



King's Research Portal

DOI:

[10.26556/jesp.v11i1.106](https://doi.org/10.26556/jesp.v11i1.106)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Owens, D. (2016). Promises and Conflicting Obligations. *Journal of Ethics and Social Philosophy*, XI(1).
<https://doi.org/10.26556/jesp.v11i1.106>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

PROMISES AND CONFLICTING OBLIGATIONS

Some obligations we create by declaration, by the exercise of normative power. In the case of a promise, we do so by communicating the intention of hereby undertaking an obligation. In the case of a command, we do so by communicating the intention of hereby binding someone else to perform. Or so I shall assume. Call these *voluntary obligations*. As I use the term, a *voluntary* obligation is not merely an obligation that results from our choice, nor even from our choice of an obligation. A voluntary obligation is the result of a specific form of choice, namely the decision to communicate the intention of hereby binding someone to perform. Of course there are various ways in which declarations or exercises of normative power may fail of their purpose but the fact that the speaker intended to bind someone by declaration is often crucial to their being so bound.

The claim that such a declaration could ever generate an obligation has been disputed.¹ Here I join other authors in assuming that normative power is a reality and so that voluntary obligations do exist but I shall also be assuming that many obligations are not in this sense voluntary. For example, parents are bound to feed their children regardless of promise, command or any other form of declaration. And this raises a question: how do voluntary obligations relate to involuntary obligations? In particular can one have a voluntary obligation to do something that is involuntarily forbidden? Here I am mainly interested in promises and thus in whether one can bind oneself to do what one is involuntarily obliged not to do but I shall also ask whether another can put me in this position by issuing a valid order.²

¹ This dispute explains the menu of theories of voluntary obligation considered in this paper. For example, I ignore expectational theories of promissory obligation that deny that promises are the products of normative power. For consideration of these theories see (Owens 2012: Chapter 9).

² Thus an order imposes a voluntary obligation on me even though it is not I who choose it.

At one extreme is the view that whilst declaration may create voluntary obligations where none previously existed, it can do so only within the limits laid down by our involuntary obligations. That is the view taken by advocates of what I shall call the 'natural right theory' of promissory obligation (see section 3). At the other extreme is the view (towards which I myself now incline) that the content of binding promises and valid commands is not constrained by our involuntary obligations as such.³ There is a related issue about whether our present voluntary obligations are restricted by those we have previously assumed (e.g. our past promises). I shall consider the latter issue in the second section and the former in the third but before getting down to business, we need a clearer view of what it is to be bound by an obligation.

1. Obligation

How one thinks about conflict of obligation will inevitably depend on how one thinks about obligation itself and there is no way of setting up the debate about the former without making claims about the latter, claims that some will dispute.⁴ In this section, I describe three marks of obligation. A *mark* of obligation is condition that any obligation must satisfy and furthermore a condition that helps to differentiate an obligation from other considerations that might persuade one to act. I shall further assume that these three marks suffice to differentiate obligations from all other practical considerations, so that any practical consideration that bears them constitutes an obligation. But the issue as to whether practical considerations bearing these marks can conflict is, I maintain, of theoretical interest whether or not they all constitute obligations. (For example, we need to settle whether you might find yourself in a situation in which guilt will be apt whatever you do.) So those who doubt that the presence of these three marks is enough for obligation may still engage with the argument of the next three sections.

³ On this point my own views have changed. Compare (Owens 2006: 72 n.28) with (Owens 2012: Section 48).

⁴ The claims about obligation made in this section are all defended at greater length in (Owens 2012: Chapter 3).

So what is distinctive of *obligation*? I'll start with the following claim: it *makes sense* for you to do something simply because you think yourself obliged to do it. If you think yourself bound by your promise to go on holiday with me then it makes sense for you to go on holiday with me even though you can see nothing else to be said for so doing, don't feel like going and so forth. Intentional activity is activity that makes sense at least to its agent, activity with some apparent point and your promise is enough to give your travelling with me a point. This is our first mark of obligation.

Now various considerations can make sense of our activities. What is special about obligations? Obligations don't just recommend fulfillment, don't just count in favour of fulfillment, they demand it. A demand does not leave it up to you to weigh all the considerations that favour fulfillment against those recommending breach. To comply with a demand is not merely to do what is demanded, it is to decide to do what is demanded without considering a certain range of factors recommending non-fulfillment. A conscientious person would not act on these considerations and so they are to be excluded from any practical deliberations you conduct on the point. So, for example, a fully conscientious promisor would not seriously consider the attractions of a subsequent invitation when determining whether to holiday with me as promised, they would not weigh the charms of this alternative entertainment against the need to keep their promise; they'll *exclude* these charms from their practical deliberations. This is our second mark of obligation.⁵

Thus far we have it that binding obligations make sense of performance and furthermore they make sense of your excluding various considerations counting against performance from your deliberations. Not that the promise excludes anything at all that counts against fulfillment: if I suffer a moderately serious injury just as I am about to leave for our holiday, I should indeed consider my health in determining whether to keep my promise to you. And if I determine that these non-excluded considerations justify non-fulfillment, then the promise no longer binds me and I am not wronging you when I fail to go. On the other

⁵ On obligation and exclusion see (Owens 2012: 85-95) and (Raz 1999: 35-48). If obligations can indeed conflict the excluded considerations may well include other obligations.

hand, suppose the injury (though non-excluded) is not especially serious and would not by itself justify breach but the other (excluded) attractions of the alternative holiday are both strong and unaffected by the injury. Here holidaying with you may be inadvisable in the light of the full range of relevant considerations, both excluded and non-excluded: I would be justified in breaching my obligation. Nevertheless, I am wronging you if I don't go because I am acting on considerations that my promise excludes. Here it would make sense for me go (despite my injury) even though I would be justified in breaching.

