

## CHAPTER 15

---

# CONSEQUENTIALISM AND PROMISES

---

ALIDA LIBERMAN

---

## 1. INTRODUCTION: FRAMING THE DEBATE

---

It is widely assumed by both philosophers and ordinary folks that we are generally morally obligated to keep our promises. A perennial worry about consequentialist moral theories is that they cannot account for the seemingly strict obligation to keep our promises in a wide range of cases, including those in which better overall consequences would result from promise-breaking than from promise-keeping. Can consequentialism adequately explain and accommodate the moral force of promissory obligation? And what impact does the answer to this question have on the plausibility of consequentialism as a moral theory?

Consequentialists (and their critics) will obviously care about the answers to these questions. But other ethicists have at least two reasons to be interested in this debate as well. First, assessing the relationship between consequentialism and promises sheds light on our understanding of promissory obligation more broadly. And second, it informs what approach we should take to moral theorizing from imagined cases.

In this essay, I argue that the accommodation of promissory obligation raises serious challenges for consequentialist views of various kinds, but that this is simply one strike against them, which does not by itself entail that the views are implausible. In section 2, I explain straightforward act consequentialist accounts of promising. I discuss three important challenges for these accounts in section 3. I introduce and assess alternative act consequentialist theories in section 4. I outline rule consequentialist theories in section 5 and assess them in section 6, before closing with a brief discussion of the upshots of this debate in section 7.

## 2. STRAIGHTFORWARD ACT CONSEQUENTIALIST ACCOUNTS

---

Straightforward act consequentialists determine the morality of actions by assessing whether they result in the best overall consequences.<sup>1</sup> On such accounts, promissory obligation stems from the fact that promise-making raises the promisee's expectations that the promisor will act as she has pledged to.<sup>2</sup> On these views, we are morally obligated to keep our promises because, and only to the extent that, fulfilling these expectations leads to better results than does failing to fulfill them. There are a number of ways in which promise-keeping can increase happiness or well-being. Most obviously, the promisee generally wants the promised action to occur and is made better off if the promise is kept than if it is broken. This is especially so if the promisee has (perhaps detrimentally) relied on the promise in her planning—say, if A has promised to drive B to the airport and B has failed to make alternate transportation arrangements. As Mill puts it, “few hurts which human beings can sustain are greater, and none wound more, than when that on which they habitually and with full assurance relied, fails them in the hour of need; and few wrongs are greater than this mere withholding of good; none excite more resentment, either in the person suffering, or in a sympathising spectator” (1863, 60–61).

Act consequentialists point out that, in many cases, *promising* to perform a desired action has greater utility than does simply stating your intention to perform it, or performing it without first articulating any plans to do so. This is in part because promising makes performance more likely than it would otherwise be, since promisors risk social sanctions if they do not follow through. Hume takes this observation a step farther and argues that promise-making is essential for attaining mutually beneficial reciprocal exchanges with others—such as your helping me with my harvest this week in exchange for my helping you with yours next week—that would not be possible without a robust convention of promising.<sup>3</sup> Individual acts of promise-keeping support and maintain

<sup>1</sup> There are also straightforward consequentialist assessments of whether rules or practices lead to the best overall consequences, including rules about promise-keeping. (This differs from rule consequentialism, which directly assesses rules in terms of their consequences and then determines the morality of actions by assessing whether they adhere to utility-maximizing rules.) For ease of discussion, I will focus on straightforward consequentialist assessment of acts.

<sup>2</sup> For example, see Bentham (1840, 444–445), Mill (1863, chap. 5), Sidgwick (1874, book 3, chap. 6), Pickard-Cambridge (1932), Narveson (1971), and Singer (1972).

<sup>3</sup> Hume argues that our natural selfishness and limited inclinations toward gratitude and generosity entail that we will be inclined to refrain from holding up our end of the bargain once we have received the benefit. Predicting this, others will not be likely to make bargains with us, and everyone will be worse off than they would be had we cooperated. We can avoid this problem only by creating and maintaining a social practice of promising. Because promise-breakers subject themselves “never being trusted again in case of failure” (*Treatise* 3.2.5.10, SBN 522), enlightened self-interest motivates people to keep their promises to ensure ongoing mutually beneficial exchange. Unlike most utilitarian accounts of promising, Hume's view is generally classified as conventionalist rather than expectationalist, which means that he takes the source of promissory obligation to be not the utility of satisfying the expectation, but rather the widely accepted and mutually beneficial convention itself.

this convention, and thereby increase everyone's well-being. The formation of expectations can also be valuable in itself, insofar as this eases the promisee's mind about whether the action will occur. As Jan Narveson notes, the promisee receives increased utility not just from the performance of the promised action but also from "being able to look forward to it, to plan on it, to adjust your activities in such a way as to make them harmonize with the to-be-enjoyed activity" (1971, 216).

Act consequentialists argue that grounding promissory obligation in expectations gives us a good explanation (and, in some cases, a better explanation than deontological accounts can offer) of some of the core features of our promising practice. Although P. S. Atiyah does not endorse an act utilitarian account of promising, in discussing such accounts he argues that the utilitarian value of satisfying expectations can explain all of the following: (1) why promises must be communicated to be binding; (2) why promises must be accepted to be binding; (3) why promises vary in strength on the basis of how robust the raised expectations are;<sup>4</sup> (4) why promises should sometimes not be kept (e.g., to avoid moral disaster); and (5) why promises stemming from fraud, coercion, misunderstanding, and the like do not create moral obligations (1981, 45–48). Failing to respect any of these constraints would make our promises less useful or beneficial than they could otherwise be.

### 3. CRITICISMS OF STRAIGHTFORWARD ACT CONSEQUENTIALIST ACCOUNTS

#### 3.1. Counterexample Cases of Inappropriate Promise-Breaking

An important desideratum for a theory of promissory obligation is that it captures the right set of cases. The most central and enduring objection to act consequentialist accounts of promising is that they entail that agents are morally permitted or even obligated to break their promises in cases in which it intuitively seems that the promises should be kept. An influential version of this criticism comes from W. D. Ross (1930, 34–39). Ross invites us to imagine a promise whose keeping creates 1,000 units of good for A (to whom the promise was made) and whose breaking creates 1,001 units of good for B.

<sup>4</sup> Walter Sinnott-Armstrong argues for consequentialism by highlighting how consequentialist theories can easily explain the variable weight of promises. He offers cases in which the moral weight of the obligation to keep a promise corresponds to how harmful breach of promise would be, and he argues that "this correlation supports the hypothesis that the harms of violating it are what makes the moral obligation as strong as it is" (2009, 440). Michael Gill offers an alternate explanation, arguing that the strength of a promissory obligation varies not according to the degree of harm imposed, but according to how much the promisee values the promise being kept (2012). Nikil Mukerji offers a different reply, claiming that the strength of promissory obligation depends on the amount of harm that breach would cause the promisee in particular, rather than on the amount of general harm that would result (2014).

