* the proposed-to party *refuses* to comply with the proposer's demands to do A. Again, we might note that in the case of

⁵⁰ Anderson, 'Coercion'.

⁵¹ For examples illustrating this point, see Sachs, 'Why Coercion Is Wrong When It's Wrong'.

threats, C is normally something that the proposer believes will make the recipient worse off, although this is not a necessary feature on the account that I am outlining here.

Accordingly, contrary to Anderson's analysis, there is a structural difference that we can appeal to in order to distinguish threats and offers. On this structural approach, the slave-owner's proposal in Nozick's example qualifies as an offer. As I illustrated above, non-normative approaches that endorse the content-based view of the distinction between threats and offers get the correct result here. However, they do so for the wrong reasons; the proposal is not an offer just because C is a 'good' consequence for the slave. Rather, the reason that the proposal is most appropriately understood as an offer is by virtue of the structural feature that the slave-owner will refrain from beating the slave *iff* the slave *complies* with his owner's demand to do A.

Compare this to a case in which the slave-owner will bring about a very *bad* consequence *iff* the slave complies with his demand to do A, say an even *worse* beating than usual. Such a proposal would be quite bizarre, and would be highly unlikely to be successful in motivating the slave to do A. Regardless of the bizarre nature of the proposal though, the badness of the consequence does not render the proposal a threat, as the content-based view would imply. Rather, such a proposal would most plausibly be understood to constitute a very unattractive offer, by virtue of its structural features.⁵² By appealing to the nature of the consequences that the proposer announces a conditional intention to bring about in each case, the content-based account is focusing on something that is only *contingently* true of *most* threats and offers.

This structural understanding of threats and offers serves to elucidate the salience of interpersonal interference to the way in which coercion undermines voluntariness, as I began to describe at the end of the previous section. When P makes a threat, they make it the case that Q can no longer act in a way that is entirely independent of P's interference. Recall the options available to Terry (the original thief target), as I described them above, options B and C.

Option B: (Hand over wallet) + (My money is stolen).

Option C: (Don't hand over wallet) + (My money is safe) + (I am killed).

Notice that regardless of which option Terry chooses, the thief's influence will be apparent, either because Terry will suffer the consequences that the thief will bring about if he refuses to comply (in option C), or because he will comply with the thief's demands (in option B).

Contrast this with the case of an offer. On the structural account, an offer announces a conditional intention to bring about a certain consequence C if and only if Q *complies* with P's demands. In the case of an offer then, P's influence does not infect every option available to Q; P has only influenced Q's choice set by adding a further option of Q's performing some act for an inducement (i.e. consequence C). However, unlike the case of a threat, there will be no consequences for Q if she refuses

⁵² Sachs calls such a proposal a 'ridiculously bad offer'.

to comply with the demand. As such, Q's pre-proposal status quo option remains intact in the post-proposal situation; in contrast, in the case of a threat, P's proposal takes away Q's freedom to maintain her pre-proposal status quo situation without further consequence.

Accordingly, we may say that in issuing an offer, the offering party influences the recipient's choice domain, but that this influence does not amount to the exertion of controlling influence. The reason for this is that offers leave open an option that is free from P's influence, namely Q's pre-proposal status quo option. If Q chooses to accept the offer, it is because she believes that compliance with what is demanded by the offering party in return for certain consequences is better than her extant circumstances, over which P will typically not exert influence.

However, the situation is different in the case of a threat. In issuing a threat, the threatening party exerts their influence through the terms of their proposal over *all* of their recipient's available options, in order to further their own ends. If Q refuses, P will intrude upon Q's sphere of autonomy by bringing about some consequence C that would otherwise not have occurred, and that P will typically not want to obtain (at least in most cases). If Q chooses to comply, it is only on the basis that acting in accordance with P's demands is better than this alternative way in which P would otherwise intrude into the sphere of Q's autonomy.

Whilst both successful offers and threats involve a third party creating the reason that is operative in the recipient's post-proposal decision about how to act, it is only in the case of a threat that the creation of this operative reason requires an intrusion into all of the recipient's available options in that context, robbing the recipient of the freedom to maintain their status quo situation. In so far as this is true, the terms of a threat involve the directed domination of one party's choice domain in the interests of another, in a way that offers do not.