Have we exhausted the practical significance of obligation? Don't I also have a positive reason to do what I am obliged to do, a fact not obviously captured just by saying (a) that it makes sense for me to do it and (b) that it makes sense for me to exclude various conflicting considerations from my deliberations about whether to do it? To put it another way, if I'm really obliged to do it, doesn't that provide at least some grounds for thinking that I *ought* to do it? There are a range of positions here, the most extreme of which is the view that I'm not really obliged to A unless I ought to do A all things considered. Now if being obliged to A did indeed imply that one ought to do A, that one had sufficient reason to do it, then it would follow that obligations couldn't make conflicting demands on us, for it surely can't be the case that we ought to do A and ought also to refrain. True I might be justified in either eating an ice cream or not eating it as I please (here I am 'rationally permitted' to do either) but when on the whole I ought to be eating an ice cream, it can't also be the case that on the whole I ought not to be eating it (and *vice versa*).

We should follow Hart in resisting this view:

There is no contradiction or other impropriety in saying "I have an obligation to do X, someone has a right to ask me to, but I now see I ought not to do it." It will in painful situations sometimes be the lesser of two moral evils to disregard what really are people's rights and not perform our obligations to them. This seems to me particularly obvious from the case of promises: I may promise to do something and thereby incur an obligation just because that is one way in which obligations (to be

distinguished from other forms of moral reasons for acting) are created; reflection may show that it would in the circumstances be wrong to keep this promise because of the suffering it might cause, and we can express this by saying "I ought not to do it though I have an obligation to him to do it" (Hart 1955: 186)

Hart's point here is that there is a notion of obligation, one familiar from ordinary life that is not tied to what one ought to do *tout court*. This pre-theoretical notion is theoretically interesting because it both bears the marks of obligation that I am currently outlining and appears to leave room for conflicts of obligation. Whether there are other (perhaps theoretically motivated) notions of obligation on which the fact that you are obliged to do something entails that you ought to do it is not something I shall discuss.

One might think that one ought to do what one is obliged to do because the function of obligation is to guide action and an adequate guide won't give advice that should not be followed.⁶ But action guidance need not take this decisive form. For example, a genuine obligation may give me a weighty reason to fulfill it, a reason I might be justified in setting aside but only *in extremis*. Unlike its predecessor, this view does not exclude the possibility of conflict of obligation but it does give all such conflicts a tragic aspect.⁷ This overdramatizes the phenomenon. When I must choose between breaking my promise to have lunch with you or leave my child to wait around for a lift home from school, neither you nor my child face any great sacrifice but an obligation must be breached nonetheless. The truth is that some genuine obligations are quite trivial because the interests they protect are relatively unimportant (e.g. my duty to avoid a

⁶ Thomson calls this the Simplifying Idea and rejects it. She agrees with Hart that one can be obliged to do what one ought not to do and maintains that whilst obligations may conflict, oughts may not (Thomson 1990: 82-7). On the other hand, Thomson seems to think that if you ought to breach a promise then it follows that it is 'permissible' for you to breach that promise, even though the promisee still has a 'claim' against you (Thomson 1990: 98). I make no use of the notion of permissibility here since, in my view, it suffers from an unhelpful ambiguity. An action may be permissible in the sense that it is not the case that one ought not to do it, whilst also being impermissible in the sense that one wrongs someone by doing it, guilt is apt and so forth.

⁷ In his influential papers about moral dilemmas, Williams foregrounded tragic choices, choices which hard to make in that they are both hard to get right and hard on those affected by the choice (e.g. Williams 1973: 173). Neither feature is crucial to my discussion.

minor invasion of your personal space).⁸ Where obligations strongly recommend performance that is not simply because of their character as obligations but rather because of the weight of the interests they protect and this is as true of promissory obligations as of any other.

Still must obligations justify their fulfillment to at least some degree? Must there be some reason to fulfill any genuine obligation, some case for thinking that one ought to do what one is obliged to do? Whether this is so will depend on whether we can, merely by declaration, make it the case that something ought to be done (other things being equal). Since a positive answer would do nothing to rule out conflicts of obligation, I here remain agnostic, resting content with the claim that a promise can make sense of your doing something (i.e. can enable you to do it intentionally) by obliging you to do it.

We are not yet done with the marks of obligation. Obligation bears not only on how we should deliberate, it also affects how the agent and others should react to what is done. Suppose the agent breaches a binding obligation. Then *ceteris paribus* we'd expect the agent to feel guilty and others to blame them for breach. Guilt and blame are, in general, appropriate reactions to breach of promise for instance.⁹ (Indeed guilt and blame may be in place even where the promisor consulted their own convenience and performed for the wrong reason.) And wherever blame and guilt are in place, the issue of forgiveness arises: the promisor can ask for it and the promisee can dispense it because the promisor wronged the promisee.

Does this third mark of obligation throw any doubt on the possibility of conflict of obligation? Again various claims might be made. The most extreme is that one simply can't be a position where, whatever one does, it is appropriate to feel guilty about doing it and so conflict of obligation is impossible. Again the extreme view is implausible: when I've knowingly borrowed more than I can repay, shouldn't I feel guilty (and require forgiveness) whomever I end up repaying?

⁸ On this point see (Raz 1977: 223) and (Wolf 2015: 244-6).

⁹ These claims all invite multiple qualifications: the promisor has an excuse (e.g. is very depressed) or else we all have better things to worry about and so forth. I'll take such qualifications as read.

More plausible is the idea that I can find myself in such a situation only when I am at fault: I shouldn't have borrowed so much money.¹⁰ Perhaps I ought now to repay X rather than Y but since I got myself into this bind, I lack any justification for letting Y down and am to blame whatever I do. This more moderate view allows that our reactive attitudes can conflict but only because, at some point, we didn't discharge our obligations: a conscientious person will never find themselves in this situation. In what follows, I shall urge that even this weaker claim is too strong. One can have conflicting obligations through no fault of one's own and so can be bound to do something (e.g. keep a promise) that one is fully justified in not doing.