(As he notes, this goodness calculation will have to be nuanced enough to include not only the promisee's disappointed expectations, but also the extent to which the broken promise weakens our valuable social convention of promising or leads the promisor to more readily break promises in the future, both of which will lead to a loss of net utility.) Utilitarians (at least, of the sort Ross is criticizing) argue that it is self-evident that the right thing to do is whatever leads to the best consequences. But Ross points out that this is not true in this case: to the contrary, common-sense morality holds that a much greater difference in benefit is necessary to justify promise-breaking, and that the promise to A should be kept even though breaking it would bring about greater benefit.<sup>5</sup>

More generally, it has been frequently objected that consequentialist theories that ground promissory obligation in the value of satisfying the promisee's expectations cannot explain why we are morally obligated to keep promises even in cases in which no bad consequences would result from promise-breaking. These include what David Owens calls "bare wrongings," in which breaking the promise would not lead to harm and keeping the promise would not lead to benefit. Other examples include promises made in secret where the promisee (and others) will never know whether the promise is kept, as well as promises made to people on their deathbeds.<sup>6</sup> Ross (1939, 104–105) offers a case of this sort: B promises a dying A that he will give A's property to C. But C knows nothing about the promise, and D would make better use of the property. Utilitarianism demands that B give the property to D, but Ross claims that "this utilitarian way of considering such a case is not the way in which honest men actually would consider it" (105). Rather, "most thinking people" know "that there is a *prima facie* duty to fulfil promises, distinct from the *prima facie* duty to produce what is good" (105).<sup>7</sup>

### 3.1.1. Consequentialist Responses

Some act consequentialists happily grant the existence of cases in which promise-keeping is not required on consequentialist grounds. For example, Peter Singer argues that there is no need for an absolutist rule requiring promise-keeping; rather, "all that is necessary is that there be habits of telling the truth and keeping promises unless there is a clear disutility in doing so which outweighs the benefits of preserving the useful practices and fulfilling the expectations aroused" (1972, 102). And Jan Narveson claims that most people would not bother to keep a promise if it were clear that the promisee no

<sup>5</sup> Ross's discussion of promise-breaking in *The Right and the Good* is sometimes taken to be a direct argument against utilitarianism, but it is better understood as an argument against the self-evidence of the utilitarian position. See also Ross's defense of his intuitionist account of promise-keeping against utilitarian critiques raised W. A. Pickard-Cambridge (Ross 1939, chap. 5). Thanks to Luke Robinson for discussion of these points.

<sup>6</sup> In a paper critical of expectationalist views of promissory obligation, Páll Árdal (1976) suggests that deathbed promises are so unusual that they cannot tell us anything significant about our normal promising practices (although he goes on to note that there are utilitarian reasons to keep such promises, since a person's dying moments will be happier if she knows that deathbed promises are generally kept). However, this is not a compelling defense, since it will not cover all cases. And as Atiyah (1981) notes, secret deathbed promises remain problematic for consequentialists.

<sup>7</sup> Ross uses "prima facie" in roughly the way that contemporary philosophers use "pro tanto."

longer cared about their doing so, claiming that they “do not believe that their obligation is independent of the actual good they are doing the promisee, and actual disappointment that would be caused by their non-performance” (1971, 220).

Other act consequentialists argue that the problem is not as stark as their critics make it out to be, and that promise-keeping will be required in almost every case for purely consequentialist reasons. As Sidgwick puts it, “the importance to mankind of being able to rely on each other’s actions is so great, that in ordinary cases of absolutely definite engagements there is scarcely any advantage that can counterbalance the harm done by violating them,” although he grants that promises should be broken in certain sorts of unusual cases (*ME* book 4, chap. 3, sec. 4). Such claims doubt whether there will ever be real-life cases in which more good is done by breaking a promise, but it is obvious to all that the promise ought to be kept anyway.

Some act consequentialists attempt to explain away our contrary intuitions in other ways. For example, Alistair Norcross (2011) outlines a case in which John promises his dying grandfather that he will return a statue to the long-abandoned temple from which the grandfather stole it years ago. On his way to return the statue—which is worth little—a merchant offers John a moderate sum of money for it. Assume that it would maximize utility for John to sell the statue. Norcross grants that we are not inclined to blame John for keeping his promise, and we may be inclined to think that this is what he should do. But he offers two explanations of why this may be the case. First, it could be that “as a matter of psychological fact, it would be difficult for someone with a morally appropriate commitment to keeping promises to break one in such circumstances,” even though this would in fact be the right thing to do (2011, 232–233). While Norcross is not explicit about this, he seems to be presuming that a person of good character would be unable or unlikely to break a promise in John’s situation. Perhaps our intuition that John ought to keep his promise is being confused with the intuition that promise-keeping would exhibit virtue or would reflect well on his character.

Second, it could be that John’s general disposition to keep promises to his grandfather is utility-maximizing, such that had he “been able on this occasion to perform the objectively right action, he would have been less committed to keeping his promises. If he had been less committed to keeping his promises, perhaps he would have done less good in the long run” (231). Norcross argues that it is possible for John to remain committed to objective consequentialism (or acting in ways that actually maximize utility) while adhering to a nonconsequentialist decision procedure.<sup>8</sup> This possibility does not refute the alleged counterexamples raised by the critic; it remains the case that John’s promise-keeping is *in fact* wrong, and this is precisely what is up for dispute. However, if it were

<sup>8</sup> Norcross argues that this need not involve problematic self-deception, but grants that there remains “tension between the belief that x is morally required and a commitment to doing something other than x (such as keeping a promise)” (2011, 235). He suggests that we can ease this tension by rejecting the idea that consequentialism demands that we always maximize the good, and instead accepting scalar consequentialism, which he defends elsewhere. According to this view, in keeping the promise to his grandfather John does not do the morally *best* thing, but he has not failed to fulfill a moral requirement.

true that John would be able to successfully keep his promises in general (and thereby maximize utility) *only if* his habit of promise-keeping were so robust that it led him to keep a non-utility-maximizing promise to his grandfather, keeping this promise would be justified on utilitarian grounds. While this may be true for some cases, I am skeptical that this explanation will work across the board.

### 3.1.2. *Are Our Intuitions about Cases Reliable?*

A major challenge for thinking about this line of criticism is that the counterexample cases raised by critics (i.e., non-utility-maximizing promises that seem morally obligatory to keep) are usually extremely abstract. It is doubtful that our intuitions about such abstract cases are very reliable. But it is not easy to come up with specific cases that clearly get the postulated results.<sup>9</sup> If the additional good accrued to the promise-breaking option is very tiny, it will not be obvious that it causes more good to break it after all, and it will accordingly not be obvious that the promise should be broken on consequentialist grounds. But if the additional good is very large, it may seem obvious to even nonconsequentialists that we should break the promise. As Norcross says, “it is difficult to describe a case in detail, such that it is obvious both that breaking the promise has better consequences than keeping it and that it would be wrong to break it” (2011, 220). These concerns are exacerbated by uncertainty: it is often unclear whether breaking or keeping a promise will do the most good in a particular case. And promisors themselves—with their human tendencies toward self-interest—are generally not good judges of this. When faced with uncertainty about particular cases, the best option may be to default to rules of thumb that are more generally reliable. As Atiyah puts it, “there are so many circumstances in which it is in practice impossible to be sure what is best on the whole that rules like the rule that promises should be kept, become almost conclusive [on utilitarian grounds] over a wide spectrum of activity” (1981, 77).<sup>10</sup>