To conclude this part of the discussion, recall that proponents of the content-based view explicate the moral wrongness of coercion and its effects on voluntariness by appealing to the fact that coercion involves a certain kind of preference frustration; furthermore, they note that individuals are willing to receive (non-coercive) offers, and that this represents the key difference between threats and offers with respect to voluntariness. The above discussion of the structural account illuminates the fact that there are in fact *two* dimensions underlying the rational agent's unwillingness to receive a coercive threat. On the one hand, such threats tend to (but need not) announce a conditional intention to bring about bad consequences. This is what the content-based view emphasizes. However, coercive threats also involve the frustration of a different kind of preference, one that is grounded by a desire to be free from domination. The reason for this is that threats involve the exertion of third-party influence over every available option in the recipient's choice set in a particular context. By emphasizing this supplementary feature of threats, the structural account makes salient the interpersonal aspect of coercion's moral wrongness.

Similarly, there is often a dual basis for why rational agents would be willing to receive offers. First, offers tend to (but need not) announce a conditional intention to bring about a consequence that would be good for the recipient if she complies with the demand. However, the recipient might also be willing to receive an offer insofar

as it is often (though not always) advantageous to have more options available to one.⁵³

This structural account of the distinction of threats and offers has important implications for our understanding of coercive offers. To see why, consider first the following example:

Lecherous millionaire: B is in an otherwise hopeless condition from which A can rescue her if she gives him what he wants. He will pay for the expensive surgery that alone can save her child's life provided that she becomes for a period his mistress.⁵⁴

On all of the accounts that I have surveyed in the first sections of this chapter, the proposal in this example uncontroversially qualifies as an offer. The key area of contention is whether we should understand it to constitute a *coercive* offer that undermines voluntariness. On the preference-based non-normative account, the question turns on whether we can construe the millionaire as preventing the woman from being in a feasible alternative pre-proposal situation. It might be claimed that the millionaire is preventing the woman from being in this situation by omission. Yet, Zimmerman rules out this interpretation: he claims that an offer is only genuinely coercive if the offering party *actively* prevents the recipient from being in their preferred feasible alternative pre-proposal situation.⁵⁵ Merely omitting to provide Q with the means to be in an alternative preferable situation is not sufficient for genuine coercion. As such, the non-normative account denies that the lecherous millionaire's proposal is a coercive offer.

Joel Feinberg, who originally posed the example, claims that the offer *is* coercive. In contrast to the preference-based approach, Feinberg endorses what he terms a 'compatibilist approach', according to which offers can be coercive even if they *enhance* freedom overall. On the compatibilist approach, an offer is coercive if it meets three conditions. First, the offer must have been made with coercive *intent*; that is, the offering party P must have made the offer with the intention of getting the recipient Q to do what P wants her to do. Second, the offer must have a coercive *effect*; that is, the offer must succeed in getting Q to do what P wanted her to do.⁵⁶ Feinberg goes on to suggest that in order for an offer to be coercive in effect it must impose a sufficient degree of 'differential coercive pressure', where differential coercive pressure is understood to refer to 'the gap between the value tag of what is offered and the price tag of what is required';⁵⁷ the greater the gap, the more coercive the proposal.⁵⁸ Finally, in addition to imposing a sufficient degree of differential coercive pressure, in order to qualify as a coercive offer, the proposal in question

⁵³ This is not to say that this is *always* advantageous; see Dworkin, *The Theory and Practice of Autonomy*, ch. 6, and my discussion in Chapter 5.

⁵⁴ Feinberg, *The Moral Limits of the Criminal Law*, 229.

⁵⁵ Zimmerman, 'Coercive Wage Offers', 132.

⁵⁶ Feinberg, *The Moral Limits of the Criminal Law*, 233–5.

⁵⁷ *Ibid.*, 234.