It might be urged that, in addition to the above marks of obligation, performance of an obligation must also be *owed to* someone, someone who is wronged by non-performance. This is extremely plausible in the case of promissory obligations: the promisor owes performance of the promise to the promisee and breach of a binding promise always wrongs the promisee. One mark of the latter is that the promisee is in a position to forgive the breach. Nor does this fact throw any doubt on the possibility of conflicting obligations: as already indicated, you can be so situated that whatever you do you will wrong someone by breaching a promise you have made them. Though I shall indeed be assuming that breach of promise is a wrong that takes the form of a wronging, I shan't write this feature into the very notion of an obligation because I'm doubtful that all breach of obligation involves wronging someone but this feature could be included without affecting the argument of the paper.¹¹

From here on the concept of obligation outlined above is taken as the basis for discussion and the three marks of obligation are used to test claims about the possibility or impossibility of conflict of obligation. Consultation of our intuitions about particular cases isn't the only test available. We might also ask a more theoretical question: what is the basis of our normative powers and why should the obligations they purportedly generate be taken at all seriously? Writers who

¹⁰ This is the position taken by (Foot 2002: 48) and (Thomson 1990: 97).

¹¹ In particular I want to leave it open whether failure to obey an order wrongs the authority that issues the order (or indeed anyone else).

share our concept of voluntary obligation disagree widely in their answers to these questions. In the rest of the paper, I'll review two such theories – the natural right theory and the social practice theory – and ask whether either of them provide any rationale for supposing that a promise (or an order) can only supplement and not countermand our prior obligations.

2. The Natural Right Theory: *Conflicting Promises*

According to the natural right theory, promises involve our exercise of a natural (pre-conventional) power to change the normative situation by declaration, a power that is rooted in a natural right to determine how we ourselves shall behave. Take my promise to have lunch with you today. I start off with the natural right to have lunch with you or not as I please. At a minimum this means that I am at liberty to lunch with you (or not) i.e. I can lunch with you (or not) without wronging you or anyone else. Having promised to lunch with you, I have lost that aspect of my liberty in that I can't fail to have lunch with you without wronging at least you. Many natural right theorists claim that there is a *transfer* of right here, that my power to bind myself has been acquired by you: you now have the power to determine whether I would be wronging you by failing to show up in that you can release me from my promise or else hold me to it. The notion that there is a literal transfer here – that some normative item that used to be the promisor's is now the promisee's – is questionable (Owens 2014: 79-81). My interest is in the more fundamental idea that the power to promise is grounded in the promisor's prior natural liberty to do what they are promising to do and I shall use 'natural right theory' to denote this claim.¹²

The natural right theory makes room for what I called voluntary obligation: it gives us a power to determine who is obliged to do what simply by declaring our choices in this matter but it also limits the scope of this power in a way that might be thought to preclude the creation of conflicting obligations. If our power

¹² Reinach holds that we have a natural power to make a binding to promise to commit murder etc, a power that is *not* grounded in any natural right to do the thing we are promising to do. He also distinguishes the power to promise from the power to transfer a right (Reinach 1983: 45-8 and 66-7).

over obligation can operate only in areas of liberty, only in regions that are currently free of obligation then it looks as if we cannot, by exercising that power, create conflicting obligations for ourselves.¹³ Or so it might appear. In the event, most natural right theorists adopt a more nuanced approach.

We find a relatively clear statement of the natural right theory in Grotius. Grotius, like many authors of the period, is a transfer theorist. He thinks of our (natural) rights and liberties as things that we (naturally) own and compares a promise to the alienation of an item of property:

And this is a complete promise, as having the same effect as the alienation of a man's property. For it is either the introduction to the alienating of a thing, or the alienation of some part of our liberty. To the former belong our promises to give, to the latter our promises to do something (Grotius 2005: 704).

This passage raises many questions but I want to focus on one issue: does the natural right theory imply that you cannot make a binding promise to do that which you are not at liberty to do?

This is what Grotius says:

To make a promise firm, it is requisite that the thing promised either now is, or may be, in the power of the promisor; wherefore in the first place, it is certain, that no promise can oblige us to do that, which is in itself unlawful; for no man has a power to do any such thing, or can have. But a promise ... receives its force from the power of the promisor, nor does it reach any farther.¹⁴

The waters are muddied by the phrase 'in itself unlawful'. Grotius appears to be distinguishing a promise to do what is not intrinsically impermissible from a promise to do what is intrinsically impermissible. My having lunch with you is not intrinsically impermissible whilst our feasting on human flesh is. I suspect

¹³ If the natural right theorist requires that the promisor have not just a liberty but also a claim to do what they are promising to do then conflicts between obligations taken on by *different* people may also be avoided.

¹⁴ (Grotius 2005: 714). See also (Locke 1988: 284).

Grotius is inviting us to conclude that, while I can't promise to share a cannibal lunch with you, I can make conflicting social arrangements and so I might after all make a binding promise that can be fulfilled only by wronging someone.¹⁵

Not all transfer theorists concede this possibility. For Hobbes there is no such thing as an intrinsically impermissible act. All Hobbesian obligations originate in a promise and so impermissibility is always extrinsic. Hobbes furthermore insists that binding promises cannot conflict and so that conflict of obligation as such is impossible. Recent natural right theorists like Shiffrin and Watson are less blank, drawing a Grotian distinction between promises to do something intrinsically immoral - promises they deem invalid - and conflicting promises that are sometimes valid. But before assessing these recent views, let's consider Hobbes more closely.

A former covenant makes void a later. For a man that hath passed away his right to one man today, hath it not to pass tomorrow to another; and therefore the later promise passeth no right, but is null.¹⁶

Here are two claims: first that it is impossible to make binding promises that conflict – only one such promise can bind; second that where I attempt to make conflicting promises, it is the prior promise that binds.¹⁷ I'll take the latter claim first.