This highlights an important theoretical point: if we are going to appeal to intuitions about cases in assessing the plausibility of a moral theory, we should do our best to ensure that the cases are as concrete and realistic as possible. Otherwise, we risk generating intuitions that are not reliably tracking our genuine moral reactions to cases, but are simply reflecting our existing theoretical commitments or building in our own idiosyncratic presumptions about uncertain outcomes. I am more confident than is Norcross about whether we can construct concrete cases that generate the critic’s intuition. For example, suppose Alyssa and Brendan are strangers sitting next to each other

<sup>9</sup> Thomas Carson articulates this worry as follows: “Ross’s second example presupposes that our judgments about the values of the consequences of actions have a far greater degree of precision than they, in fact, have. I cannot think of a case in which one course of action will produce exactly 1000 units of good and an alternative course of action which involves breaking a promise will produce exactly 1001 units of good. We need concrete examples against which to test our intuitions” (2005, 145).

<sup>10</sup> This is especially so given the possibility of unexpected “chain reactions” (e.g., breaking a promise to re-pay a loan and then having bad credit for years); although “the risk [of a chain reaction] is no doubt a very small one in most circumstances . . . the consequences of such a breakdown would be so very great that it is undesirable to take such risks unless the gain manifestly outweighs the dangers” (Atiyah 1981, 79).

on an airplane that has been stranded on the tarmac for several hours. An exasperated Brendan asks Alyssa to promise him that she will write a one-star online review of the airline when she gets off of the flight. Alyssa makes the promise. Suppose that no negative consequences will result from her failing to write the review: Brendan will never find out about it and so cannot be disappointed, no one else overheard her make a promise, Alyssa is already so inclined to break casual promises that breaking yet another will not affect her future reliability, and one more negative review will not impact the airline's overall rating or bottom line. Assume also that Alyssa has no independent moral obligation to keep this promise.<sup>11</sup> Writing the review will slightly decrease net utility: while it will be easy, it will take a few minutes, and will cause Alyssa to dwell momentarily on her frustrations. It nevertheless seems to me that Alyssa has at least a weak moral obligation to write a review, simply in virtue of having promised to do so. Act consequentialist theories cannot explain the existence of a moral obligation in cases like this.

However, I grant that our intuitions about such cases might be murky: while we may be confident in our judgements about whether promises ought to be kept when much is at stake, we may not have strong reactions in cases where the stakes are small. And committed act consequentialists will likely not share my reaction to this case or others like it; one philosopher's modus ponens can be another's modus tollens. In the end, this general critical strategy—that is, appealing to counter-intuitive cases as alleged counterexamples to consequentialist theories—will inevitably remain unconvincing to those with strong pre-existing theoretical commitments.

Singer frames a similar observation in this way:

Most of the criticism [of act utilitarianism] has been inconclusive because it has consisted of the outlining of unusual situations, in which the application of act-utilitarianism is said to give results which conflict with our "ordinary moral convictions." This method of argument can never move anyone who has greater confidence in the act-utilitarian principle than in his "ordinary moral convictions." Whenever the conflict is a real one, and not merely an apparent conflict, dependent on the omission of factors which the act-utilitarian can and should take into account, the genuine act-utilitarian will be prepared to jettison his "ordinary moral convictions" rather than the principle of act-utilitarianism. (1972, 94)

Singer suggests that appealing to intuitions about cases in which a consequentialist account of promising leads to counter-intuitive results will get us nowhere, because the utilitarian will never accept these intuitions. We must also note that the critic of consequentialism will not be prepared to jettison her intuitions about cases in order to accommodate a theoretical principle that she does not antecedently find compelling—for it is

<sup>11</sup> T. M. Scanlon (1998) describes a promise that does not raise expectations (your "Profligate Pal" asks you for a loan, and you give him the money without any expectation of repayment), but we have the intuition that keeping the promise is morally obligatory anyway, because it ought to be kept for independent moral reasons (i.e., the Pal should return the money because he feels gratitude toward you for a gift disguised as a loan). Assume that such alternate explanations do not apply in my case.

likely that the act utilitarian's theoretical commitments are at least partially grounded in intuitions that are not shared by nonutilitarians (e.g., about whether there is a moral difference between doing vs. allowing harm, about the sources of intrinsic value, etc.).

### 3.1.3. *Underlying Theoretical Disputes*

Consequentialists and their critics seem to be at an impasse: nonconsequentialists offer alleged counterexamples that they find compelling, and consequentialists insist that any such cases are unproblematic (either because promise-keeping is obligatory on consequentialist grounds in that case or because the intuition that promise-keeping is obligatory in that case is misguided). One way to move forward is by defending a general philosophical account of which kinds of intuitions are a reliable starting point for moral theorizing, and then arguing that the intuitions elicited by nonconsequentialist counterexample cases (such as my airplane case) are or are not among the class of reliable ones (e.g., because they are or are not widely shared, or are self-evident, etc.). However, I think that a more productive approach to moving past this impasse requires better understanding the underlying and often unarticulated theoretical disputes between consequentialists and nonconsequentialists about the nature and stringency of promissory obligation.

First, consequentialists and nonconsequentialists fundamentally disagree about *what kind of reason* grounds the obligation to keep a promise. The consequentialist presumes more generally that all moral reasons are grounded in the beneficial consequences of actions, while critics of consequentialism disagree and suggest that distinctively promissory obligations must be grounded in something that is less contingent and more essential to the making of the promise (such as an invitation to trust in the promisor, the transfer of a normative power to exercise decision-making authority over the promised action, contractualist agreements that no reasonable person could dissent from, etc.).<sup>12</sup> This is related to the more general dispute between consequentialists and nonconsequentialists over rights. Nonconsequentialist theories of promissory obligation often presume that A's promising B to  $\Phi$  at time  $t$  involves A granting or transferring to B the right to decide whether A will  $\Phi$  at  $t$ . Under this picture, wrongfully breaking a promise involves a rights violation. On straightforward consequentialist accounts—which do not take rights seriously—breaking a promise must be wrong for a different sort of reason, a reason that does not have to do with anything that A owes to B specifically.

