

NIKO KOLODNY AND
R. JAY WALLACE

Promises and Practices Revisited

Promising is clearly a social practice or convention. By uttering the formula, “I hereby promise to do X,” we can raise in others the expectation that we will in fact do X. But this succeeds only because there is a social practice that consists (inter alia) in a disposition on the part of promisers to do what they promise, and an expectation on the part of promisees that promisers will so behave. It is equally clear that, barring special circumstances of some kind, it is morally wrong for promisers to fail to do what they have promised to do. What is perhaps less clear is how the moral wrongness that is involved when promises are broken is related to the social practice that makes promising possible in the first place.

This question is at the center of T. M. Scanlon’s pioneering recent work on promising. Fundamental to the account he offers—first in “Promises and Practices” and more recently in *What We Owe to Each Other*—is his rejection of the view that the “wrong involved in breaking a promise is a wrong that depends essentially on the existence of a social practice of agreement-making.”¹ On Scanlon’s account, we can understand why it is wrong to break one’s promises without making any reference to the fact that promising is a social practice or convention.

We agree with Scanlon that the moral wrong involved in breaking promises is not simply the wrong involved in flouting the terms of a social practice. But we do not agree that the moral obligation to fulfill one’s

We are grateful to Seana Shiffrin, Samuel Scheffler, and two anonymous readers for *Philosophy & Public Affairs* for extremely helpful comments on earlier versions of this article.

1. “Promises and Practices,” *Philosophy & Public Affairs* 19 (1990): 199–226; and *What We Owe to Each Other* (Cambridge, Mass.: Harvard University Press, 1998), chap. 7. The quotation is from the latter, at p. 295. All subsequent page references to *What We Owe to Each Other* are given in the text.

promises can be understood without reference to the fact that promising is a social convention. Our aim in the present article is to defend these conclusions. We show that Scanlon's favored theory is vulnerable to a circularity objection that he himself formulates, an objection that derives from his rejection of the practice view, which Scanlon's ingenious and subtle arguments do not succeed in deflating. After developing these conclusions we sketch a hybrid account, which builds on the conventional character of promising at a crucial point. We argue that the hybrid account can avoid the circularity objection to which Scanlon's theory is vulnerable, without sacrificing his basic insight that the obligation to fulfill one's promises is not simply the obligation to abide by the terms of social practices. The result is an improved framework for understanding the distinctive way in which promising to do something generates a moral obligation of fulfillment.

I. THE PRACTICE VIEW

At its core, the social practice of promising consists in the fact that members of a certain group have, and are known by one another to have, a policy of doing X (unless certain excusing conditions obtain), whenever they have performed communicative acts of some recognized type, such as uttering the formula, "I hereby promise to do X." This social practice is an instrumental good. It is, as John Rawls puts it, "a rational means whereby men can enter into and stabilize cooperative agreements for mutual advantage."²

To arrive at such agreements, party A typically needs to assure party B that if B cooperates, A will also cooperate. In some instances, A can assure B of this by communicating to B that A already has sufficient reason to cooperate if B does his part. For example, they may be working on some joint project, which requires the contributions of both. In this case, A can assure B that because A wants the joint project to succeed, he has reason to do his part, provided B does his. Often, however, A cannot communicate to B that A has reason to cooperate, if B does his part, because it is clear that he has no such reason. Consider the example at the crux of David Hume's discussion of promising, in which farmer A wishes to persuade farmer B to help with A's harvest today in return for

2. John Rawls, *A Theory of Justice* (Cambridge, Mass.: Harvard University Press, 1971), p. 345.

A's helping with B's harvest tomorrow, but is unable to do so, because B is aware that A will have no reason to help him tomorrow, after having already reaped whatever benefits are coming to him from their agreement. What makes promising especially valuable, as Hume observed, is that it allows A to assure B that A will do his part even when he has no independent reason to do so.³ Because B knows that A, as a participant in the social practice, has a policy of fidelity—of doing X whenever A performs a communicative act of the appropriate type, such as uttering the formula, “I hereby promise to do X” —A can assure B, by uttering that formula, that A will do X. A assures B in this way, we might say, by “engaging” his standing policy of fidelity.

There is more to the social practice of promising than this, however. By uttering, “I hereby promise to do X,” a promiser not only engages a policy, but also incurs a moral obligation. It is a constitutive feature of promising, which distinguishes it from other means of agreement making, that the promiser assures the promisee that he will do X by performing a communicative act that both know morally obligates the promiser to do X. The promiser influences the attitudes of the promisee by reflexively invoking an obligation that is incurred by the very act of promising itself. A fundamental question is, What accounts for this obligation?

According to the practice view, the obligation can only be explained by invoking the social practice of promising. Hume, for example, holds that the wrongness of promise breaking is rooted in impartial disapproval of acts that flout the practice, disapproval that arises from a recognition of the instrumental value of the practice for its participants collectively.⁴ Although Hume is not entirely clear on how breaking a promise abuses the practice, and thus deserves this disapproval, there appear to be two general possibilities: breaking a promise may undermine the practice or exploit it. The social practice exists as an instrumental good only insofar as people reasonably expect that others generally adhere to a policy of fidelity. If one's failure to adhere to a policy of fidelity undermines this expectation, then one deprives others of an important instrumental good. Even when one's failure to adhere to a policy of fidelity does not undermine the practice, as is more often the case, it nevertheless exploits the practice. One free-rides on expectations generated

3. David Hume, *A Treatise of Human Nature*, ed. L. A. Selby-Bigge, P. H. Nidditch, 2nd ed. (Oxford: Oxford University Press, 1978), pp. 520–21.

4. *A Treatise of Human Nature*, bk. III, pt. II, chap. V.

by others' adherence to the policy without doing one's part to sustain those expectations oneself. Rawls's explanation of the wrongness of promise breaking, for example, focuses on the wrongness of this kind of free-riding. The duty to keep a promise, he argues, is an instance of the more general duty, under the "principle of fairness," to do one's part in a practice whose benefits one voluntarily accepts. In making a promise, one opts to avail oneself of the benefits of the practice, and thus morally obligates oneself to do what its rules require, which is to do what one promised.⁵ The wrong of breaking a promise, according to these accounts, just is the wrong of undermining or exploiting a social practice, and so cannot be explained without reference to a social practice.

These general moral prohibitions on undermining or exploiting valuable practices also attach to other practices of agreement making. Scanlon describes a nonmoral "Reindeer code," for example, which requires members of the "Sacred Brotherhood of Reindeer," by appealing to their sense of honor, to do what they swear to do in the name of the Brotherhood (p. 298).⁶ The Reindeer code, like promising, is a social practice that consists in its being common knowledge among a group of people that they each have a policy of fidelity: i.e., a standing intention to do X whenever they have engaged in a certain kind of performance. To the extent that this code is a valuable practice, the general moral prohibitions that we have discussed forbid Reindeer from violating their policy of fidelity.

What is distinctive of the social practice of promising, therefore, is not that its participants have moral obligations to adhere to its policy of fidelity. In part, promising differs from invoking the Reindeer code in that promisers signal their recognition of these moral obligations as their motivation for adhering to the policy of fidelity that is partly constitutive of the practice. Reindeer, by contrast, signal their sense of honor, not their recognition of a moral obligation not to exploit the practice of invoking the Reindeer code. But the difference goes deeper than this. Suppose that Reindeer started to signal, along with their sense of honor, their recognition of their moral obligation not to exploit the practice of invoking the Reindeer code. Would their signaling of their recognition of this moral obligation count as promising? No, because it would be the

5. *A Theory of Justice*, pp. 344–50.