⁵⁸ Feinberg lists a number of other measurements of coercive pressure and their effects on voluntariness, including what he calls 'coercive force', 'total coercive burden', 'the coercive minimum'. However, in discussing coercive offers, he suggests that if our ultimate concern is to decide whether a recipient's consent was valid, the relevant measure is differential coercive pressure. *Ibid.*, 254.

must offer a 'prospect that is not simply much preferred, but one which is an exclusive alternative to an intolerable evil'.⁵⁹

The compatibilist approach adopts the broadly Aristotelian position, which claims that forces of hazard can exert coercive pressure just as intentional agents can; on the compatibilist approach, the woman in the lecherous millionaire case is understood to have essentially been coerced by her circumstances. I rejected this Aristotelian view of coercion above following my discussion of the examples of Alan and Bernie; it is highly revisionist to claim that non-agential forces can coerce in a way that invalidates consent in bioethics. Feinberg too seems to recognize this, and goes on to amend the compatibilist position in a manner that seems to bring it broadly into line with the preference-based view. He goes on to claim that on the compatibilist view, not all coercive offers undermine voluntariness to the extent of invalidating the recipient's consent: rather they will tend to do so if P's coercive offer is made after P has already *created Q's circumstances of vulnerability*. However, P's coercive offer frequently will not invalidate Q's consent when he was not responsible for creating those circumstances.⁶⁰

The preference-based view and the amended compatibilist view are thus in broad agreement about a crucial necessary condition of what it is for an offer to be coercive in a manner that invalidates consent; the preference-based view's active prevention condition and the amended compatibilist's condition about the creation of the victim's circumstances of vulnerability seem to identify a very similar feature of coercive offers. The main difference between the two is that the compatibilist view holds that some offers can be coercive without diminishing voluntariness to the extent that is sufficient to invalidate consent.

In their defence of these accounts, Zimmerman and Feinberg both say little to defend their respective conditions, instead relying on their intuitive appeal. Yet, as I mentioned in the previous section, it is difficult to see how the utilitarian approach to cashing out the wrongness of coercion on the preference view can accommodate the significance of this. I am now also in a position to further explicate the problems with the Kantian approach. Recall that the Kantian approach claims that the wrongness of the way in which coercion frustrates the victim's desires lies in the fact that it fails to acknowledge the victim's full status as a rational being. The problem is that even if this claim is plausible, it does not explain why an offer will only *invalidate consent* if the offering party is actively preventing the recipient from being in their preferred feasible situation. The reason that this is problematic is that offers can plausibly fail to acknowledge the victim's status as a rational being without meeting the aforementioned necessary condition of coercive offers that invalidate consent. Indeed, the lecherous millionaire's proposal is a striking example of this. The proposal plausibly fails to acknowledge the recipient's full status as a rational being; however, as both the preference view and the amended compatibilist view hold, we may plausibly deny that the lecherous millionaire's offer invalidates the recipient's consent. The Kantian approach to understanding the wrongness of coercion on the preference view fails because it leaves us with the question of why

⁵⁹ Ibid., 234.

⁶⁰ Ibid., 244.

offers that fail to acknowledge the recipient's full moral status only invalidate consent in some cases but not others.

In contrast, the new structural account of the distinction between offers and threats, and my discussion of freedom from domination, can help to elucidate the importance of the conditions that the preference view and the compatibilist view each adopt to accommodate valid consent in cases like the lecherous millionaire. To phrase this using the terminology of the compatibilist approach, when the offering party creates the recipient's circumstance of vulnerability, their proposal is similar to a coercive threat, insofar as the proposing party can then be said to have exerted influence over all of the options in the recipient's choice set in this context. In the case of normal non-coercive offers, I suggested that P's influence need not infect every option available to Q, since Q can simply choose to maintain her pre-proposal status quo situation, free from Q's influence, and without further consequence. This is not possible in the case of a threat. Moreover, it is also not possible in cases in which P makes an offer to Q having *already* placed Q into circumstances of vulnerability. In such a case, although the recipient may choose to remain in her pre-proposal status quo situation following the offer, this situation is one in which her will is *already* subjugated to her coercer's, by virtue of the fact that the coercer has created her current vulnerable circumstances. This is the situation in which the slave finds himself in Nozick's example.

One might object to the account that I have outlined here by suggesting that it gets the wrong conclusion in the lecherous millionaire case; contrary to the analysis that I have offered here, perhaps it might be claimed that the voluntariness of the victim in the lecherous millionaire case has been diminished in a manner that is sufficient to invalidate her consent. In support of this alternative interpretation, we might note that we would be reluctant to enforce a contract based on this kind of exploitative offer.