Hobbes assumes that where one is obliged to do something, one ought to do it. Now when faced with conflicting promises, it is very often the case that I ought to fulfil one of them and not the other but it is not generally true that the promise I ought to fulfil is the one I made first (Thomson 1990: 89). Suppose I promised to have lunch with you and later promised to help John move house at the same time. Here I may have forgotten my earlier promise when I make the later one but equally I may not: perhaps I just ignore my earlier promise in the light of John's need for help. Still it might be that ought to break my social engagement

¹⁵ This may also be the best way to interpret what Rosati says on the matter (Rosati 2011: 128-9). As Rosati says, I myself took something like this view in earlier work: see note 3.

¹⁶ (Hobbes 1994: 86). (Pufendorf 2003: 118) says the same.

¹⁷ (Pink 2009: 404-6) argues that the transfer theorist is (implausibly) committed to both claims. (Shiffrin 2011: 163-5) denies this.

with you in order to help John move house. The fact that I first agreed to meet you for lunch doubtless counts for something but the fact that John won't be able to move at all unless I help out counts for more. I ought to break the prior promise I made to you. Thus Hobbes can't maintain that only the earlier promise binds.

Might Hobbes simply abandon the idea that prior promises nullify their successors - allowing that a later promise can invalidate an earlier - whilst maintaining that both promises cannot bind, that you can't surrender your liberty in favour of both parties?¹⁸ In determining what to do tomorrow lunchtime, I may reflect that my later promise was more solemn than my earlier, that the costs of breach are greater and so forth. Perhaps I can't entirely get myself off the hook of having lunch with you just by solemnly agreeing to meet another - I still owe you an explanation, a free lunch and so forth. But, one may think, I no longer have a promissory obligation to lunch precisely because of my subsequent promise.

Unfortunately the marks of an obligatory lunch remain. First, it would still make sense for me to have lunch with you simply because I promised to have lunch with you. This would be the wrong thing to do and likely involve a display of weakness on my part but it would hardly be an unintelligible act. Furthermore it would make sense for me to ignore (say) my own convenience in deciding which promise to keep, to exclude such considerations from deliberation. Finally it would be appropriate for me to feel guilty about any breach of promise and to seek your forgiveness, not simply to express regret.

The last point is merits expansion. Suppose you have a limited quantity of a life-saving drug to distribute amongst a group of sick people with whom you have no prior connection. All are equally deserving but not all can be saved. Here you bitterly regret having to decide who lives and who dies. You may also wish that this task had fallen to someone else or, more charitably, that nature would take the matter out of your hands. Once the choice is made you will (quite reasonably) feel awful about the resulting deaths, not just *qua* compassionate

¹⁸ Perhaps a later property transfer can also invalidate an earlier one, provided delivery (i.e. performance) has not yet occurred.

observer but as someone who brought those deaths about, feelings not fully assuaged by your confidence that you made the right choice. *Still* in so far as you are convinced that you did the right thing, you won't think of yourself as having wronged those who died and so you won't feel guilt. You might well ask for their understanding but not for their forgiveness.

Now suppose that you actually promised the drug to one group of people. You then discover to your horror that more lives may be saved by giving it to another previously unknown group. Even if you are convinced that you should save the greater number, doesn't your promise affect how you would think and feel about what you must now do? (After all, using the drug to save other lives is exactly what this promise was meant to exclude.) Wouldn't you vastly prefer that you had not made this promise at all? True, you'd feel awful about having to deny the drug to anybody but when reallocating the drug involves a breach of promise, there is a further element of distress, one that registers the fact that you are now (albeit justifiably) wronging those people. The distress here is not just a deeper regret, as if you'd discovered that the group who must be sacrificed is larger than you thought. It is a feeling of a different sort, one that looks for forgiveness. A less biting form of the same feeling is apt when you stand one person up to go help another because you promised to help.

Grotius says that we cannot be bound by a promise unless it is in our power to keep it and this might be taken to imply that moral necessity is as much of a shield against guilt as physical necessity (Foot 2002: 43-4, 48-9). On this view, if I can't keep the promise because I must save the lives of others then it is just as if some physical incapacity prevents me from keeping my promise. Were this so I should be indifferent between a situation in which I must choose to break a promise to some in order to save the lives of others and a situation in which nature forces my arm. Yet I'm far from indifferent: nature's intervention relieves the guilt of breach of promise in a way that mere moral necessity (like saving more lives) never could. Though moral necessity justifies a breach of promise it does not dissolve the obligation to keep it, nor relieve the burden of guilt at breach.

Given all this the following question arises: when one breaches such an obligation, what is the significance for one's relations with other people of the fact that one is justified in breaching it? We are supposing that such a justification does not abolish the obligation and that we are, for instance, still wronging the promisee by breaching our promise. So does it matter to anything other than our regard for our own practical rationality whether we are also justified in breaching it? People standardly offer a justification as a way of shielding themselves from blame and where one's breach of promise is justified this does indeed mean that one should not be *blamed*, even by the promisee. That is usually why people wish to be justified in what they do – they thereby deflect this form of criticism.¹⁹ However it remains the case that you have wronged the promisee and might well ask for their forgiveness. That request makes sense because it is still appropriate to feel *guilty* about breach of promise, even when blame would be out of place because your breach is justified. To ask for forgiveness is not just to express regret: it is to seek to relieve yourself of the burden of guilt (Owens 2012: Section 17).

The natural right theorist should acknowledge that promissory obligations can conflict where we overcommit ourselves. Still our voluntary obligations might at least be constrained by our involuntary obligations. In the next section we'll address this further issue.

3. The Natural Right Theory: *Wicked Promises*

Natural right theorists (Hobbes excepted) are agreed that there is a certain class of intrinsically impermissible acts, acts that one is not at liberty to perform and therefore cannot make a binding promise to perform. A promise to help rob a bank will serve as an instance of what might be called call a *wicked* promise i.e. a promise to do something that is intrinsically (and so involuntarily) impermissible.²⁰ Later on I'll confess a lack of clarity about 'intrinsic

¹⁹ And deflect it in a way that (unlike excuse) does not impugn their rationality.