6. We follow Scanlon in holding that the code is nonmoral in character. One reason for this is that its scope is restricted to members of the Brotherhood. A second reason is that it draws on different motivations: a sense of honor, rather than a desire to act only in ways that would be justifiable to the people who would be affected.

practice of invoking Reindeer honor, not the practice of promising, that Reindeer would be signaling that they have a moral obligation not to exploit. In promising, one signals one's recognition of a moral obligation not to undermine or exploit not merely *some* social practice of agreement making, but specifically the social practice of promising: the practice that consists in participants' signaling that they adhere to a policy of fidelity because they recognize moral obligations to adhere to the policy.⁷

7. Here we likely depart from earlier versions of the practice view. At times, Hume and Rawls describe the social practice of promising in ways that make it seem parallel to the revised Reindeer practice. First, there is a freestanding practice of agreement making, like the original Reindeer practice, that consists in participants' being known to adhere to its policy of fidelity out of nonmoral motivations. Second, there is a moral obligation not to undermine or exploit this practice, by appeal to the recognition of which participants can (further) assure others that they will adhere to this policy. According to Hume, the freestanding practice is sustained by self-interested concern for one's reputation as a cooperator. This freestanding practice is not the practice of promising itself; the practice of promising emerges only when people acquire and begin to communicate a sense of obligation not to undermine the freestanding practice. Rawls, by contrast, labels as "the practice of promising" the freestanding practice, which is constituted by compliance with a nonmoral requirement to do X if one has uttered (in relevant conditions), "I promise to do X." But this label seems misplaced; it would imply that one can participate in the practice of promising without signaling any awareness of a moral obligation. Rawls then distinguishes this practice-constituting, nonmoral requirement from the nonconstitutive, moral obligation, under the principle of fairness, to comply with such practice-constituting, nonmoral requirements in general.

It is unclear, however, whether the initial practices really could be freestanding: that is, sustained independently of a sense of obligation. There are familiar problems—deathbed promises and the like—with Hume's claim that promising is sustained by concern for one's reputation. Perhaps the sense of obligation is supposed not only to reinforce participants' initially self-interested adherence to the policy of fidelity, but also to extend the policy's application to cases, such as death-bed promises, in which reputation effects are not at issue? But it is obscure how the sense of obligation can do this, if it is simply a sense of obligation not to undermine the freestanding practice, which does not apply to these cases in the first place. Rawls leaves the normative force of the nonmoral requirement that constitutes the practice unexplained. If the social practice exists, then people must reliably be motivated to comply with its requirement. What supplies this motivation? It is not enough to say that the requirement is constitutive of the practice, like the rules of chess, so that anyone who regularly fails to comply with the requirement is participating in a different practice, or no practice at all. Even if we set aside the worry that this analogy would define broken or lying promises out of existence, it fails to explain why people should care whether they participate in the practice of promising. It is clear why a grandmaster would want to count as playing chess, and thus care about abiding by its constitutive rules: he finds it an engaging pursuit in its own right, it offers him opportunities to exercise and be recognized for his skills, and so on. Yet why should anyone care about abiding by the constitutive rules of promising? If the practice *were* already established, one might want to *appear* to follow its rules, in order to convince others that one will do as one says. But this is not the same thing as wanting to follow the rules themselves, and motivations of this kind would be insufficient to establish a practice constituted by those rules.

There may seem to be an air of bootstrapping about this. In most cases, I am morally obligated to continue to adhere to the policy of fidelity only if my ceasing to adhere to it would exploit the social practice of promising. This social practice exists to be exploited only if enough others continue to adhere to the policy of fidelity. If the foregoing is correct, then other participants adhere to the policy because of their awareness of moral obligations of the same kind as mine: namely, obligations not to exploit the social practice of promising.⁸ In other words, I know that others will continue to adhere to the policy of fidelity only if each of them believes that he is obligated to continue to adhere to it, and each of them will believe that he is so obligated only if he believes that enough others will continue to adhere to it. In sum, I am obligated to continue to adhere to the policy of fidelity only if I have reason to believe that enough others believe that they are similarly obligated to continue to adhere to it, and each of them will believe that he is obligated to continue to adhere to it only if he believes that enough others believe that they are similarly obligated to continue to adhere to it, and so on.

This system of interdependent expectations, however, is no more paradoxical than that which underlies the social institution of money. I am willing to continue to accept these intrinsically valueless pieces of paper in exchange for real goods and services only because I believe that enough others are similarly willing to continue to accept them, and each of them is willing to continue to accept them only because each of them believes that enough others are willing to continue to accept them, and so on. While there is a question about how such regimes of expectation come to be established, there is no mystery about how, once established, they perpetuate themselves. So long as nothing that might compromise the regime has come to light, one has reason to believe that it is

8. Insofar as others signal that they adhere to a policy of fidelity out of nonmoral motivations, they are not participating in the same practice. The practice of promising, to repeat, consists in participants' communicating that they adhere to a policy of fidelity out of a sense of moral obligation. It might be suggested that the appearance of bootstrapping could be avoided by supposing that others are deceptive: that they signal that they adhere to the policy out of moral motivations, thereby counting as participants in the practice of promising, when in fact they adhere to the policy out of other motivations. We can leave aside the questions whether a practice marked by systematic deception of this kind could count as the practice of promising, or whether it could even be sustained. It would be enough of an objection to the practice view that it had to assume such deception in order to avoid the bootstrapping problem.

still in place and hence to act accordingly. In the case of promising, so long as nothing suggests that others have suddenly come to doubt their obligations to adhere to the policy of fidelity, one has reason to believe that they will continue to adhere to it, and hence one is obligated to continue to adhere to it oneself.

II. SCANLON'S CRITICISM OF THE PRACTICE VIEW

Scanlon does not doubt that there is a social practice of promising (p. 296). Nor does he deny that practice-involving considerations, such as those that Hume and Rawls offer, provide "*a moral reason*" for keeping promises (p. 316). He rejects the practice account instead because he perceives, accurately in our view, that these considerations fail to capture what is most fundamentally wrong about breaking promises.

This failure manifests itself in two ways. First, one can lead someone to rely on one's doing something without making a promise or invoking any social practice. What makes one's failure to perform wrong in these cases seems the same as what makes one's failure to perform wrong in cases in which one promises. In Scanlon's "state of nature" example, two hunters find themselves on opposite banks of a river, the first with the second's errant boomerang, the second with the first's errant spear. The first gets the second to believe that he will return the second's boomerang, if the second returns his spear. The wrong involved in failing to return the boomerang, after the second has performed, seems identical with the wrong that would have been involved had the first promised to return it (p. 297). Whatever makes promise breaking wrong also makes these other failures to perform wrong and is not to be explained by reference to social practices.

Second, according to the practice view, the obligation to keep promises is owed to all those who stand to benefit from, or who have contributed to, the practice as a whole. In breaking a promise, one wrongs them collectively, either by depriving them of an important benefit or by free-riding on their contributions to the social practice. By contrast, we ordinarily think that the obligation to keep a promise is an obligation owed specifically to the promisee (p. 316). This belief is evident in our tendency, when we have broken a promise, to see ourselves as having reason to apologize primarily, if not exclusively, to the promisee. One might put the point in terms of Scanlon's contractualism, according to

which what makes a particular kind of conduct wrong lies in the reasons that people who would be adversely affected by it would have for rejecting principles that permit it. When we engage in such conduct, we wrong the people whom we adversely affect in the relevant way. If the practice view is correct, everyone who benefits from or has contributed to the practice would equally have reason for rejecting principles permitting promise breaking; hence, on the practice view, promise breaking wrongs them all. As Scanlon persuasively argues, however, promise breaking principally wrongs the promisee. What makes promise breaking wrong ought accordingly to lie in the reasons that potential promisees, in particular, would have for rejecting principles permitting it.

III. SCANLON'S ACCOUNT

Scanlon begins by dismissing the initially appealing suggestion that the reasons for rejection stem from the loss that potential promisees would suffer from relying on broken promises.⁹ Instead of justifying an obligation to keep promises, Scanlon observes, these reasons support the more general principle of Loss Prevention, or L:

If one has intentionally or negligently led someone to expect that one is going to follow a certain course of action, X, and one has good reason to believe that that person will suffer significant loss as a result of this expectation if one does not follow X, then one must take reasonable steps to prevent that loss (pp. 300–01).