In response, the first thing to note is that there can be reasons for not enforcing a contract that have nothing to do with voluntariness of an individual's decision to accept its terms. My claim that the victim in the lecherous millionaire case provides valid consent thus does not commit me to the claim that we ought to enforce the terms of the contract. Morality might require that the millionaire simply saves the child without making any other demands. The woman might also claim that the contract is unduly exploitative. Yet this concerns a different set of issues about fairness that are distinct to those concerning whether the woman can autonomously agree to the terms of the millionaire's offer at the time that it was made.

More broadly, I am prepared to bite the bullet in maintaining this interpretation of Feinberg's example because of the importance of allowing for the possibility that people can make voluntary decisions in very poor circumstances. Indeed, to claim otherwise is to raise a grave threat to the autonomy of vulnerable individuals. Consider this permutation of the lecherous millionaire case: Suppose the woman is about to agree to the millionaire's offer in return for saving the life of her child. Suppose that we intervened and stopped this contract from being signed, citing the woman's lack of voluntariness. We surely could not congratulate ourselves for safeguarding the woman's autonomy as she now watches her child perish. The lesson here is that we should allow for the possibility that people can make autonomous

decisions in desperate circumstances, and that our moral obligation in such cases is to alleviate these circumstances of vulnerability. If we fail in that obligation, this does not mean that we have the right to prevent the vulnerable individual from exercising a right to determine how to manage the terrible situation she finds herself in, and which we have failed to alleviate.

I suggest that we are now in a position to answer the questions delineated at the beginning of this chapter as follows: First, whilst acknowledging Aristotle's observation that coercion is mixed with regards to voluntariness, we can understand it to undermine decisional autonomy by virtue of the fact that the coercing party dominates their victim's will. The coerced agent may thus be understood to choose voluntarily in the sense that they retain reflective autonomy with respect to their choice, but they lack the sort of interpersonal voluntariness that is connoted by retaining freedom from domination. Second, on an adequate, structural understanding of the distinction between threats and offers, both kinds of proposal can be coercive. We should reject the claim that coercion necessarily involves the use of threats. Although offers typically enhance the recipient's freedom, some offers can be coercive if the offering party is actively preventing the recipient from being in a preferable alternative pre-proposal situation, or by creating their circumstances of vulnerability; this interference with the recipient's circumstances is necessary to ensure that the offer serves to dominate the recipient's will, even if it can be understood to enhance their freedoms in their post-interference circumstances. This conclusion echoes the conclusions of both the preference view and the compatibilist view; however, I have provided a new explanation of why this is the right conclusion to draw, that links our understanding of threats and offers to an account of the relationship between coercion and voluntariness that is absent in the preference and compatibilist views. In the final section, I shall explain how this theoretical analysis can help us in practical discussions of apparently coercive offers, with reference to the cases delineated in section 1.

5. Practical Applications

The first thing that should be clear from my above analysis is that the quoted justifications for claiming that the offers outlined in section 1 are coercive are unconvincing. Offers are not coercive just because they are made against the background of a substantial loss of freedom; to claim otherwise would jeopardize the possibility of informed consent in various medical contexts. However, there are a number of things we can learn from considering the practical application of theoretical accounts of coercion and autonomy.

First, the terminology of coercion in practical debates does not always follow the philosophical use, or map on to the kind of wrong that the philosophical use aims to capture. Consider for instance the case of the TGN1412 trial. On the philosophical analysis I have provided, the proposal to withhold the incentive of payment following withdrawal from the study can be understood as the negative expression of an offer. The initial offer amounted to a proposal to provide payment (which would otherwise not have been provided) *iff* the proposed-to party *complies* with the proposer's demand to fully participate in the study. Proposing to withhold the incentive if the

proposed-to party does not comply is to restate the proposal in negative terms. Having established that the proposal in question is an offer, one can then ask whether it is coercive. On the account I have explored above, for this to be so, the body running the study would have had to have been responsible for ensuring that the participants were in a circumstance of vulnerability prior to the offer being made.