Second, nonconsequentialists presume that A's promising B to  $\Phi$  and B's accepting this promise generates a pro tanto moral obligation for A to  $\Phi$ . Pro tanto obligations are *strict*, or such that you are required to act on them when all else is equal, and you go wrong in some way if you do not. If you have promised me that you will meet me for lunch at Spiral Diner at noon, you *must* do so, unless you have a good excuse. Contrast this with mere reasons, which are not strict; you have a reason to eat lunch

<sup>12</sup> See Friedrich and Southwood (2009), Shiffrin (2008) and Owens (2012), and Cholbi (2002) for examples of views grounding promissory obligation in trust, the exercise of normative power, and contractualism, respectively.

at Spiral Diner because the food is delicious, but do not go wrong if you do not act on this. Pro tanto obligations are also *overridable*: sometimes all else is not equal, and the demand is outweighed or undermined by other demands. All else is not equal when excusing circumstances arise, such as needing to break a promise to satisfy a more important conflicting obligation or discovering that the promise was based on deception.<sup>13</sup>

On the surface, counterexamples to consequentialist accounts of promising are a species of a much broader objection to consequentialism: namely, that consequentialism has unintuitive implications and requires us to act against the demands of common-sense morality. However, these objections also reflect deeper theoretical disputes about whether promises generate strict pro tanto moral obligations, about whether promise-breaking involves rights violations, and about whether the moral force of promising stems exclusively from the contingent benefits of promise-keeping or (also) depends on some other noncontingent source. Adjudicating this dispute requires more than intuitive reflection about whether promise-keeping is obligatory in particular concrete or abstract cases. Rather, it requires a broader theoretical assessment of the plausibility and explanatory power of strict pro tanto moral obligation as a conceptual category, and of what the source of such obligations might be. Reflection on promissory obligation will surely play an important role in this discussion. But it cannot be the only focus.

### 3.2. Are Act Consequentialist Theories Circular?

Another strategy for criticizing act consequentialist accounts of promising involves claiming not that they lead to counterintuitive results, but that they are inherently circular (or even self-refuting).<sup>14</sup> Because this critical strategy does not appeal to intuitions about cases, it avoids the impasse encountered earlier. The basic worry is that there is a good utilitarian reason to make and keep a promise to  $\Phi$  (as opposed to simply stating an intention to  $\Phi$  and then so acting) only if this leads the promisee to form a utility-increasing expectation that the promisor will  $\Phi$ . But the promisee will form such an expectation only if she antecedently believes that the promise obligates the promisor to  $\Phi$ , thereby making it likely that the promisor will  $\Phi$ . And the promisor will be so obligated (and will accordingly be likely to  $\Phi$ ) only if the promisee expects the promisor to  $\Phi$ . But the promisee will not have this expectation unless she believes that promise-keeping

<sup>13</sup> See Liberman (2015, chap. 7) for a discussion of the excusing conditions on promises.

<sup>14</sup> D. H. Hodgson articulates a strong version of the circularity objection in chapter 2 of *Consequences of Utilitarianism* (1967), in which he argues that utilitarianism is self-refuting, because a utility-maximizing practice of promising would be unable to get off the ground in a fully utilitarian world. While Hodgson's argument was influential when it was first published, it has been subject to a number of forceful criticisms that have seriously blunted its force; for example, see Narveson (1971) and Singer (1972).

is obligatory for the promisor. And promise-keeping will not be obligatory unless the promisee forms an expectation: the theory runs into a circle.<sup>15</sup>

I do not think that circularity as such is a serious problem for the act consequentialist. For act consequentialists need not be committed to the claim that expectations *alone* explain why we ought to keep promises. Rather, they can bolster their accounts by appealing to utility-maximizing social conventions. If creating and generally adhering to a social convention of promising that enables us to form promissory expectations makes us all better off, this convention will itself be justified on utilitarian grounds. The convention gives promisees reason to expect that promisors will perform as promised, thus getting us out of the circle.

However, this solution suggests a new worry for the act consequentialist, which is closely related to the previous objection. While adhering to a promising convention that demands satisfying promisee expectations will *generally* lead to the best overall consequences, this won't always be so. For there will surely be some cases—and perhaps many—in which violating the norms of the convention and leaving promissory expectations unfulfilled leads to better results than does keeping the promise. Promisors will have to assess whether the generally valuable promising convention ought to be followed in any particular case. While the convention enables expectations to be formed and thus makes promising possible, it does not provide an act consequentialist justification for promise-keeping in particular cases; only direct appeal to the good or bad consequences of promise-keeping can do that. And direct assessment by a promisor of whether keeping a particular promise leads to the best results overall is both burdensome and potentially unreliable, biased as we all tend to be in favor of our own interests. This concern motivates some to move to rule consequentialism instead, discussed in section 5.

### 3.3. Do Act Consequentialist Accounts Prove Too Much?

Finally, we should consider an interesting and underappreciated worry raised by Michael Cholbi (2014). Cholbi argues that the most plausible act utilitarian accounts ground promissory obligation in the value of assuring the promisee that an action that the promisee takes to be valuable on independent grounds will be performed. In most cases, keeping such promises maximizes utility and is accordingly obligatory. But in many cases in which keeping a particular promise maximizes utility, *making* that promise—and thereby creating an expectation which it is then utility-maximizing to fulfill—will *itself* be utility-maximizing compared to refraining from making the promise. Cholbi offers an example of Matilda and Ned, a couple who are perfect for each other but are hesitant to make a permanent commitment due to their past romantic histories. They would both be much happier if they married each other, as their relationship would feel more secure. Marriage involves making promises to each other; it follows that they

<sup>15</sup> For examples of the circularity objection to expectationalist accounts of promising, see Prichard (1940) and Robins (1976). Kolodny and Wallace (2003) raise a circularity objection to Scanlon's (1998) contractualist expectationalist account of promising and respond by developing a hybrid expectationalist/conventionalist view.

are obligated on utilitarian grounds to make marriage promises to each other. Cholbi concludes that it follows from this that “a wide range of seemingly discretionary human choices, choices usually seen as matters of personal prudence rather than impersonal morality, become matters of moral obligation on a utilitarian view” (264).<sup>16</sup>

But this is not how we generally construe promise-making: philosophers and lay people alike widely presume that promises are optional or discretionary, and that we are generally not morally required to make them in the first place, even if we are required to keep them once made.<sup>17</sup> Utilitarianism is often accused of being overly demanding,<sup>18</sup> and Cholbi argues that this is an additional way in which utilitarianism demands too much: if the value of assurance grounds a utilitarian obligation to *keep* promises, it will frequently also ground a utilitarian obligation to *make* promises. The utilitarian defense of promise-keeping comes saddled with a commitment to the obligatoriness of promise-making.

I suspect that the situation is not as dire as Cholbi makes it out to be, because I doubt that promise-making will often successfully increase utility if the promisor is not already voluntarily inclined to make the promise. Cholbi presumes that the utilitarian value of promising stems from the promisor’s satisfying the desires of the promisee. We often desire that others act in certain ways toward us because they have voluntarily chosen to, and not because they are obligated to. For example, I might prefer that a dinner guest say nothing rather than compliment me on the meal I have served simply because she believes that this is what etiquette demands. Similarly, Ned might prefer that Matilda refrain from marrying him rather than marry him because she believes that doing so is morally obligatory. In such a case, Matilda’s utilitarian obligation to make a marriage promise is narrower than Cholbi makes it out to be: she is obligated to promise *for discretionary reasons*. Either she is already motivated by discretionary reasons, in which case the obligation is inert and can do no motivational work, or she is not already motivated by discretionary reasons, in which case making the promise would not maximize utility after all.