L does not require one to do X. One can satisfy L equally (i) by doing X, (ii) by giving a timely warning that one will not X, or (iii) by compensating for losses incurred by relying on the expectation that one would do X.¹⁰ The duty to keep promises, by contrast, is not neutral between warning and

9. He credits this suggestion to Neil MacCormick, "Voluntary Obligations and Normative Powers I," *Proceedings of the Aristotelian Society*, supp. vol. 46 (1972): 59–78.

10. It might be suggested that even this disjunctive obligation is implausibly strong. In certain commercial transactions, for instance, it seems possible to signal an intention to do X in the kinds of circumstances described, without incurring an unqualified obligation of fulfillment or compensation. Perhaps you have bought my house in part because of my truthful claim that I do not intend to sell the adjacent property; but my circumstances change, and three years later I need the money, so I sell the property to someone who builds a gas station on it, without providing you any compensation. It has been maintained that

compensation, on the one hand, and performance, on the other. It is precisely a duty to fulfill the expectations that one has created in certain conditions, unless the promisee consents to one's not fulfilling them.

The reasons for rejection stem instead from what Scanlon calls the "value of assurance." He illustrates this value with the example of the Guilty Secret (pp. 302–03). Harold, an old acquaintance you have run into at a party, asks you to promise him that you will not reveal a past indiscretion of his—which you consider fairly inconsequential, but which he finds extremely embarrassing—while you are visiting at Harold's campus for a term. Whether or not you promise, revealing the secret would be wrong, for the straightforward reason that it would gratuitously hurt Harold. But if you do promise, then it is wrong for a further reason. What makes it wrong cannot be that Harold might suffer some loss by relying on the expectation that you will keep his secret. There is no way for Harold to rely on it, and thus L has no application in this case. Instead, Scanlon suggests, what makes it wrong is that you would be undermining something that Harold has good reason to want: namely, assurance that you will keep his secret. This is what would underlie Harold's complaint against your revealing his secret and, analogously, potential promisees' reasons to reject principles permitting you to do so.

Scanlon's appeal to the value of assurance seems plausible, but it is also somewhat elusive. On the one hand, the value to Harold of the assurance that you will not reveal his secret is not simply the value to him of his believing that you will not reveal his secret. Scanlon guards against the assumption that "the value of assurance is purely experiential, that it consists merely in the values of freedom from worry, increased ability to sleep at night, and so on. . . . What people in Harold's position, and in

my behavior, although perhaps regrettable, has in no way wronged you. Cf. Charles Fried, *Contract as Promise* (Cambridge, Mass.: Harvard University Press, 1981), p. 9. This conclusion may require some qualification: we are not certain, for instance, that there is no moral obligation to take "reasonable steps" to avert or compensate for the loss in at least some cases of this kind. But if the obligations specified in L do not seem to apply here that is because of the special nature of commercial transactions, which provide legally enforceable mechanisms for extracting commitments from both buyers and sellers. If I fail to provide a contractual guarantee that the adjacent property will not be developed without your permission, then my representations regarding my plans for the property should be taken by you merely as honest claims about my present intentions (with an implicit rider: "but of course my plans might change"), and any losses that might stem from changes in my plans are risks that you assume in buying the property on the terms proposed. *Caveat emptor!*

many other positions, reasonably want is not mere freedom from worry; they also want certain things to happen (or, as in Harold's case, not to happen)" (p. 303). On the other hand, the value to Harold of the assurance that you will not reveal his secret is not simply the value to him of your not revealing his secret either. That value provides you with reasons not to reveal it, as we saw above, but these reasons are independent of having promised anything to him.

The value of this assurance to Harold would seem to be some combination of the value to him of his secret's being kept and the value to him of the belief that his secret will be kept. The value to him of assurance that you will keep his secret, one might say, is something like the value to him of knowledge that you will keep his secret. We might expect Harold to say, "Please don't just *tell* me what I want to hear. I need to *know* that you will not reveal this to anyone." Against this, it might be objected that one cannot, strictly speaking, be said to know that X if the event or state of affairs described by X lies in the future. If it is impossible to provide someone with foreknowledge of future events and states, then assuring Harold cannot consist in providing him with genuine foreknowledge of your future actions. But one might gesture toward the idea in this way: to take seriously the value of his being assured that you will keep his secret is to see yourself as having reason to endeavor to make his belief that you will keep his secret as close to knowledge as is possible.

It is a further question why assurance, understood in this way, should be thought valuable. What is it about the prospective knowledge that other people will act in a certain way that makes this a condition that we have reason to want? Most of the time, we believe, the answer to this question builds on the significance of assurance for human agency. The value of assurance goes beyond the value of the states of affairs that we want to be assured about, insofar as it includes a subjective component: our awareness that the valued state of affairs will in fact obtain. This typically matters to us because of its manifold effects on our capacity to plan our activities through time and to execute our plans in action. If I am assured that you will do X, then I can take that as a fixed point in my deliberations about what to do myself, in ways that greatly facilitate my ability to realize my goals and values in the conduct of my life. I can *rely* on your doing X, as Scanlon puts it (p. 303), focusing my thoughts and efforts on other matters once I obtain assurance that you are going to do X, for instance, and coordinating my activities effectively with yours.

But the case of the Guilty Secret suggests that this interpretation of the value of assurance does not apply in all cases. On Scanlon's account, Harold really can't rely on your assurance that his secret will not be divulged: there is nothing that he can concretely do about the matter one way or another. And this is connected to Scanlon's contention that there is no concrete reliance loss incurred by Harold through your divulging the secret, of the sort that would open up the possibility of compensation. This contention becomes more plausible if we change the example in a crucial point, imagining that Harold is a colleague at your home university, who taught long ago at the institution you are going to be visiting for a semester, and who asks you before you go not to divulge to anyone there the incident he has now come to regret. If we assume that Harold has little or no continuing contact with his former colleagues, then the assurance he asks you to provide can make little difference to his deliberation and planning for the future—it won't even lead him to let his guard down, for instance, since he no longer interacts with the former colleagues from whom his secret is to be kept. Yet Scanlon seems correct to claim that assurance might be valuable to Harold, even under the circumstances described. He has reason to want that his former colleagues not have their conception of him colored by learning of the embarrassing episode from his distant past. And he has reason to want the prospective knowledge that this state of affairs will obtain, because (for instance) this subjective condition will affect his peace of mind, even if it fails to impinge in any discernible way on his deliberation and planning.

All the same, it seems that in the vast majority of cases, the value of assurance is connected to its effects on agency. In familiar coordination cases, for example, my assurance that you are going to do X has no value to me independent of its role as a basis for deliberation and planning: the only reason why I want to be assured that you are going to come around to my office at noon tomorrow is that we need to get together, and I need to integrate our meeting with my other activities and goals. Scanlon apparently worries that if we understand the value of assurance in terms of this kind of reliance, it will remain mysterious why there should be any pressure to go beyond the principle of Loss Prevention in understanding our obligations to those whose expectations for the future we have deliberately influenced. But this concern seems misplaced. Even in the familiar cases that involve effects on our agency and

planning, assurance represents a value sufficient to justify principles stronger than that of Loss Prevention.

To see this, we need only compare assurance in Scanlon's sense with a weaker condition, what we might call conditional assurance: the prospective knowledge that another person will either do X, or provide fair warning or compensation in the event that X is not in fact done. This weaker condition is one that we typically have reason to want to obtain, for reasons similar to those canvassed above. But if conditional assurance is good, full Scanlonian assurance is still better, even if we restrict our attention to situations in which reliance and agency effects are at issue. Our ability to plan effectively our own activities, and to coordinate them with the activities of others, is clearly enhanced when we obtain Scanlonian assurance that another person will do X, and not merely conditional assurance that fair warning and compensation will be provided if X is not done.

This account of the value of assurance helps us to make sense of the moral obligation to fulfill one's promises, which as we have seen is not neutral between performance and compensation. In particular, Scanlon suggests that the value of assurance justifies the Principle of Fidelity, or F:

If (1) A voluntarily and intentionally leads B to expect that A will do X (unless B consents to A's not doing so); (2) A knows that B wants to be assured of this; (3) A acts with the aim of providing this assurance, and has good reason to believe that he or she has done so; (4) B knows that A has the beliefs and intentions just described; (5) A intends for B to know this, and knows that B does know it; and (6) B knows that A has this knowledge and intent; then, in the absence of special justification, A must do X unless B consents to X's not being done (p. 304).