Some of the ethicists who made the claim that the participants of the trial were coerced focused on the large amount of money that had been offered to participants.⁶¹ As my discussion of the nature of offers above should make clear, the problem with claiming that it is coercive to offer individuals large amounts of money in return for some service is that we often offer people inducements to perform tasks, and the recipients of these offers can clearly consent to these transactions. Indeed, in some cases, it may be irrational for them not to; I might plausibly be described as irrational if I refused the offer of £1million to give a lecture, say.⁶² Naturally, the amount of money offered is relevant to our understanding of whether an offer is going to be sufficiently attractive to succeed in changing the recipient's beliefs about what they have most reason to do. However, such an offer can only be said to amount to a domination of the will of the sort that coercion entails if it is considered as part of a larger scheme in which the offering party has already placed the recipient in a circumstance of vulnerability.⁶³

A plausible explanation of why ethicists might be concerned with the amount of money offered in these cases is that they are concerned about a different moral phenomenon. It might be suggested that being offered a large amount of money may lead the recipients of the offer to fail to adequately attend to other material information about the trial (such as the degree of risk to which they will be exposed); on this understanding, the offered amount may have overwhelmed the recipients into accepting the offer, no matter the cost.⁶⁴ Whether or not this was the case in the TGN1412 case is an empirical matter that I cannot address here. However, the point that I want to make is that this sort of interference with autonomous decision-making bears more resemblance to informational manipulation than the coercive interventions with which I have been concerned in this chapter. On this understanding, the offer serves to negatively affect the salience attributed to certain kinds of reasons, namely those associated with the risks of the trial. This may undermine autonomy; but it is, I suggest, obfuscatory to describe this intervention as coercive.⁶⁵

Second, determining the involvement of third-party agency is central in claims regarding coercive offers, and this can be difficult to determine in practical contexts.

⁶¹ Schonfeld et al., 'Money Matters'.

⁶² On this point, see also Wilkinson, 'Biomedical Research and the Commercial Exploitation of Human Tissue'.

⁶³ Emanuel and Miller agree with this broad conclusion that the offer in the TGN1412 trial was not coercive, but do so by appealing to a normative approach to coercion. Emanuel and Miller, 'Money and Distorted Ethical Judgments about Research'.

⁶⁴ Wilkinson calls such cases 'enormous offer' cases. Wilkinson, *Bodies for Sale*. Both Macklin and Nelson et al. seem to implicitly support the view that such offers can be coercive. See Macklin, 'The Paradoxical Case of Payment as Benefit to Research Subjects'; Nelson et al., 'The Concept of Voluntary Consent'.

⁶⁵ On a similar point from a normative theorist of voluntariness generally, see Wertheimer, 'Voluntary Consent', 248–9. See also Largent et al., 'Misconceptions about Coercion and Undue Influence'.

Determining agency is central to claims regarding coercion because a key aspect of coercive offers (at least those that invalidate consent) on the account that I have developed in this chapter is that the offering party is actively preventing the recipient from being in a preferable pre-proposal situation, or has intentionally created their circumstances of vulnerability. Only then can such offers amount to the sort of domination that undermines interpersonal voluntariness. When the recipient's less preferable pre-proposal situation has been created by forces of hazard, offers made to them cannot coerce in a manner that would invalidate their consent.

In the philosophical thought experiments I have been considering, it is easy to ascertain whether the proposing party was engaged in such active prevention. However, coercive offers in practice raise far more difficult questions in this regard. To illustrate, it is easy to imagine a clean philosophical thought experiment in which Jones offers poverty-stricken Smith £10,000 for one of his kidneys; and we can say that this offer would only be coercive on the account I have developed if Jones had already somehow actively made it the case that Smith were financially destitute, and thus in a position in which refusing the offer was not eligible for choice given his circumstances. Part of the reason that this case seems so artificial is that it would rarely be the case that one individual would be responsible for another's poverty in this way. However, when we broaden the scope of our consideration in the context of real world markets for human organs, it might be claimed that structural injustices in the global economy mean that rich nations can appropriately be construed as actively preventing people in poor nations from escaping poverty.⁶⁶ Similarly, in the context of medical trials, even if the team organizing the study were themselves not responsible for creating the financial situation of those who signed up to the study, the research may have been funded by a governmental body that may at least be complicit in creating the circumstances of vulnerability that lead individuals to agree to participate in risky trials.