²⁰ We can illustrate the same issue with much more mundane examples. Suppose the philosophy department ought to know that there is very little chance of our getting a new line but I have

impermissibility' and thus about the notion of a wicked promise but for the moment let us assume that we understand it.²¹

Presumably one is obliged not to rob banks. Can one nonetheless acquire a voluntary obligation to rob a particular bank? Specifically mightn't one have a promissory obligation to rob this bank, a promise one is (involuntarily) obliged not to fulfil?²² To settle this question, we should look for the marks of obligation noted earlier. Could it make sense for me to drive the getaway car simply because I had freely promised to do so, had not been released etc.? And could it make sense for me to exclude consideration of, say, personal risk from my deliberations about whether to keep my promise? On the face of it, the answer is 'Yes': I solemnly bound myself to an immoral enterprise and so I feel that I must follow through, its immorality notwithstanding. I may well be wrong about what I ought to do but I'm neither incoherent nor robotic. Considerations other than the promise itself may count in favour of keeping it. Perhaps I know that without a good getaway driver, my cronies are liable to be caught up in a bloody battle and this prospect is by itself enough to get me to participate in the bank-raid. But the thought of such further consequences is not needed to make sense of my fulfilling the promise and in any case my sense of responsibility for them turns largely on the fact they would be caused by my breach of promise.

The other mark of obligation may also be present. If I don't show up for the raid, it might well apt for me to feel guilty about letting my cronies down. True a decent person would be unlikely to focus their emotional attention on the wrong of breach rather than on the wrong avoided by breach but one simply indifferent to the fact that they made the promise, and to the reactions that breach of it provokes, would not be admired. A certain emotional ambivalence is usually called for: remorse about making and then breaking the arrangement as well as

promised the Chair (who wishes to keep up morale) that I won't tell anyone. Can I be bound by this promise?

²¹ I've already noted one reason for treating this notion with caution, namely the ambiguity of 'permissibility' (see note 6). In this section I'll be expressing reservations about the idea of *intrinsic* impermissibility.

²² Here it is unlikely that one's promise will justify one's fulfilling it but that is only because the interests the involuntary obligation protects are so weighty. Presumably I also have an involuntary obligation not to shove you aside as I rush to the bus stop but a sufficiently serious promise might well justify such behavior.

relief that you didn't follow through on it. Again it might be thought that one's guilt must be focused on the adverse consequences of breach but this need not be so (as when there are no consequences²³) and where one does pay attention to consequences, you impute them to yourself largely because you promised. Those who must disappoint other people's expectations feel much better about doing so where these expectations were not aroused by a promise.

While insisting that a wicked promise does not bind, recent natural right theorists tend to concede that the promisor often behaves badly in breaching a wicked promise and that this badness need not be a function solely of the further consequences of breach. True they must insist that guilt at breach is out of place whenever it is my scruples that lead me to back out of the bank raid but if I fail to show up and drive the getaway car simply because I get cold feet and bottle out, guilt at breach remains apt. One might wonder whether shame rather than guilt would be the right reaction to a display of cowardice that supposedly wrongs no one. Be that as it may, this suggestion won't work. Suppose my cronies come to realise that I'm a bit of a coward and so not a very good prospect as a getaway driver. Once I show signs of backing out, they decide that I'm not worth the trouble and release me from my promise before I breach. That should be a relief to me: I've avoided a guilt-worthy act. Yet this act can't be the supposed wrong, already committed, of manifesting my cowardice; it must instead be the wrong of manifesting it by actually breaking faith with them. But why would that act have been a further wrong unless I was bound by my promise?

Shiffrin suggests that one who makes a wicked promise instead commits a wrong of misrepresentation: they represent themselves as hereby taking on an obligation that they could not possibly be taking on. Note this wrong is not meant to depend on my cronies being misled for they may happen to agree with Shiffrin that I can't bind myself to rob a bank and so are not deceived. Still, Shiffrin maintains, there is something wrong about my so representing myself. This view raises various questions: what sort of wrong is this (a wrong of insincerity?) and who (if any one) is wronged by it? But we need not settle these issues in order to

²³ Foot concedes that were there no adverse effects, we might still feel 'a distress that we thought rational' at the breach of a solemn promise (Foot 2002: 47).

see that Shiffrin's proposal can't resolve the present difficulty. Suppose (once more) that having made my dishonest promise to rob the bank, I am released from it before I have the chance to breach. If Shiffrin is right, I must still feel bad about having misrepresented my moral powers but the situation has changed to my advantage. My culpability is reduced and my guilt assuaged because there is a further wrong that I can no longer commit, namely the wrong of breach, a wrong distinct from any wrong of misrepresentation. How could this be so unless my promise bound me?

Thus far I have been taking for granted that we can spot a wicked promise when we see one but in fact the notion of an intrinsically impermissible act needs elucidation. Emptying a bank vault is not intrinsically impermissible, nor is sticking a knife into someone; the promises of security truck companies and surgeons often bind. Perhaps 'theft' and 'assault' are intrinsically impermissible but such notions are not needed to capture the intended content of many a wicked promise. Should you promise to 'rob a bank' for me, I wouldn't normally complain if it turned out that the bank's money was somehow free for the taking. On the contrary, I'd be rather pleased that you managed to fulfil your promise by removing the money from the bank without having to rob it. Only if the theft were part of an anarchist project would I require you to *rob* the bank. I assume the natural right theorist thinks of ordinary criminal conspiracies (and not just anarchist plots) as involving wicked promises. So how exactly are they going to pick out the class of promises that are meant to be invalid simply in virtue of their content?