Given condition (1), an obligation is triggered under F only if one gives the hearer assurance that one will do X. One gives this assurance only if one gives the hearer reason to believe that one will be motivated to do X. To do this, one might appeal to any number of motivations. In the Guilty Secret case, for example, you might remind Harold that you have no interest in embarrassing him, or that revealing the secret would reflect poorly on you.

To give Harold this kind of assurance, however, would not be to promise. In promising, Scanlon correctly observes, one assures the hearer and thus triggers an obligation under F in a special way. One leads the hearer

to believe that one will do X precisely because (i) one is motivated to avoid wrongdoing, and (ii) one knows that it would be wrong, having obligated oneself by F to do X, not to do X. “When I say, ‘I promise to help you if you help me,’ the reason that I suggest to you that I will have for helping is my awareness of the fact that not to return your help would, under the circumstances, be wrong: not just forbidden by some social practice but morally wrong—disallowed by the kind of moral reasoning that lies behind Principle F” (p. 306). As we saw above, this feature is connected to the distinctive utility of promising: that it allows the promiser to assure the promisee that she will do X without appealing to any independent reason to do X. In promising to do X, the promiser generates assurance in the promisee by reflexively invoking a moral obligation that is brought into being by the very act of promising itself. This is crucial in cases, such as that of Hume’s farmers, in which the promisee desires assurance that the promiser will do X and the promiser desires to give such assurance, but both are aware that the promiser has no independent reason to do X.

IV. THE CIRCULARITY PROBLEM

Scanlon observes that his account, like others that dispense with the practice view, may seem to be mired in a kind of circularity (pp. 307–08).¹¹ Is it really possible for A to promise B to do X simply by communicating to B (i) that A will be motivated to avoid wrongdoing, and (ii) that A will know that it would be wrong, having obligated herself by F to do X, not to do X? Consider:

(a) In order for B to be assured—by appeal to F, in the way in which Scanlon describes—that A will do X, A must first be obligated by F to do X.

(b) In order for A to be obligated by F to do X, condition (1) of F must first be satisfied.

(c) In order for condition (1) of F to be satisfied, B must first be assured that A will do X.

11. He attributes this objection both to Hume and to G.E.M. Anscombe, “Rules, Rights, and Promises” in *Ethics, Religion, and Politics: Collected Philosophical Papers* (Minneapolis: University of Minnesota Press, 1981), pp. 97–103. An especially clear statement of the general worry can be found in G. J. Warnock, *The Object of Morality* (London: Methuen, 1971), at pp. 99–100. We discuss Warnock’s proposed solution, which does not appeal to social practices, in Section VII below.

From these claims, it follows that:

(d) In order for B to be assured—by appeal to F, in the way in which Scanlon describes—that A will do X, B must first be assured—by some other means—that A will do X.

This is not to deny that there often are other means by which A can assure B that A will do X and thus trigger F. First, A might assure B that A will do X on the grounds that she has what we might call a “prior reason” to do X, that is, a reason to do X, typically of a nonmoral character, that A has prior to and independent of having made any promise. For instance, A may lead B to believe that she will go to the opera on Sunday by indicating that the cast features her favorite singer, that she has had tickets to the performance for months, that arrangements have been made for a babysitter, and so on. To trigger F solely by signaling that one has a prior reason of this kind, however, is not to promise. In promising, one distinctively appeals to a motive that one will have only as a result of having made the promise, and that engages one’s sense of duty. One need not have any prior reason to perform, and even if one does, the communication of this reason is not what makes one’s action a promise.

Second, A might assure B by indicating that, as a result of A’s communicative performance, A will have a “nonmoral practice-based reason” to do X (an “nmp” reason, as we shall henceforth call these sorts of reasons). To take Scanlon’s example, A might swear to B on his honor as a member of the “Sacred Brotherhood of Reindeer” to do X. In this case, A appeals to a nonmoral code that requires Reindeer to do what they swear to do in the name of the Brotherhood (p. 298). Again, however, to trigger F solely by signaling that one has an nmp reason of this kind is not to promise. Even though A appeals to a motive that he will have only as a result of his communicative action, this motive does not engage his sense of impartial moral duty. In cases of promising, it seems, A must lead B to believe that A will do X because A has a “moral practice-based” reason to do X, that is, because A’s communicative action morally obligates her to do X, in virtue of a duty not to exploit or undermine social practices. This is precisely the practice view.

There are several considerations that might lead one to think that the circularity problem is an illusion. First, one might wonder why B should not simply go ahead and believe that A will do X. “It is not as though if B believes that A will do X, B’s belief will be false. If B believes that A will

do X, then A will be obligated under F to do X, and so A will do X.” The proposal, in other words, is that B should form an initially groundless belief that A will do X, which will cause A to do X, thus making the belief true. One problem with this “power of positive thinking” proposal is that it is not under B’s voluntary control to form the belief that A will do X. She needs to have evidence for it.¹² Another problem is that it is not true that if B forms the groundless belief that A will do X, then A is obligated under F to do X. A is obligated under F only if A has led B to believe, on the basis of evidence, that A will do X. This condition would not be satisfied if, *per impossible*, B simply willed herself to believe it.

Second, one might observe that in order to convince herself that A will be motivated by a sense of obligation to do X, B does not need to believe that A is obligated. B only needs to believe that A *believes* that she is obligated. B knows that if A believes that she is obligated, then her sense of duty will kick in. Suppose, therefore, that B pretends to believe that A will do X, thereby leading A to believe that B believes it. Suppose that B succeeds, knowingly, in convincing A that B has this belief. So B believes that A believes that condition (1) is satisfied. So B believes that A believes that A is obligated, under F, to do X. So now B does believe that A will do X. So now condition (1) really is satisfied, and A is in fact obligated to do X. B’s deception puts a virtuous circle in motion; it generates a genuine obligation without A having any motivation other than her sense of duty. It may be doubted, however, that this deception would actually satisfy F. As with the previous proposal, A has not *led* B, in the required sense, to believe that A will do X, nor do the conditions of “mutual knowledge” embodied in clauses (4) through (6) seem to be met. In particular, A’s (eventually true) belief that B believes that A will do X does not amount to knowledge that B believes that A will do X. If A reflects on the situation, then she ought to realize that B is pretending, given that B has no reason to believe that A will do X.

Third, the circularity problem might appear to arise only when we attribute to potential promisees an unnatural degree of reflective self-consciousness. In normal cases, surely, promisees accept that promisers are going to do what they promise without thinking very clearly about the promisers’ reasons for so acting. They accept that promisers are

12. A point compellingly made by Michael Pratt, “Scanlon on Promising,” *Canadian Journal of Law and Jurisprudence* 14 (2001): 143–54, at p. 153.

motivated by a concern to avoid violating obligations that have been incurred by the very act of promising, in the sense that if asked to reflect on why they expect that promisers will do what they promise, they will offer this motivation. If they lacked this much background knowledge, they would not even count as participants in the social practice of promising, a practice that we have seen to be constituted in part by awareness that promisers incur obligations by the very act of promising. But their understanding of this obligation, and, in particular, of the feature of the act of promising that generates it, may be indeterminate. More importantly, typical promisees give little thought to this obligation and its motivational significance in coming to believe that the promisers will do what they promise. Promises normally trigger beliefs in promisees automatically, without any explicit reflection on their part. There seems no reason to insist, especially when one takes into account the background knowledge that promisees have, that beliefs produced in this way fail to qualify as genuine assurance. Thus, in normal cases, promises assure promisees without their engaging in any reflection on promisers' reasons for doing what they promise.