This of course is a far more complex question that I cannot hope to address here, and much will depend on the details of particular cases. However, in the light of my analysis of coercion, defining the scope of what constitutes the 'offering party' in these transactions (at an individual, organizational, or national level), and their role in intentionally creating the underlying circumstances of vulnerability amongst recipients of certain kinds of offers, is central to understanding whether we should construe these practices as coercive. It is also central to distinguishing the wrongs that coercion involves, from the wrongs involved in manipulation and exploitation. The less direct, and the less intentional the offering party's contribution to their target's underlying circumstances of vulnerability, the less appropriate the charge of coercion becomes, and the more likely it is that the wrong involved in making an offer does not primarily concern autonomy. Instead, if such offers are wrong, it is likely for the reason that by omitting to remedy their target's circumstances of vulnerability without further cost to themselves, the offering party fails to fulfil a

⁶⁶ See, for example Rippon, 'Imposing Options on People in Poverty'; Annas, 'Life, Liberty, and the Pursuit of Organ Sales'. For a response to arguments of this sort, see Wilkinson and Moore, 'Inducement in Research'.

moral duty of easy rescue.⁶⁷ In any case, contrary to what is widely claimed, it is these issues regarding the relationship between the offering party and their recipient's circumstances of vulnerability which are where the real questions about autonomy lie for coercive offers, rather than the size of the inducement per se.

We may also notice that those who endorse normative accounts of coercion have to answer a different question in this context. Rather than ascertaining whether such markets are coercive by asking whether rich nations are actively preventing individuals in poor countries from escaping poverty, they have to ask whether rich nations are *morally failing* these individuals by failing to alleviate their poverty. Although I have suggested that we ought to reject normative accounts, this shows how different theories of coercion can lead to different normative interpretations of pressing issues in practical ethics.

Similar issues concerning the role of the offering party in creating circumstances of vulnerability also arise in the context of offering incarcerated criminals chemical castration in return for a reduced sentence. Normative accounts can easily deny that the offer is coercive, presuming that the offender has been legitimately incarcerated in an establishment that respects the fundamental rights of prisoners; in such circumstances, the recipient's rights are not being illegitimately violated in his baseline state of affairs, against which the offer is made to him. However, on non-normative accounts, in order to determine whether the offer is coercive, we have to ascertain whether the proposing party created the prisoner's circumstances of vulnerability (or whether they are 'actively preventing the recipient from being in a preferable alternative situation' to use the framing of the preference view). This seems implausible when we understand the proposing party as the individual psychiatrist who makes the offer to the offender; however, it seems more plausible if we understand the criminal justice system as a collective to constitute the proposing party. Contrary to Green's analysis quoted at the beginning of the chapter, the key issue in considering the coerciveness of this sort of offer is thus not the offender's deprivation of liberty per se; the key issue is whether the proposing party should be construed as the same party that is actively preventing the prisoner from being released.

A further interesting aspect of this example that my account helps to elucidate is the relevance of the state's intentions. John McMillan has argued that the coerciveness of the chemical castration offer partly depends on the intention underlying it: Does the proposing party make the offer with the intention that the recipient will accept it? If not, then it is hard to see how the offer is coercive.⁶⁸ I agree that the proposing party's intentions are important here, since it seems plausible that intending one's offer to be accepted may be necessary for that offer to constitute domination of the recipient's will. As I have claimed, this is central to understanding the moral wrongness of coercion. If, on the other hand, the proposing party does not care whether or not their offer is accepted, it is difficult to see how they are dominating the recipient's will, in so far as domination implies directing another to a particular end.

⁶⁷ For more on this duty, see Rulli and Millum, 'Rescuing the Duty to Rescue'.

⁶⁸ McMillan, 'The Kindest Cut?' See also Shaw, 'Offering Castration to Sex Offenders'.

Conclusion

In this chapter, I have mapped out some of the theoretical complexity surrounding the concept of coercion and its relationship to voluntariness. In doing so, I have developed a structural account of threats and offers that accords with our common understanding of these kinds of proposal. Moreover, this account can naturally be incorporated into a plausible explanation of why coercion undermines decisional autonomy, in a manner that makes salient its implications for interpersonal voluntariness. In turn, this has important implications for our understanding of putative coercive offers, a concept that has been widely invoked in a number of bioethical debates. In order for such offers to be coercive in a manner that invalidates consent, they must involve a violation of the target's freedom from domination. I have suggested that offers can involve such a violation if they are made to an individual who has been placed into a circumstance of vulnerability by the offering party.