However, while I am not convinced that act utilitarians will be faced as many morally obligatory promises as Cholbi suggests, I do not doubt that there will be some cases in which promise-making is morally obligatory on act utilitarian grounds (because the promisee does not have the preference just described, or because the benefits of making the promise outweigh the promisee’s preference that the promise be made only if it is voluntary). Act consequentialists must accept that this is a feature of their view that

<sup>16</sup> Cholbi also argues that utilitarianism implies that we will sometimes be obligated to promise to do what is antecedently obligatory for us, but that such redundant promises are unnecessary, and perhaps a sign of an agent’s weakness. However, I am not convinced that redundant promises are problematic. See Shiffrin (2011) for a discussion of how redundant promises might make an imperfect duty more determinate, counterfactually commit a promisor to an action, or give greater weight to the promisee’s perspective in deliberation as a form of partiality.

<sup>17</sup> Cholbi does not suggest that common-sense morality entails that there is *never* a moral obligation to make promises. For example, some role relationships (such as friendship or a romantic partnership) might require that the parties at least occasionally make promises to each other to properly fulfil their roles.

<sup>18</sup> See Sobel, Chapter 11, this volume.

many will find counterintuitive.<sup>19</sup> This is the flip-side of the worry from section 3.1 about whether act consequentialist views require promise-keeping in too few cases: perhaps these views also require promise-making in too many cases. And a similar theoretical impasse over conflicting intuitions is likely to arise.

Ultimately, the three objections I have discussed in this section show us that act consequentialist theories of promissory obligation are always to some extent *revisionary* of our common presumptions about promising: they require jettisoning the presumption that promises create strict pro tanto obligations, admitting that we are at least sometimes not morally obligated to keep our promises, and accepting that promise-making is sometimes morally required. These revisions do not by themselves conclusively establish that simple act consequentialist views are misguided; maybe our common presumptions about promising are problematic instead. But the need for revisions is a cost for simple act consequentialist theories. In light of this, it is worth looking at more complex act consequentialist accounts to see if they can avoid such revisionism.

## 4. ALTERNATE ACT CONSEQUENTIALIST ACCOUNTS

---

### 4.1. Multiple Intrinsic Goods

So far, I have been assessing act consequentialist accounts that aim to maximize a single good (such as pleasure, or well-being construed more broadly). Can consequentialist views that allow for multiple intrinsic goods offer a more plausible or less revisionary explanation of promissory obligation? Philip Pettit (2018) argues that we should expand our conception of what can properly count as the consequences of an action for the purposes of weighing outcomes in a consequentialist theory. He includes goods that are constitutively part of the action rather than a causal effect of it, and he argues that the agent's motivation or disposition in acting can partially determine the value of the action. In making this argument, Pettit presumes that a plausible consequentialist theory should include multiple intrinsic goods, and he suggests in passing that promise-keeping could be one of them. A consequentialist who accepted that promise-keeping was an intrinsic good could avoid many of the counterexample cases pointed to in section 3.1. For the intrinsic value of keeping of the promise itself would often tip the scales in favor of promise-keeping even in cases in which it appeared at first glance that more good could be done by promise-breaking (although there would still be some cases in which breaking the promise was necessary for the maximization of other, weightier intrinsic goods).

<sup>19</sup> Note that this same counterintuitive result might also apply to deontological theories (such as Rossian pluralism) that require maximization of benefit whenever possible as one of their pro tanto moral rules.

consequentialist. However, Peter Railton (2018) offers a reply in which he explains why consequentialists should be hesitant to expand their list of fundamental intrinsic goods to include things like promise-keeping. Railton presumes that a fundamental feature of consequentialist theories is an explanatory structure that “account[s] for the various forms of moral appraisal, such as moral rightness and moral goodness or virtue, in terms of tendencies to realise fundamental, intrinsic non-moral good” (34).<sup>20</sup> He notes that pleasure is a fundamentally nonmoral good, as is friendship. But he argues that the value of keeping a promise is fundamentally moral; promise-keeping is good only because and to the extent that it manifests moral goods like integrity, loyalty, or gratitude. If Railton is right about this, then accepting promise-keeping as an additional intrinsic good imposes a serious theoretical cost on a consequentialist theory: it changes the basic explanatory structure of the theory.

Advocates of promise-keeping as an intrinsic good could respond in (at least) two ways. First, they could claim that something other than the reduction of the moral to the nonmoral is fundamental to the structure of consequentialist theories. For example, Pettit (1997) has argued that consequentialism essentially requires that agents *promote* values (rather than *honoring* them).<sup>21</sup> If this is the right way to understand what is distinctive about consequentialism, this lessens the theoretical cost of accepting multiple fundamentally moral intrinsic goods. Second, they could accept Railton’s assumption about the explanatory structure of consequentialism and seek to identify other, uncontroversially nonmoral goods that the value of promise-keeping bottoms out in, such as the value of building trust between strangers, or of sustaining a personal relationship, or of enabling group cooperation.<sup>22</sup> Ultimately, the core issues raised by this debate—that is, whether promise-keeping can plausibly be considered an intrinsic good, and whether the value of promise-keeping is fundamentally moral or nonmoral—are underexplored and well worth further consideration.

## 4.2. Directedness and Agent-Relative Values

Agent-neutral consequentialist accounts struggle to explain how promisees should be prioritized over others in the provision of benefits because of their status as promisees. As Atiyah puts it, “the felicific calculus simply fails to pay due regard to the identity of the parties concerned” (1981, 70). Ross takes the failure to accommodate the personal character of duty to be a core defect of utilitarian views: while the utilitarian presumes

<sup>20</sup> Stephen Darwall articulates a similar idea in this way: “Consequentialist moral theories start with a *non-moral value theory* . . . Consequentialist moral theories all agree that the moral rightness and wrongness of acts are determined by the non-moral goodness of *relevant consequences*” (2003, 27).

<sup>21</sup> To honor a value is to remain committed to realizing that value yourself (even if this leads to less of the value being instantiated in the world); to promote a value is to pursue its realization impartially in both yourself and others. Pettit argues that “the consequentialist says, first, that values determine rightness in the promotional, not the honoring way” (129).

<sup>22</sup> Thanks to Josh Crabill for this suggestion.

that “the only morally significant relation in which my neighbors stand to me is that of being possible beneficiaries of my action” (1930, 19), we must also recognize the moral significance of other relationships such as promisor and promisee, and we should not be neutral between benefitting a promisee or benefitting a third party. Could an agent-relative version of act consequentialism accommodate this insight?

Michael Smith takes this worry about agent-neutrality very seriously, noting that consequentialist accounts of promissory obligation grounded in agent-neutral values entail that “an agent’s obligations turn out not to be especially targeted on the promises that *he* makes. This is remarkable because promissory obligation, at least as ordinarily understood, *is so targeted*” (2011, 209, emphasis in original). This version of the directedness objection takes this concern a step further: not only does agent-neutral consequentialism imply that promisors should not automatically prioritize their promisees over third parties, but it also implies that promisors should not automatically prioritize keeping the promises that they themselves have made.