The natural right theorist must weaken the notion of a wicked promise so as to include promises to do what is only contingently impermissible but without deeming invalid all promises to do the contingently impermissible (otherwise conflicting promissory obligations would also be ruled out). In an effort to invalidate a normal criminal promise to rob a bank whilst also allowing for conflict of voluntary obligation, it might be maintained that a promise is invalid whenever it is envisaged that the fulfilment of that promise will wrong someone. Envisaged by whom? It can't be the promisor doing the envisaging, otherwise it would be impossible for someone to knowingly make conflicting promises that

both bind in the way described in the last section. It can only be the promisee. This is indeed the line taken by Thomson and endorsed by Watson but it is hard to see how switching our focus from the promisor's faults to the promisee's will help.²⁴

Both Thomson and Watson assume, along with almost all natural right theorists, that a binding promise (like a gift) must be accepted by the promisee.²⁵ Thomson then suggests that where a wicked promise is invalid, this is because the promisee was 'at fault' for accepting it. On this basis, Thomson maintains that a promise to participate in a terrorist plot is not binding, yet Thomson also allows that if one promises to lend someone a book and later discovers that it is owned by a third party, the promise still binds even though it could be fulfilled only by stealing the book. A likely difference between these cases is that the addressee of the book loan promise might not have known that the promisor could fulfill it only by wronging someone, whilst one who accepts a confederate's promise to participate in a terrorist plot could hardly fail to realise that its execution may involve killing people. Hence, Thomson concludes, the acceptance of the latter promise but not of the former is illicit.

But is the promisee really 'at fault' for accepting a promise simply because they know that the fulfillment of this promise will likely involve wronging someone? Working out how to keep the promise is usually the promisor's business. Suppose A needs the loan of C's book a great deal more than B does. A asks C for the book, C demurs and soon after agrees to lend it to the more charming B instead. Can't A go to C before C hands over the book and seek a promise that the book will be loaned to them after all, even though they know such a promise can be fulfilled only if C wrongs B by breaching their previous promise? Can't A take the view that C ought not to have promised the book to B in the first place and so

²⁴ (Thomson 1990: 314-5) and (Watson 2009: 176-7).

²⁵ Grotius, Hobbes and Pufendorf all require acceptance. See also (Thomson 1990: 302) and (Watson 2009: 156-7). (Shiffrin 2008: 491-3) is the exception. She does not consider Thomson's proposal and (I assume) would not endorse it. It is unclear to me how Shiffrin would distinguish promises to do something 'inherently immoral' from promises one can't fulfill without wronging someone but which are nevertheless valid (Shiffrin 2011: 164-5).

the ensuing snafu is entirely C's responsibility?²⁶ It certainly looks as though any promise C makes to A on this point would be valid, even though A knows perfectly well that it can be fulfilled only by wronging B. Were it said that this is only because the promised act is not 'intrinsically impermissible', we should ask once more what that involves beyond foreseeable wrongness.

It is often unclear how a promise is going to be fulfilled but that fact alone need not prevent us from accepting it. Suppose you volunteer to be my best man even though, as we both know, you will be spending the year in Australia. I wonder aloud how you will manage the expense of travel and your other commitments to work and family. Nevertheless you insist and I accept your promise. Have I done wrong? Must I have good grounds for believing that you have secretly worked out some way of reconciling all these conflicting demands on your time and money? When you later try to back out, I might be unsurprised but still annoyed. Perhaps I am obliged to release you from the promise, perhaps not; either way, release is needed otherwise the promise binds. You can't get out of it simply by claiming that I shouldn't have accepted your offer in the first place.

It looks as if the natural right theory fails on its own terms in that it draws no plausible line between the wicked promises supposedly invalidated by the immorality of their content and those that are not. This failure suggests that there may be no such line, that the sheer impermissibility of what you are promising to do never by itself invalidates your promise. But drawing that conclusion would be a little hasty: the fault may instead lie with the natural right theory itself. Will it help to change our intellectual framework? In the next section I'll consider how the issues raised by 'wicked' and conflicting promising look against the background of a rather different theoretical approach, though one that, like the natural right theory, treats promissory obligation as the product of an exercise of normative power.

²⁶ Here C may or may not have wronged A by not lending them the book in the first place. As we saw in the previous section, that sort of thing need not affect the validity of C's promise to B.

4. Social Practice Theories

On the natural right theory, we have the power to promise because we have the natural right to control our own lives and the ability to bind ourselves is as natural as the right of self-control from which it derives: neither the right nor the power depends for its existence on whether it happens to be recognized or respected, on whether it is embodied in an institution like the law or in some more informal social practice. According to an opposing tradition, promissory obligation is conventional: it exists only where the power to promise is recognised and its exercise treated as binding the promisor. I shall not go into what recognition of a power to promise involves except to rehearse our three marks of obligation. Where there is a social practice of promising, agents will tend to think that it makes sense to do something simply because they have promised to do it, to exclude certain considerations from their deliberations about whether to do it and will tend to feel guilt apt if they don't. Furthermore both promisees and by-standers will think it right to blame them for breach.

How is the establishment and maintenance of a practice of promising to be justified? Given that its rules bind us only when justified, this amounts to asking: under what circumstances do the rules of such a practice actually bind us? One answer goes as follows: the practice is to be justified by the interests that it serves.²⁷ That general formula leaves much scope for variation. On the most familiar way of filling it out, the existence of a practice of taking promises seriously has various desirable consequences. For example, where people are inclined to keep the promises they make, to blame those who breach and so forth, society has fashioned an extremely useful device of social co-ordination and it is the benefits of having such a device that ensure that those who make promises are obliged to keep them (Hume 1978: Book 3, Part 2). Approaching the issue in this way, how should we think about both conflicting promises and wicked promises?

²⁷ There has been a vigorous debate over what more is required to ensure that we are bound to comply with the rules of a desirable practice: perhaps the benefits and burdens created by compliance with its rules must be fairly distributed amongst the affected parties, perhaps the parties must also agree to shoulder its burdens or at least do nothing to avoid receiving its benefits. For further discussion, see (Rawls 1999: 93-99 and 301-8), (Nozick 1974: 90-5) and (Scanlon 1998: 338-42).