These remarks, although correct, do not succeed in blocking the circularity objection. The question is: Can an account of the obligation of fidelity spell out clearly the conditions that trigger that obligation, and would the obligation survive under explicit awareness on the part of both promisers and promisees of the conditions that provide its basis? It seems a fair demand on Scanlon's account, and a demand he accepts, that the account be able to describe at least one instance of promising that is in this way stable under reflection. One might put the challenge in this way: Could a group of people who were convinced by Scanlon's account of promising, and who knew themselves to be so convinced, continue to succeed in making promises to one another?

Finally, it might be suggested that the circularity problem derives from an implausibly stringent interpretation of condition (1) in Principle F. Instead of requiring A to assure B that A will do X, condition (1) might be weakened, to require only that A assert that A will do X or announce an intention to do X. A would then be able to say: "I will do X. Now you can rest assured that I will do X, because by telling you that I will do X, I have obligated myself under F to do X." In this scenario, A appeals to F to produce assurance, but that assurance has been produced is not a condition of F. Some of Scanlon's formulations may seem to suggest that he is

attracted to this weaker interpretation of condition (1); at one point, for instance, he suggests that in uttering the formulation, “I promise to do X,” a person is in part “claim[ing] to have a certain intention” (p. 307). But on reflection it seems clear that this strategy should be resisted. If F can be triggered in cases in which no assurance has been produced, then it becomes hard to maintain that the justification of F lies in the value of assurance.¹³ It is this thought, it seems, that underlies Scanlon’s denial that F applies in cases in which the would-be promisee has not

13. Michael Pratt explores the possibility of a similar weakening of condition (1) of F in a draft of “Promises and Perlocutions,” *Critical Review of International Social and Political Philosophy* (forthcoming). He considers the suggestion that the contractualist justification of the revised F would be at least as strong as the contractualist justification of the original F. This suggestion, however, depends crucially on the premise that B’s believing that A will do X, as opposed to A’s actually doing X, is an insignificant aspect of the value of assurance. As we argued in Section III, this premise is mistaken. The value of assurance, in most cases, is rooted in its role in B’s deliberation and planning, a role that it has only in virtue of B’s believing that A will do X.

A different justification for the revised F might seem to be provided by the very value that underlies the moral practice-based obligation: the value *to agents* of the normative power to give themselves, by announcing an intention to do something, a compelling reason to do that very thing. Given that this normative power is something of value to agents, why can’t the fact that the revised F would confer it on agents suffice to justify that principle directly, without appeal to a social practice? Joseph Raz, “Promises and Obligations,” in P.M.S. Hacker and Joseph Raz, eds., *Law, Morality, and Society: Essays in Honour of H.L.A. Hart* (Oxford: Clarendon Press, 1977), pp. 210–28, suggests that the obligation to keep promises might be justified along these lines (although he emphasizes the constitutive value of this normative power in allowing one to enter into intrinsically valuable special relationships, rather than its instrumental value in securing cooperation).

Let us consider this line of argument more closely, adopting the contractualist framework that Scanlon assumes in defending his approach. Suppose that I have announced an intention to do X, but now I am inclined not to do X. Does anyone have reason to reject a principle permitting me to do this? It might be argued that we all potentially do, insofar as we have reason to want the normative power to bind ourselves by announcing an intention. A principle that permitted me not to do X (having announced an intention to do so) would permit everyone not to do X (having announced an intention to do so), rendering this normative power generally unavailable. Assume for the sake of argument that considerations of this kind are sufficient to justify the revised F; who in particular would be wronged by behavior that violates this principle? The answer, if the above justification is correct, is that behavior of this kind could be construed as wronging the *agent* who engages in it. It is, after all, the agent who is concretely deprived of the value that underlies the general obligation of fidelity, according to the line of argument we are here considering. If that value is the ability to bind myself normatively through the announcement of an intention, then in failing to deliver I disadvantage myself, depriving myself of something I have reason to want. Quite apart from its intrinsic peculiarity, however, this account delivers a distorted picture of the way in which failures of fidelity affect the persons who are implicated in them. It seems

been assured that the would-be promiser will make good on his announced intention.¹⁴

V. SCANLON'S RESPONSE

Scanlon proposes a different response to the circularity objection. His proposal appeals to the Principle of Due Care, or D, which belongs to the same family of principles "regarding the expectations that we lead others to form about what we intend to do" (p. 295) to which F and L belong:

One must exercise due care not to lead others to form reasonable but false expectations about what one will do when one has good reason to believe that they would suffer significant loss as a result of relying on these expectations (p. 300).

Scanlon suggests that B can arrive at his belief that A will do X on the grounds that (i) A abides by D, and (ii) A is openly attempting to get B to believe that A will do X. B should reason:

- (a) If A does not believe that he will do X, then his attempting to get me to believe that he will do X violates D.
- (b) Since he would not violate D, he must believe that he will do X.

clear that in reneging on a promise, I am wronging people other than myself, above all the promisee; when I make a promise, I accordingly produce assurance in the hearer by communicating my awareness that I would be wronging *others* if I did not perform.

Scanlon himself sometimes suggests that the justification of F lies partly in the reasons that potential promisers have to want a normative power to bind themselves. Yet Scanlon is clear that the reasons that potential promisers have to want this normative power derive from the reasons that potential promisees have to want to be assured. Potential promisers have reasons to want to be able to assure potential promisees simply because such assurance matters to promisees, not because it is a means to or constituent of something else that matters to promisers. "Given the reasons that potential promisees have for wanting assurance," Scanlon writes, "potential promisers have reason to want to be able to provide it" (p. 303).

14. One such case is that of the "Profligate Pal," who habitually fails to repay debts to his friends (p. 312). Now ashamed of this and desperate for cash, the pal pleads with you to give him one more loan. He sincerely believes that "he has turned over a new leaf" and will pay you back this time. "You do not believe this for a minute, but out of pity you are willing to give him the money he needs" (p. 312). Scanlon maintains that the pal's sincere claim to have a certain intention is not enough to trigger F. He needs to convince you. If the pal has an obligation to repay you, therefore, it is not an obligation under F, but most likely a debt of gratitude for a "gift, tactfully described as a loan" (p. 314).

(c) The fact that A believes that he will do X is a good reason to think that A will do X.

(d) Therefore, A will do X.

Since B now believes that A will do X, condition (1) is met, and A is obligated by F to do X.¹⁵

It might be thought that this response tacitly concedes the circularity objection. B's reason for believing that A will do X is not, at least in the first instance, the reason that is supposed to be distinctive of promising: that A will be *obligated by F* to do X. It is instead that, given A's responsiveness to principle D, A would not attempt to get B to believe that A would do X, if A did not believe it himself. But any successful account of promising will have to concede at least this much to the circularity objection. In order for principle F to become engaged in the first place, there has to be some reason for B to believe that A will do X that is independent of A's being obligated by F to do X, but that is not a prior or nmp reason in the above senses (that is, a reason that A had prior to making the promise, or that is grounded in nonmoral practices). To appeal to a prior or nmp reason, again, is not to promise. It is distinctive of promising that A's public performance itself reflexively provides B with a reason to believe that A will do X that is connected to A's concern to avoid wrongdoing, but that is not yet grounded in an appeal to Principle F. Once assurance has been generated in this way, Principle F can become engaged, and A's concern to avoid violating F can provide B with a further reason for assurance that A will do X, through a kind of "moral multiplier effect" (p. 308); but it is important that assurance initially be secured without assuming that Principle F requires A to do X. Viewed in this context, Scanlon's appeal to Principle D has considerable attraction, suggesting an initial basis for B's belief about A's behavior that is at once independent of Principle F, and connected in the appropriate way to A's distinctively moral motivations. It is a basis, moreover, that is intelligible independent of the fact that promising is a social convention, and that therefore represents a genuine alternative to the practice view.

15. Note, however, that as D presently reads, it does not apply to cases like the Guilty Secret, where the promisee will not "suffer significant loss as a result of relying on these expectations" and the promiser has no reason to believe that he will. It seems that Scanlon needs to add the rider, "or has good reason to believe that they would value the assurance that those expectations might constitute."