In response, Smith suggests turning to an agent-relative consequentialist theory.<sup>23</sup> A common argument against agent-relative theories (such as hedonistic egoism) is that they are incoherent: if you have granted that pleasure is intrinsically valuable, it is inconsistent to grant that *my* pleasure is valuable and helps determine what I ought to do without also granting that *your* pleasure is valuable and also helps determine what I ought to do. Smith argues that this objection equivocates on the meaning of “valuable,” which is fundamentally relational: my pleasure is *valuable-to-me*, while your pleasure is *valuable-to-you*. It follows that “your pleasure simply isn’t relevant to my obligations, nor is mine relevant to yours, because though reason is on the side of my desiring or aiming at my pleasure, it isn’t on the side of my desiring or aiming at yours, and vice versa” (212).

Having made room for an agent-relative theory, Smith considers how “a consequentialist who has an open mind about the possibility of there being relative values that explain the distinctive [i.e., directed or relational] nature of promissory obligation” could construct a theory (215). He starts with Scanlon’s nonconsequentialist expectationalist theory, which draws on contractualism to ground promissory obligation in the value of satisfying expectations that you have deliberately raised in others. Smith concludes that Scanlon’s view is best understood as implicitly consequentialist, “because the facts about obligation that it describes are all ultimately grounded in, and thus reduce to, facts about the relative value of agents meeting the reasonable expectations that they knowingly create” (214).

This approach offers a theory that is technically consequentialist and can avoid revisionary conclusions about the directedness of promissory obligation. But Smith’s agent-relative consequentialist defense of promising is very revisionary in other ways. Many philosophers presume that agent neutrality is a central and distinctive feature of

<sup>23</sup> See also Hurley, Chapter 2, this volume.

consequentialism.<sup>24</sup> Switching to an agent-relative account requires revising a common assumption about the structure of consequentialism, which will likely be taken by those who accept this assumption to be a serious theoretical cost. Other sophisticated versions of act consequentialism that avoid counterintuitive intuitions about the nature and stringency of promissory obligation are likely to have to make a similar trade-off: additional theoretical complexity (and arguably, implausibility, or a lack of distinctiveness compared to deontological theories) is the price to pay for adherence to common-sense moral assumptions. As Brad Hooker notes, theories that attempt to force act consequentialism to have the same results as deontological theories must generally “*postulate* so much (e.g., agent-relative value, desert, etc.) in order to squeeze everything into act-consequentialism that the resulting theory is left *explaining* very little” (2011, 214).

## 5. RULE CONSEQUENTIALIST ACCOUNTS OF PROMISING

The challenges act consequentialist theories face in explaining promissory obligation have led some to embrace rule consequentialism instead. In “Two Concepts of Rules,” John Rawls articulates a basic rule utilitarian theory of promissory obligation.<sup>25</sup> Rawls notes that keeping promises will not always maximize utility. But he argues that it does not follow from this that utilitarians should refrain from keeping such promises, because the utilitarian justification does not apply to individual promises themselves. Rather, our *practice* of promising is justified on utilitarian grounds, and the practice has its own internal rules (analogous to the rules of a game like baseball or chess) that demand that promises are generally kept. We can deliberate about how to construct this practice (e.g., what conditions should count as excuses, how we should penalize promise breakers, etc.).

<sup>24</sup> For example, McNaughton and Rawling argue that “the distinction between consequentialism and deontology is best drawn, we maintain, in terms of the distinction between the agent-relative and the agent-neutral” (1992, 836). Howard-Snyder (1992) argues that deontological theories can be agent-neutral, but insists that consequentialist theories cannot be agent-relative. Pettit argues that a key characteristic of consequentialism is that “the values which determine rightness are all neutral values, not values that have a distinctively relativized reference” (1997, 129). And Zong (2000) defends the claim that consequentialism (as a criterion of rightness, rather than as a decision procedure) must be agent-neutral with respect to both agency and values. For an argument that act consequentialism is compatible with agent-relativity, see Portmore (2001).

<sup>25</sup> Rawls does not use this terminology, which was not coined until Brandt (1959, chap. 15). In *A Theory of Justice* (1971), Rawls argues that fairness requires us to act in accordance with the just institutions and practices whose benefits we voluntarily accept. To fail to do so is to unfairly (and, ipso facto, morally problematically) free-ride on others’ participation in the practice. Promising is a just and valuable practice, and making a promise involves voluntarily participating in that practice. It follows that promise-keeping is morally obligatory, lest one unfairly free-ride on a beneficial and voluntary social practice.

But once we are participating in the practice, we cannot engage in such deliberation, because “it is a mistake to think that if the practice is justified on utilitarian grounds then the promisor must have complete liberty to use utilitarian arguments to decide whether or not to keep his promise. The practice forbids this general defense; and it is a purpose of the practice to do this” (1955, 16).

Other rule consequentialists have refined this approach.<sup>26</sup> I will focus on Brad Hooker’s (2011) version of rule consequentialism, since he has offered the most explicit, sophisticated, and detailed defense of promissory obligation in particular on rule consequentialist grounds.<sup>27</sup> According to Hooker’s version of rule consequentialism:

An act is wrong if it is forbidden by the code of rules whose internalization by the overwhelming majority of everyone everywhere in each new generation has maximum expected value in terms of well-being, with some priority for the worst off. (222)

The code of rules is fixed not by the rules people actually have internalized, but by the rules that would lead to the best consequences if people were to internalize them. Hooker proposes that such a code would include a rule that generally forbids us from breaking promises. Moreover, he claims that the practice of promising demanded by this rule would meet six theoretical desiderata that are typically thought the province of deontological theories of promising.

The first of these desiderata is most relevant for our present discussion. It states that promissory obligation is not contingent on benefits to the promisee or others.<sup>28</sup> This is in part because Hooker shares the intuition that promises create strict pro tanto obligations; he presumes that “the idea that promising creates duties to do things other than maximize good . . . and the idea that promising involves transferring a right from the promisor to the promisee are ideas that I think any moral theory must accommodate if it is to be plausible” (240). Hooker argues that a rule according to which promises are binding only when keeping them is beneficial would be nonideal, because it would induce insecurity in the promisee. Because “keeping a promise is very often disadvantageous to the promisor,” promisors are likely to weasel out of promises whenever they can; it follows that “promisees need assurance that promisors are not left with much room for talking themselves into believing that their promises aren’t binding” (250).

Hooker claims that if the promising rule permitted us to break promises whenever doing so failed to be beneficial, promisees would not have such assurance. His reasoning is not explicit here, but he seems to assume that if the promising rule contained a loophole allowing promises to be broken whenever they were not beneficial, self-interested promisors would too often exploit this loophole by convincing themselves that their

<sup>26</sup> For example, see Urmson (1953) and Brandt (1959, 1992).

<sup>27</sup> This discussion of Hooker’s view and my criticisms of it (in section 6.2) are also addressed in Liberman (2015, chap. 2).