Can conflicting promises bind? On the present approach that looks like an issue that cannot be resolved *a priori* for it all depends on the costs and benefits of various normative systems. One might take the line that when finding oneself 'committed' to being in two places at the same time, one should simply weigh up all the relevant interests and do what seems best without remorse. On this view, the significance of obligation lies purely in its role in determining what one ought to do and any 'conflict of obligation' may be resolved by sound deliberation. But we need to acknowledge a broader range of human interests. People don't just have an interest in what will be done but also (a) in what it makes sense to do (b) in how it makes sense to conduct deliberation and (c) in whether blame, guilt etc. are apt reactions to what is done or to how deliberation is conducted. And, I would argue, the point of promising (and other normative powers) is to give them some control over these things.

Our interest in (a), (b) and (c) might still be purely instrumental. For example, it might be that guilt feelings even in people who justifiably broke their promises are apt only where they make it more likely that those people will keep other promises that ought to be kept. And tendencies to exclude otherwise relevant considerations from deliberation might be justified in a similar fashion (i.e. as maximizing the probability that you will conform with reasons across the whole range of cases). On this basis we can argue that by recognizing genuine conflicts of obligation (i.e. conflicts that cannot be resolved simply by settling what you ought to do) a system of promissory obligations serves our interest in social co-ordination. Again it is a contingent matter which set of rules best promotes social co-ordination (or the other interests at stake) and therefore a contingent matter whether a binding system of rules should recognise conflicts of obligation and the same applies to cases where only one of the conflicting obligations is the product of a promise. The logic of obligation will not settle these questions for us.

28

²⁸ On the last point I part company with Gilbert and Searle who affirm the validity of all these promises *a priori*, whilst allowing that the question of whether you ought to keep them remains open (Searle 2001: 193-4) and (Gilbert 2006: 143-64 and 227-34). Neither Gilbert nor Searle appeal to the distinctive interests that promising serves.

We can illustrate the approach to conflict of obligation that I just associated with Social Practice Theories by considering 'wicked commands'. These raise issues analogous to those raised by so called wicked promises but they arise in a context where the conventional nature of the obligations involved is much less controversial. Few people think that political or military authorities have a natural power of command. Some maintain that the basis of their authority lies in promise or consent but, on the currently most popular view, their authority derives from the desirable consequences of granting them the relevant power of command. This authority might also serve the commander's personal interests but the function of the relevant social structure is not to serve their personal interests.

When a social authority (say a military officer) issues a command, they seek to impose an obligation on their subordinates and they do so precisely by declaring their intention of hereby imposing this obligation. Given that, it is natural to compare promises which require the violation of some (other) obligation with commands which require the violation of some (other) obligation and to ask whether our authority to make and accept promises is limited in whatever way an authority's ability to issue binding orders is limited. With any social authority, we must settle the scope of its authority. On the one hand, a military officer is entitled to order their troops into battle. If they issue such an order, and the troops fail to obey, the troops are violating an obligation. The order may be foolish enough that the troops are justified in disobeying (all things considered) but even so they often remain under an obligation to obey. Given this it would make sense for them to obey, to exclude consideration of their personal safety from their deliberations about whether to obey and to feel guilty about not obeying. On the other hand, should an officer order one of their soldiers to marry a particular person or to sell their (private) car, this order is simply *ultra vires* and void *ab initio*.

Mightn't 'wicked orders' and 'wicked promises' be invalid because *ultra vires*?²⁹
The point of recognizing powers of command within a military hierarchy is not

²⁹ (Altham 1985: 4-7) assumes that wicked commands do not bind and suggests that wicked promises are equally invalid.

to serve the interests of those individuals to whom such powers are assigned. Rather the organization as a whole has (we suppose) a worthwhile social purpose, a purpose best discharged if certain individuals are given the power to issue instructions to others. Here it is important that these instructions be sources of obligation, excluding various considerations recommending disobedience from the deliberations of the subordinate. The conscientious soldier will not treat an officer's command as one factor to be weighing against all the other costs and benefits of the proposed action in determining what they should do. Rather they will exclude a range of factors (including for instance their personal safety) from their deliberations about whether to obey and often will not deliberate at all, at least about *whether* rather than *how* to obey. This all seems sensible where the organization works best if the subordinate does not second guess the superior's deliberations but simply acts as if the superior had already correctly weighed all the relevant practical considerations (Raz 1999: 40-5, 62-5).

On this model of military discipline, it is clear why certain orders will be simply *ultra vires*. The needs of the military are not served by giving an officer authority over whom their soldiers marry or over what car they own. On these matters the soldiers should act in accordance with their own view of all the relevant considerations. But I doubt that a general invalidation of orders to commit 'impermissible' acts can be supported in the same way. The realm of the morally dubious (like that of the tactically inadvisable) is broad and includes many matters crucial to the effective functioning of an army. Suppose, for example, that an officer orders a soldier to commandeer a private vehicle for military purposes. A reasonable soldier might well doubt whether he ought to do this but, as it seems to me, a reasonable soldier might also set these doubts aside simply because they have been ordered to do it by someone in authority over them.

There are limits to the scope of the exclusion here: when ordered to kill a large number of innocent civilians, the soldier must take a view about whether such killing can be justified – no reasonable soldier could regard this matter as none of their concern. (Perhaps there are also cases – the Charge of the Light Brigade? – where tactical idiocy invalidates an order.) The limits of the exclusion are set

by the legitimate purposes of the organisation and will vary from social context to social context. Doubtless the interests which ensure that an act breaches some obligation will on occasion also suffice to ensure the invalidation of any instruction to commit that act. But no viable military organisation could countenance the invalidation of an order just because a soldier reasonably thinks that it requires them to breach another obligation, even if they happen to be right. Some social organisations (competitive sports perhaps) can serve their function without requiring you to do such things but an army is not one of them: the scope of arguable 'impermissibility' is just too broad.