A different objection is brought by John Deigh, who argues that Scanlon's response cannot account for promises in circumstances in which deception is permissible, such as war and bargaining.¹⁶ In surrendering, for example, one promises to cease hostilities. According to Scanlon's response, one can promise only in situations in which D is in effect; but D is not in effect in war, for it is permissible to lead one's enemies to have false expectations about what one is going to do. Scanlon, therefore, cannot explain how surrendering might constitute a binding promise. In reply, Scanlon observes that Deigh's conclusion only follows if one:

assumes that the permission to mislead in situations of the kinds he describes would be a blanket permission. . . . [H]owever, it is more plausible to suppose that this permission is selective. In wartime it is permissible to deceive one's opponents on the battlefield in certain ways. But why suppose that this extends to a permission to deceive them about one's sincerity in entering into a peace agreement? What purpose would this permission serve? Similarly, why think that the permissibility of misleading others about one's "bargaining position" (by saying such things as, "This is my final offer") extends to misleading them about one's willingness to abide by the offers one makes? Again, what purpose would this serve?¹⁷

Indeed, we think Scanlon is entitled to a slightly more assertive reply. People have reasons to want the ability to exit special contexts such as war, negotiation, and bargaining on mutually acceptable terms, when such terms become available. It is in fact constitutive of negotiation, for example, that participants aim to exit it on mutually acceptable terms, that is, to reach a negotiated settlement. Negotiating in bad faith and negotiating for its own sake (if there is such a thing) are parasitic on this standard case. There would be little point in negotiation, therefore, if there were no reliable way to reach a negotiated settlement. Similarly, there would be little point in bargaining if there were no reliable way to strike a bargain. War is somewhat different, since it can be rational to wage war even when there is no hope of ever arriving at a mutually acceptable peace. Defending ourselves against a relentless

16. "Promises under Fire," *Ethics* 112 (April 2002): 483–506.

17. "Reasons, Responsibility, and Reliance: Replies to Wallace, Dworkin, and Deigh," *Ethics* 112 (April 2002): 507–28, at pp. 527–28.

but evenly matched aggressor may be the best option available to us, even if it means war without end. Nevertheless, we and our enemies still have reason to want to be able to exit the context of war on mutually acceptable terms, should such terms ever become available. In particular, we and our enemies have reason to want the ability to surrender, should we and they, for our different reasons, ever find it preferable to the continuation of armed conflict. Thus, while some kinds of deception may be permissible within these special contexts, there are compelling reasons, intelligible within a contractualist framework, why one kind of deception should never be permissible: namely, deception about one's intention to abide by a particular resolution that one has explicitly proposed or consented to, as a way of exiting the context on mutually acceptable terms. Unless we can rely on such announcements, we have no effective way to exit these contexts when mutually acceptable terms become available, which is something that we generally have reason to want.

VI. OUR OBJECTION

A more fundamental problem with Scanlon's response, we suggest, is this: If A's appeal to D assures B that A will do X, then it does so by providing B with reason to believe that A has some prior or nmp reason to do X. We will begin by defending this claim, and then go on to explain how it undermines Scanlon's response.

To a first approximation, recall, the reasoning that A is supposed to induce B to follow is:

- (a) If A does not believe that he will do X, then his attempting to get me to believe that he will do X violates D.
- (b) Since he would not violate D, he must believe that he will do X.
- (c) The fact that A believes that he will do X is a good reason to think that A will do X.
- (d) Therefore, A will do X.

Let us look at this chain of reasoning more closely. Why is A's belief that he will do X a reason for B to conclude that A will in fact do X? To start, we might note that the belief in question has as its subject matter something that is under A's control, namely his doing (or failing to do) X. If A believes that he will do X, it must be because A has formed the intention

of X-ing, and this in turn might seem to make it reasonable for B to conclude that A will actually do X when the time comes.

In fact, however, the situation is more complex. B must assume not simply that A intends to do X, but that A intends to do X on the basis of compelling reasons. To see this, note that A can intend to do X without taking himself to have any compelling reason to do X. A may simply intend to do X on a whim. If A forms the intention to do X on this basis, however, then it is not clear that B has good reason to believe that A will do X. Forming, at some earlier time, an intention to do X at time *t* does not give A a reason to do X at *t*. Instead, having at time *t* an intention to do X at *t* rationally requires A at *t* to do X, making it irrational for A at *t* not to do X.¹⁸ Thus, the fact that A has formed a prior intention to do X at *t* is not itself a reason against abandoning that intention and doing something else. What we can say is this: *if* A does not abandon his prior intention to do X at *t*, so that the intention survives until *t*, *then* A is rationally required at *t* to do X. If B has reason to believe that A behaves morally, as has been assumed throughout this discussion, then it would seem that B has reason to believe that A behaves rationally. It follows that B has reason to believe that if A's present intention persists, then A will do X. The question is whether B has reason to believe that A's present intention will persist.

If we assume that A has no independent reason to act as he presently intends, then he has no reason not to revise his intentions should a reason to do so, or merely a contrary whim, occur to him. In that case, B's only basis for believing that A's present intention will endure would seem to be a general belief in the inertia of intentions: a belief that, other things equal, intentions, once formed, are more likely than not to persist. This no doubt gives B *something* to go on; it may justify him, for instance, in assigning a higher probability to A's doing X than he would be justified in assigning in the absence of A's announcement. The question is whether it suffices to *assure* B that A will do X. If we focus on typical cases, in which the value of assurance lies in the basis that it provides for deliberation and planning, then the question becomes whether A's present intention suffices to justify B in *relying* on A to do X.¹⁹ Only if it does

18. Compare John Broome, "Normative Requirements," *Ratio* (new series) 12 (1999): 398–419.

19. In the cases like the Guilty Secret, in which reliance is not at issue, the analogous question might be whether A's present intention sufficed to justify B in, say, feeling relieved

will A have succeeded in producing the distinctively valuable kind of assurance that it is the point of F to protect, and thus have obligated himself under F to do X.

There are perhaps some cases in which A's present ungrounded intention would suffice to justify reliance. If A's reasons against doing X are weak enough, and if the potential loss to B from A's not doing X is sufficiently small relative to the potential gain, then the higher probability that A's present intention entitles B to assign to A's doing X might make it reasonable for B to rely on A's doing X. The problem is that promisers often depend on promises to secure assurance in situations in which they have compelling reasons against performance and promisees risk significant loss. "Typically," as Scanlon observes, "a promise is asked for or offered when there is doubt as to whether the promiser will have sufficient motive to do the thing promised. The point of promising is to provide such a motive" (p. 322). Consider the paradigm example of Hume's farmers, in which A has weighty reasons not to perform, and in which B's potential loss is considerable. Here it is utterly implausible to suggest that the fact that A has a present intention to do X, which he has strong reasons to revise and no reason not to, would suffice to justify B in relying on A to do X. In situations such as these, A can only hope to *assure* B by leading him to believe that A's present intention is what Scanlon calls a "settled intention" (p. 308), one that is grounded in compelling reasons to do X. What kind of reasons might these be?

Notice, first, that A's compelling reason for doing X cannot simply be that, having assured B that A will do X, A will be obligated under F to do X. We are imagining that A and B are fully reflective and philosophically informed; the question we have posed is whether the obligation of fidelity would survive under explicit awareness on the part of both promisers and promisees of the conditions that provide its basis. Hence, we can and must assume that both parties are cognizant of the potential circularity problem. They know that A does not yet have an obligation under F until A has assured B by other means that A will do X. What we need to identify, therefore, is a different kind of reason for A to do X, a reason that is already available at the point when A communicates to B

and resting easy at night, or whether those responses, given what A has given B to go on, would simply be wishful thinking.

the future-directed intention to do X, and that might contribute to generating in B the kind of assurance that engages Principle F in the first place.

Notice, second, that Scanlon cannot suppose that A intends to do X because, as a result of his communicative performance, he has a moral practice-based reason to do X. It would certainly not be incoherent, if such a practice exists, for A to intend this; indeed, we will argue later that the circularity problem can most effectively be avoided by appealing to such moral practice-based reasons. But Scanlon cannot avail himself of reasons of this kind without abandoning his opposition to the practice view: the thesis that the wrong involved in breaking promises can be understood only by appeal to the fact that promising is a social practice.