<sup>28</sup> Hooker’s other desiderata are that promissory obligations: “(2) are self-imposed (autonomy); (3) are backward-looking in the sense that they depend on events in the past; (4) are agent-relative; (5) confer rights on particular others; and (6) give only some others (the promisees) the status of being wronged if the promise isn’t kept” (249).

promise lacked benefit and was therefore permissible to break, even when this was not the case. And the existence of such an easily exploitable loophole would so decrease promisees' confidence that their promises would be kept that our valuable practice of promising would be undermined. It follows that the code of rules whose internalization would best maximize expected well-being would not permit promise-breaking whenever doing so led to increases in benefit.

## 6. CRITICISMS OF RULE CONSEQUENTIALIST ACCOUNTS

### 6.1. Avoiding Collapse

Rule consequentialism is often accused of being subject to worries about incoherence or collapse.<sup>29</sup> If rule consequentialism is overridingly committed to maximizing expected value, but insists that we should stick to some rule even when doing so does not maximize expected value, the position seems incoherent. But if rule consequentialism maintains coherence by allowing that we should maximize expected value in any case in which doing so is against one of the established rules, then the rule will need to include so many exceptions that the view is extensionally equivalent to act consequentialism.

Hooker's version of rule consequentialism avoids the collapse worry because of his constraint that the moral rules be internalized. The costs associated with internalization of the rules are factored into the calculation of expected value. As rules become more complex or demanding, or as they multiply in number, the cost of internalizing them increases. There will come a point at which the additional cost of internalizing extremely complex/demanding/varied rules outweighs any increases in expected value afforded by such rules. So it is extremely unlikely that the rules will become so complex that they are extensionally equivalent to act consequentialism.

### 6.2. A Challenge for Hooker's Account

Hooker's argument presupposes that promisee assurance is so important for the success of the overall promising practice that it is worth protecting even at the cost of requiring promise-keeping in particular cases that do not maximize well-being. It also presupposes that promisors are so motivated to get out of their burdensome promissory obligations that they will be highly unreliable at assessing when a promise fails to benefit the

<sup>29</sup> See Lyons (1965, 182–195) for a classic articulation of this objection applied to promises in particular; see Gibbard (1965) and Goldman (1974) for general replies to Lyons's argument. See also the discussion in Hooker, Chapter 23, this volume.

promisee; we cannot give them any leeway in so assessing, lest we open the floodgates to promisors wrongly assuming their promises are of little or no benefit.

However, I worry that these presuppositions prove too much. Even understood as strict *pro tanto* obligations, promises are overridable and may be permissibly broken in a wide variety of cases, such as when keeping the promise is incompatible with satisfying a more important obligation or was based on a false claim. If promisors are highly motivated to get out of their burdensome obligations, won't they *also* be overly quick to wrongly assume that a promise was premised on false information, or that keeping it conflicts with satisfying a more important obligation? And won't this too lead to promisee insecurity and lack of assurance? The same reasoning Hooker employs to justify constraining the promising rule so that promises must be kept even when the promise is of no benefit to the promisee might also justify constraining the rule so that promises must be kept even when they conflict with a more important obligation or are grounded in a false belief.

Hooker presumes that maintaining promisee security is so important that it is worth the cost to overall expected well-being of occasionally requiring actions that do not benefit anyone in particular cases. If promisee security really is so essential to the successful practice of promising, it is at least plausible that maintaining promisee security is *also* worth the cost to overall expected well-being of occasionally acting on a promise grounded in a false belief, or of failing to satisfy more important obligations in particular cases. But this is a bad result; we don't want our promising rule to be so constrained. It seems that Hooker's view proves too much: a strict promise-keeping rule that forbids promise-breaking in the cases Hooker wants to accommodate risks also forbidding promise-breaking in cases in which there should be an excuse in place. If this is correct, then Hooker's rule consequentialism has revisionary implications in a ~~similar way that act~~ **way similar to that of** consequentialist views ~~do~~.

Hooker could potentially respond to this objection in the following way. Perhaps third parties are better than promisors are at determining when it is best to break a promise, all-things-considered. An optimific rule about when to blame promise-breakers would track with these judgments, and the threat of potential sanction would in many cases prevent promisors from too readily breaking promises in their own favor. Internalizing a permissive promising rule that allows for breach of promise in a fairly wide range of cases—*paired with* a rule requiring third-party sanction if the promisor misjudged—might be more optimific than internalizing a restrictive promising rule that forbids promise-breaking in most cases. More reflection on whether rule consequentialism can accommodate a rule requiring promise-keeping in all and only those cases demanded by common-sense morality is needed.

## 7. UPSHOTS AND FUTURE DIRECTIONS

Answering the question of whether consequentialism can accommodate promissory obligation requires more than just reflecting on our immediate intuitions about whether

promise-keeping is morally required in some hypothetical, non-utility-maximizing case. We must also consider a range of questions about the nature of promising as a moral phenomenon, including whether promises yield strict pro tanto obligations (and how these obligations should be understood); what the excusing conditions on promises are (and whether they line up with the cases in which consequentialist theories predict promise-keeping is not required); how and whether the strength of promissory obligation varies (and whether consequentialist theories give us a good explanation of this—see note 4); whether promise-making is ever morally obligatory (and if consequentialism problematically implies the wrong answer to this question); and so on. The answers to these questions can help shape our desiderata for any plausible theory of promissory obligation, be it consequentialist, deontological, or other.

Careful reflection on the relationship between consequentialism and promises also sheds light on the role that appeal to intuitions regarding (or speculation about the results of) imagined cases can and should play in the assessment of a moral theory more generally. Specifically, it illustrates the need for cases with a high degree of specificity whenever possible, especially for topics (such as assessing the results of promise-keeping or -breaking) that involve a large degree of uncertainty. It also illustrates the limits of this sort of reflection in isolation: intuitions about cases are likely to clash on the basis of existing theoretical and intuitive commitments, and they must be bolstered by appeals to independent criteria like those addressed in the previous paragraph.

Ultimately, I am skeptical that a traditional act or rule consequentialist theory will be able to accommodate all of our theoretical desiderata and pretheoretical intuitions about the robustness and breadth of promissory obligation. Most likely, there will always be some case in which the consequentialist must admit that a promise ought not be kept, despite deontologists' intuitions to the contrary. Consequentialists have a variety of resources to limit such cases, or minimize their forcefulness. But if counter-intuitive cases remain, this is a theoretical cost that consequentialism must deal with.

In light of this, consequentialists should continue to try to accommodate some version of promissory obligation to whatever extent they can, either by explaining how their views can accommodate a wider range of promissory obligation than it may first appear, or arguing that the cases that the views cannot explain are not problematic. One way forward is looking beyond straightforward utilitarian theories and seeing whether more sophisticated versions of consequentialism that are independently plausible can better accommodate promissory obligation. Such accommodation would be a point in favor of a theory, and philosophers developing such theories would do well to explicitly discuss what their theories imply about promise-keeping.