Back to promises. Promisors and promisees make and accept promises on a huge range of matters – there aren't the restricted spheres of competence one finds in a special purpose social hierarchy like the army. According to most Social Practice Theories, a practice of promising is to be justified by the beneficial consequences of the existence of the practice and, in particular, the role of the practice as an aid to social co-ordination. That promises have such a generic function would account for the absence of any sphere of competence restricting the making of promises. I would argue that people have another generic interest, an interest in being able to control the rights and obligations of themselves and of those around them and that they value such control for its own sake. In particular promisees have an interest in being able to determine by declaration whether the promisor is obliged to keep their promise. This normative control (or *authority*) interest is distinct from our interest in achieving social co-ordination (Owens 2012: Chapter 6). A practice of promising serves the authority interest *directly* by giving people the power to make and accept promises and so, on this view, the value of the whole practice is no longer purely instrumental.

I maintain that it is this authority interest rather than an interest in social co-ordination that fixes the rules of the practice of promising and in particular determines the status of conflicting promises and of 'wicked promises'. The validity of these promises will thus depend on whether our interest in being able to bind others to performance extends to cases where they are already obliged to do the opposite. To settle this question we need not speculate about the wider

social effects of acknowledging conflicting obligations, rather we should reflect on what we value about (promissory) obligation and control over it. I leave it open what such an inquiry might yield.³⁰ It may be that what some obligations require us to do is so abhorrent that there can be no (non-instrumental) value in having the power to create such an obligation. On the other hand, supposing promisees have a general interest in being able to bind others to performance, it is hard to see why that shouldn't extend to *some* cases where the promisor is already obliged to do the opposite (consider the book loan example from the last section for instance). If so there can be no edict of invalidation covering all such promises.

We need not choose between the various accounts of promissory obligation considered in this paper in order to reach some general conclusions about promises and conflicts of obligation. In the first section, we noted indications that people are bound by promises to do things that they are obliged (either voluntarily or involuntarily) not to do. Emboldened by some theory of promise, one might respond that people are wrong to feel guilty about breach of such a promise or confused in thinking that it makes sense to keep it. But the application of several accounts of promissory obligation has failed to overturn these indications, has failed to provide a clear theoretical rationale for invalidating all such promises. Given that, we should take these indications at face value and allow that promises to do what we are obliged not to do can bind.³¹

David Owens

Kings College London

³⁰ For some thoughts on this, see (Owens 2012: 249).

³¹ Many thanks to audiences at Princeton University, Georgetown University and Cambridge University and to Daniel Viehoff for comments on earlier versions of this paper.

Bibliography

- Altham, J. (1985) – ‘Wicked Promises’ in Ed. Hacking – *Exercises in Analysis* (Cambridge: Cambridge University Press).
- Foot, P. (2002) – *Moral Dilemmas* (Oxford: Oxford University Press).
- Gilbert, M (2006), *A Theory of Political Obligation* (Oxford: Oxford University Press).
- Grotius, H. (2005), *The Rights of War and Peace* (Indianapolis: Liberty Fund).
- Hart, H. (1955) – ‘Are There Any Natural Rights?’ *Philosophical Review* Vol. 64 No. 2 pp. 175-191.
- Hobbes, T. (1994) – *Leviathan* (Indianapolis: Hackett).
- Hume, D. (1978), *Treatise on Human Nature* (Oxford: Oxford University Press).
- Locke, J. (1988) *Two Treatises on Civil Government* (Cambridge: Cambridge University Press).
- Nozick, R. (1974) – *Anarchy, State and Utopia* (Oxford: Blackwell).
- Owens, D. (2006) – ‘A Simple Theory of Promising’ *Philosophical Review* 115: 51-76.
- (2012) – *Shaping the Normative Landscape* (Oxford: Oxford University Press).
- (2014) – ‘Does a Promise Transfer a Right?’, Eds. G. Letsas, P. Saprai and G. Klass – *Philosophical Foundations of Contract Law* (Oxford: Oxford University Press) pp. 78-95.
- Pink, T. (2009), ‘Promising and Obligation’, *Philosophical Perspectives* 23: 389–420.
- Pufendorf, S. (2003), *The Whole Duty of Man* (Indianapolis: Liberty Fund).
- Rawls, J. (1999) – *A Theory of Justice* (Cambridge: Harvard University Press).

- Raz, J. (1977) – ‘Promises and Obligations’ in Hacker. P. and Raz J. (Ed.) – *Law Morality and Society* (Oxford: Oxford University Press).
- (1999) – *Practical Reason and Norms* (Oxford: Oxford University Press).
- Reinach, A. (1983) – *The A Priori Foundations of Civil Law in Alethia 3: Philosophy of Law* (Irving, Texas: The International Academy of Philosophy Press).
- Rosati, C. (2011) – ‘The Importance of Self-Promises’ in Ed. H. Sheinman – *Promises and Agreements* (Oxford: Oxford University Press).
- Scanlon, T. (1998), *What We Owe To Each Other* (Cambridge: Harvard University Press).
- Searle J. (2001) – *Rationality in Action* (Cambridge: MIT Press).
- Shiffrin, S. (2008), ‘Promising, Intimate Relationships and Conventionalism’ *Philosophical Review* 117: 481–524.
- (2011), ‘Immoral, Conflicting and Redundant Promises’, in Ed. J. Wallace, R. Kumar, and S. Freedman - *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon* (Oxford: Oxford University Press) pp. 155–78.
- Sidgwick, H. (1981) – *Methods of Ethics* (Indianapolis: Hackett)
- Thomson J. (1990), *The Realm of Rights* (Cambridge: Harvard University Press).
- Watson, G. (2009), ‘Promises, Reasons and Normative Powers’, in eds. D. Sobel and S. Wall, *Reasons for Action* (Cambridge: Cambridge University Press): 155–78.
- Williams, B. (1973) – *Problems of the Self* (Cambridge: Harvard University Press).
- Wolf, S. (2015) – *The Varieties of Value* (Oxford: Oxford University Press).