If A cannot lead B to believe that he has F-based or moral practice-based reasons, however, then the only alternative is to lead B to believe that he has prior or nmp reasons to do X. Yet Scanlon himself seems to agree that his account would fail if B must assume that A's settled intention to do X is grounded in an "independent source of motivation" of the sort represented by prior or nmp reasons (p. 308). As we saw above, it is supposed to be distinctive of promising that the promiser incurs an obligation of fulfillment (to do X, say), without signaling that he has some prior or nmp reason to do X. It was this that was supposed to distinguish promises from the communicative actions involved both in the opera example, in which A assures B that he will go to the opera by citing prior reasons that he has, e.g., that it will feature his favorite singer, and in the Reindeer example, in which A assures B that he will do X by citing an nmp reason grounded in his honor as a Reindeer. What we now see, however, is that the appeal to D that Scanlon envisions, no less than these other communicative actions, assures B by leading him to believe that A has prior or nmp reasons. The only difference, it seems, is that in appealing to D, A does not specify what these reasons are. Instead of saying, "I'm going to the opera, because it will feature my favorite singer," A says, in effect, "I'm going to the opera, because I have compelling reasons to do so, but I will spare you the details of what they are."

Still, Scanlon might maintain that this difference is important. In the nonpromising cases, A assures B by stating the reasons that he has to do X, whereas in the promising case, A assures B by getting B to take it on trust that A has some reasons or other to do X, even though A declines to state what they are. It remains distinctive of promising, therefore, that A

assures B by reflexively appealing to a moral motivation: namely, A's concern not to deceive B. Moreover, this account explains to some extent why promising might be distinctively useful. For reasons of economy or privacy, A may prefer not to specify his reasons for doing X. Promising, as Scanlon understands it, allows A to assure B without specifying his reasons.

As we observed earlier, however, the distinctive utility of promising is not simply that it allows A to assure B that A will do X when A has prior or nmp reasons to do X that he prefers not to communicate to B, but also that it allows A to assure B when A does not have *any* prior or nmp reasons to do X at all. This is precisely what makes promising so beneficial in cases such as that of Hume's farmers, in which the parties have no prior or nmp reason to do what they have promised, and in which the lack of such independent reasons is common knowledge.²⁰ On Scanlon's account, promises can no longer perform this vital service. If A can promise to do X only by leading B to believe that A has some prior or nmp reason to do X, then in a situation in which A has no prior or nmp reason to do X, A cannot promise at all.²¹

20. It may seem patently false to say that farmer A has no prior reasons in this case, for surely he has the prior reason of wanting his crops harvested in a timely and efficient manner. Indeed, this is a prior reason, a reason that A has independently of any communicative action that he performs. However, it is not a prior reason to perform the promised action; it is not a prior reason to help B with harvesting his crops, nor even a prior reason to help B in harvesting his crops provided that B has helped A in harvesting his. It is instead a prior reason to *assure* B that A will perform the promised action. By assuring B of this, A can induce B to help him, and thus ensure that his own crops are harvested in a timely and efficient manner. This prior reason to assure B may give A reason to form an intention to help B, in order to give B some grounds for assurance. But it does not give A any compelling reason to sustain this intention until the time of action, and so it fails to give B the sort of grounds for assurance that we are now supposing that he needs.

21. In a situation of this kind, it would be possible for A to resort to deception about his reasons in an effort to generate in B assurance that he will do X. Such deception may not violate D, because A may reasonably believe that B will not incur any loss as a result of relying on those false expectations. If A's deception succeeds and B is falsely assured that A intends on the basis of some prior or nmp reason to do X, then A will have triggered an obligation under F to do X. If A intends to comply with this obligation should it be triggered, then A will not expect B to suffer any loss from relying on the false belief that A intends to do X on the basis of some prior or nmp reason. Nevertheless, this deception does violate Scanlon's Principle ML, that "one may not, in the absence of special justification, act with the intention of leading someone to form a false belief about some matter, or with the aim of confirming a false belief he or she already holds" (p. 318). The upshot is that in a situation in which he takes himself to have no prior or nmp reason to do X, a scrupulously

The core of our objection to Scanlon's account, then, is that it cannot explain how one could succeed in promising in the paradigm cases epitomized by Hume's farmers. These paradigm cases are characterized by two important conditions. First, what B stands to lose from relying on A if A fails to perform is great relative to what B stands to gain if A does perform, and, second, A has no prior or nmp reasons to perform. The first condition entails that A's communicating a bare intention to do X will not suffice to assure B, and hence that A must appeal to compelling reasons he has to do X. The second condition deprives A of the only reasons to which, on Scanlon's account, A might appeal: prior or nmp reasons. It is important to stress that these conditions are neither exotic nor insignificant. Not only are they conditions under which, as a matter of fact, promises typically succeed in generating assurance; they are also precisely the conditions that give promising its special utility and point. The first condition is what makes pressing the need for mechanisms of generating assurance, and the second condition is what requires that there be a mechanism, like promising, that does not appeal, even implicitly, to prior or nmp reasons.²²

VII. THE RETROSPECTIVE ASPECT OF THE PRINCIPLE OF VERACITY

What we need, in outline, is a moral principle that would "ratchet up" something less than the production of assurance into the genuine article. Suppose that there were a principle that required A to do X whenever he engaged in a certain sort of communicative act, whether or not it

moral agent, who abides by ML as well as D and F, cannot even attempt to promise to do X, if promises indeed function by indicating the presence of prior or nmp reasons.

22. Pratt rejects Scanlon's proposed solution in the following passage: "The Principle of Manipulation establishes that it is wrong to convince another that one intends to do something that one knows the other cares about and may rely on being done, if one does not actually have this intention. It is presumably also wrong to *attempt* to do this, and so it follows that it is wrong to utter the words 'I promise to X' in such circumstances without actually having an intention to X. But it does *not* follow that the intention of the promisor to do X is 'settled' in the sense that the promisor may not change his mind without the consent of the promisee" ("Scanlon on Promising," p. 153). Pratt's concern appears to be that although the speaker (who says, "I promise to do X," and who abides by the Principle of Manipulation) must have some settled intention to do X, the intention need not be the kind of settled intention that *promisers* distinctively acquire as a result of promising: namely, an intention backed by the recognition that he has incurred an obligation to do X, unless released by the promisee. (Here we read Pratt's "may not change his mind" as "may

produced complete assurance. Having performed an act of this kind, A would then be obligated to do X, and this fact could assure B, in turn, that A would do X. At no point in this process would A have to appeal to any prior or nmp reason to do X; A could generate assurance solely by signaling his readiness to comply with obligations incurred by the communicative act itself. The questions are: What kind of moral principle might perform this ratcheting function? And what kind of communicative act would it recognize as triggering obligations of fulfillment?

An initially attractive suggestion is that some principle prohibiting A from generating false expectations might serve this ratcheting function. We tend to think of such principles under what might be called their “prospective” aspect, i.e., as prohibiting agents from generating expectations that they have reason to believe are false. Yet these principles might also be viewed as having a “retrospective” aspect, i.e., as requiring agents who have already generated expectations to ensure that those expectations are not false, insofar as this is in their power. The suggestion is that if A has generated an expectation that he will do X, then such a principle might obligate A, under its retrospective aspect, to see to it that this expectation is true: that is, to do X. A would be obligated, one might say, not to make a liar of himself.

not, morally speaking, change his mind.” It is hard to see how else the phrase, “without the consent of the promisee,” would be relevant.) But it is not clear why, at this stage of Scanlon’s argument, the settled intention must be of this form. As we have reconstructed Scanlon’s proposal, what is communicated by the speaker (who says, “I will do X,” and who is known to abide by D) is that he has a settled intention of a different kind, a settled intention which is backed by prior or nmp reasons. The fact that the speaker has such an intention should, under normal circumstances, be sufficient to assure the hearer that he will do X. Given that the hearer has now been assured, and given that the speaker abides by E, the hearer may reasonably conclude that the speaker has a settled intention of the kind that, in Pratt’s view, goes missing: an intention that the speaker may not, morally speaking, abandon without the consent of the promisee. Our objection, to repeat, is not that this mechanism cannot succeed, but instead that it is not available to speakers in situations in which they have no prior or nmp reasons. It may be that Pratt ignores the possibility that the settled intention is based on prior or nmp reasons, because he implicitly assumes that the situation is one in which the speaker has no such reasons. But if so, then the problem is not that what the speaker has communicated is somehow insufficient, which is what Pratt’s formulation suggests, but instead that what the speaker would communicate would be deceptive, and hence that it is not available to him to communicate, so long as he is morally conscientious. Immediately following the quoted passage, Pratt offers a response on Scanlon’s behalf, which he then refutes. This response is roughly equivalent to the “power of positive thinking” proposal that we consider in Section IV, above.