That being said, I do not think that an inability to accommodate all of our intuitions about promise-keeping is a decisive strike against consequentialism as a moral theory. For our intuitions about the strength and breadth of promissory obligation are not independently robust enough to by themselves merit complete rejection of an otherwise powerful moral theory. Rejections of utilitarianism as a moral theory on this basis move

too quickly: in the end, whether consequentialism stands or falls will have to be decided on broader grounds than this.<sup>30</sup>

## REFERENCES

- Árdal, Páll S. 1976. "Promises and Reliance." *Dialogue* 15, no. 1: 54–61.
- Atiyah, P. S. 1981. *Promises, Morals and the Law*. Oxford: Oxford University Press.
- Bentham, Jeremy. (1840). 1977. *A Comment on the Commentaries and A Fragment on Government*. Edited by J. H. Burns and H. L. A. Hart. London: Athlone Press.
- Brandt, Richard. 1959. *Ethical Theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Brandt, Richard. 1992. *Morality, Utilitarianism, and Rights*. New York: Cambridge University Press, 1992.
- Carson, Thomas L. 2005. "Ross and Utilitarianism on Promise Keeping and Lying: Self-Evidence and the Data of Ethics." *Philosophical Issues* 15, no. 1: 140–157.
- Cholbi, Michael. 2002. "A Contractualist Account of Promising." *Southern Journal of Philosophy* 40: 475–491.
- Cholbi, Michael. 2014. "A Plethora of Promises—Or None at All." *American Philosophical Quarterly* 51, no. 3: 261–272.
- Darwall, Stephen. 2003. "Theories of Ethics." In *A Companion to Applied Ethics*, edited by R. G. Frey and Christopher Heath Wellman, 17–37. Malden, MA: Blackwell.
- Friedrich, Daniel, and Southwood, Nicholas. 2009. "Promises beyond Assurance." *Philosophical Studies* 144, no. 2: 261–280.
- Gibbard, Allan F. 1965. "Rule-Utilitarianism: Merely an Illusory Alternative?" *Australasian Journal of Philosophy* 43, no. 2: 211–220.
- Gill, M. B. 2012. "The Non-Consequentialist Moral Force of Promises: A Response to Sinnott-Armstrong." *Analysis* 72, no. 3: 506–513.
- Goldman, Holly S. 1974. "David Lyons on Utilitarian Generalization." *Philosophical Studies* 26, no. 2: 77–95.
- Hooker, Brad. 2011. "Promises and Rule-Consequentialism." In *Promises and Agreements*, edited by Hanoch Sheinman, 237–254. Oxford: Oxford University Press.
- Howard-Snyder, Frances. 1994. "The Heart of Consequentialism." *Philosophical Studies* 76, no. 1: 107–129.
- Hume, David. (1741). 1978. *A Treatise of Human Nature*. 2nd ed. Edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Clarendon Press.
- Kolodny, Niko, and Wallace, Jay. 2003. "Promises and Practices Revisited." *Philosophy & Public Affairs* 31, no. 2: 119–154.
- Liberman, Alida. 2015. *The Mental States First Theory of Promising*. PhD diss., University of Southern California.
- Lyons, David. 1965. *Forms and Limits of Utilitarianism*. Oxford: Clarendon Press.
- McNaughton, David, and Rawling, Piers. 1992. "Honoring and Promoting Values." *Ethics* 102, no. 4: 835–843.
- Mill, John Stuart. (1863). 2001. *Utilitarianism*. 2nd ed. edited by George Sher. Indianapolis: Hackett.

<sup>30</sup> Thanks to Eric Barnes, Philippe Chuard, Josh Crabill, Justin Fisher, Steve Hiltz, Robert Howell, Jennifer Matey, Luke Robinson, Steve Sverdlik, and Doug Portmore for helpful discussion of this paper.

- Mukerji, Nikil. 2014. "Consequentialism, Deontology and the Morality of Promising." In *Business Ethics and Risk Management*, edited by Johanna Jauernig and Christoph Lütge, 111–126. Dordrecht: Springer.
- Narveson, Jan. 1971. "Promising, Expecting and Utility." *Canadian Journal of Philosophy* 1, no. 2: 207–233.
- Norcross, Alastair. 2011. "Act-Utilitarianism and Promissory Obligation." In *Promises and Agreements*, edited by Hanoch Sheinman, 217–236. Oxford: Oxford University Press.
- Owens, David. 2012. *Shaping the Normative Landscape*. Oxford: Oxford University Press.
- Pettit, Philip. 1997. "The Consequentialist Perspective." In *Three Methods of Ethics: A Debate*, edited by Marcia Baron, Philip Pettit, and Michael Slote, 92–174. Malden, MA: Blackwell.
- Pettit, Philip. 2018. "Three Mistakes about Doing Good (and Bad)." *Journal of Applied Philosophy* 35, no 1: 1–25.
- Pickard-Cambridge, W. A. 1932. "Two Problems about Duty (II)." *Mind* 41: 145–172.
- Portmore, Douglas W. 2001. "Can an Act-Consequentialist Theory Be Agent Relative?" *American Philosophical Quarterly* 38, no. 4: 363–377.
- Prichard, H. A. (1940) 2002. "The Obligation to Keep a Promise." In *Moral Writings*, edited by Jim MacAdam, 257–266. Oxford: Clarendon Press.
- Railton, Peter. 2018. "Toward a More Adequate Consequentialism." *Journal of Applied Philosophy* 35, no 1: 33–40.
- Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rawls, John. 2005. "Two Concepts of Rules." *Philosophical Review* 64, no. 1: 3–32.
- Robins, Michael H. 1976. "The Primacy of Promising." *Mind* 85, no. 339: 321–340.
- Ross, W. D. 1930. *The Right and the Good*. Oxford: Clarendon Press.
- Ross, W. D. 1939. *Foundations of Ethics*. Oxford: Clarendon Press.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.
- Shiffrin, Seana. 2008. "Promising, Intimate Relationships, and Conventionalism." *Philosophical Review* 117, no. 4: 481–524.
- Shiffrin, Seana. 2011. "Immoral, Conflicting and Redundant Promises." In *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, edited by R. Jay Wallace, Rahul Kumar, and Samuel Freeman, 155–178. Oxford: Oxford University Press.
- Sidgwick, Henry. 1874. *The Method of Ethics*. Chicago: Chicago University Press, 1962.
- Singer, Peter. 1972. "Is Act-Utilitarianism Self-Defeating?" *Philosophical Review* 81: 94–104.
- Sinnott-Armstrong, Walter. 2009. "How Strong Is This Obligation? An Argument for Consequentialism from Concomitant Variation." *Analysis* 69, no. 3: 438–442.
- Smart, J. J. C. 1973. "An Outline of a System of Utilitarian Ethics." In *Utilitarianism: For and Against*, edited by J. J. C. Smart and Bernard Williams, 3–76. Cambridge: Cambridge University Press.
- Smith, Michael. "The Value of Making and Keeping Promises." In *Promises and Agreements*, edited by Hanoch Sheinman, 198–216. Oxford: Oxford University Press, 2011.
- Urmson, J. O. 1953. "The Interpretation of the Moral Philosophy of J. S. Mill." *Philosophical Quarterly* 10: 33–39.
- Zong, Desheng. 2000. "Agent Neutrality Is the Exclusive Feature of Consequentialism." *Southern Journal of Philosophy* 38, no. 4: 676–693.