A principle can serve the ratcheting function, of course, only if it obligates A whether or not his communicative act succeeds in inducing reliance or assurance. This means that D is not the principle that we are after. D requires A to take due care not to lead B to form reasonable but false expectations when A has reason to believe that B will suffer loss from relying on those expectations. But B will rely on those expectations only if he has already been assured, in the relevant sense, that A will do X. Since the retrospective obligation under D would be triggered only when assurance has already been produced, we cannot appeal to it to ratchet up something less than the production of assurance into the real thing.

Nevertheless, it might seem plausible that there is a principle that obligates A whether or not his communicative act succeeds in inducing reliance or assurance. To communicate a falsehood knowingly is still to lie, even if the liar has sufficient reason to believe that she will not succeed in inducing reliance or assurance in her listeners. If the general principle of veracity that prohibits lying has a retrospective aspect, then it would seem that simply by communicating to B that A will do X, A is obligated to make this communication true, whether or not the communication succeeds in inducing reliance or assurance. The Principle of Veracity, call it "V," would thus ratchet up the mere assertion that A will do X into the production of assurance that A will do it.²³

The appeal to V is miles away from Scanlon's account of promising. Not only does Scanlon not explicitly offer it, it also relies on assumptions that Scanlon does not endorse, and it would radically transform his account of the obligation to keep promises. To begin with, the appeal to V assumes that there is a kind of symmetry in the prospective and retrospective aspects of principles forbidding the generation of false expectations. Just as one is forbidden from generating an expectation that one believes to be false, so too one is required, having generated an expectation, to make it true. Yet Scanlon implicitly rejects this assumption. His Principles D, L, and E, which govern the generation of expectations about one's future actions, are asymmetric. D, for instance, which functions only prospectively in Scanlon's system, requires A, in general, not to generate an expectation that A will do X if A has reason to believe that

23. Warnock, *The Object of Morality*, pp. 94–117, presents this as a solution to what is, in effect, the circularity problem. For critical discussion, see P. S. Atiyah, *Promises, Morals, and Law* (Oxford: Clarendon Press, 1981), pp. 104–05.

that expectation would be false. There is no corresponding retrospective principle that requires A, in general, to do X when he has created an expectation that he will do X. On the one hand, F obligates A to do X when A has generated an expectation, but only when this expectation was generated in the special circumstances captured by conditions (2) through (6). On the other hand, L obligates A whenever he has generated an expectation, but it does not obligate A to do X. L is satisfied just as well by a timely warning or compensation. Similarly, Scanlon's general principle of veracity, ML, is clearly understood by him to be an exclusively prospective prohibition on acting "with the intention of *leading* someone to form a false belief about some matter" (p. 318; emphasis ours). Quite apart from these difficulties, embracing the appeal to a retrospective interpretation of V would render L and F otiose. A's simply asserting to B that he would do X would already obligate A to do X, period.

The appeal to V, therefore, is not a solution to the circularity problem that Scanlon could easily adopt. Moreover, it is unclear it ought to be adopted on independent grounds. The natural argument for the appeal to V would begin by urging that V is plausible under its prospective aspect, and then contend, perhaps for reasons of symmetry, that V ought to have a corresponding retrospective aspect. But it is not obvious, at least as a general rule, that every prospective obligation implies a symmetrical retrospective obligation. For example, one might think that prospective obligations are in general more stringent than retrospective obligations, because the burdens of complying with the latter are typically heavier. It is usually less costly to refrain from making a false assertion or inducing false assurance that one will do X than it is to make true such an assertion or such assurance by doing X. Indeed, this appears to be the main reason why the principles in Scanlon's system that approximate the retrospective aspect of D, namely, L and F, are less stringent than the "true" symmetrical retrospective obligation would be. Together, L and F permit a timely warning or compensation instead of performance, except when assurance is created under the special circumstances stipulated in the conditions of F, whereas the "true" retrospective obligation would require performance in every case. As Scanlon observes, "A principle according to which the only way to avoid obligations as binding as those specified by Principle F is to avoid voluntarily creating any expectations about one's future conduct would be too limiting. It would mean, for example, that we could never tell people what we intend to do

without being bound to seek their permission before changing course” (p. 304). This line of thought is reflected in ordinary moral consciousness, as Henry Sidgwick forcefully reminds us in the following passage: “[W]e are not bound to make our actions correspond with our assertions generally, but only with our promises. If I merely assert my intention of abstaining from alcohol for a year,”—to a stranger on a train, say, who clearly has no interest in being assured of this—“and then after a week take some, I am (at worst) ridiculed as inconsistent.”²⁴

While the retrospective interpretation of Principle V represents one possible ratcheting mechanism to overcome the circularity problem, therefore, it is not one that we are presently prepared to endorse, either as a potential amendment to Scanlon’s account, or on its own terms. We have greater confidence in a different ratcheting mechanism, which coheres with many of Scanlon’s fundamental insights, and whose foundations are clearer and less controversial. It is to this mechanism that we now turn.²⁵

VIII. A HYBRID ACCOUNT

The ratcheting mechanism that we have in mind draws crucially on the practice view. Suppose that there is a social practice consisting in the fact that members of a group adhere, and are reasonably expected to adhere, to a policy of fidelity, i.e., a policy of doing X (unless certain

24. Henry Sidgwick, *The Methods of Ethics*, 7th ed. (London: Macmillan, 1907), p. 304.

25. Judith Jarvis Thomson proposes what is in effect a different kind of ratcheting mechanism in *The Realm of Rights* (Cambridge, Mass.: Harvard University Press, 1990), chap. 12. For Thomson, promising is a species of “word-giving,” where the promiser invites the promisee to rely on the truth of a claim, and the promisee “receives and accepts the invitation (there is uptake)” (p. 298). The question is, how are we to understand this notion of “uptake”? If receiving and accepting an invitation to rely is a matter of forming the belief that the promiser will perform, the circularity problem has not been avoided (what reason does the promiser give the promisee for forming this belief?). If on the other hand one receives and accepts an invitation to rely by performing some voluntary act distinct from forming the relevant belief, the account will face different problems. For one thing, it seems that promises often become binding without the promisee’s performing any voluntary act distinct from belief. And even where an act of this kind may be performed, it is obscure what might underlie the obligation to fulfill that this act engages. At one point Thomson appeals to a kind of “natural right” or “metapower”: a power to give one’s hearer the power to generate a claim on one by accepting one’s invitation (pp. 320–21). This is reminiscent of the account by Raz that is criticized in n. 13 above.

acknowledged excusing conditions obtain) should they ever perform a certain kind of communicative act, such as uttering the formula, "I hereby promise to do X." Suppose that A were not to adhere to this policy: that having uttered the formula, "I hereby promise to do X," A were to fail to do X (in the absence of excusing conditions). In exceptional circumstances, A's failure to do X might actually undermine the expectations that constitute the practice. Since the practice has great instrumental value to its participants, A has a moral obligation not to undermine it, and thus to do X. In the vast majority of cases, however, A's failure to perform would not undermine the practice. Nevertheless, he would be free-riding on the contributions that others have made to the practice. The practice exists only because others do their part, by adhering to the policy of fidelity; A is therefore morally obligated, under a general proscription on free-riding, to do his part, by adhering to the policy himself. In other words, having uttered the formula, "I hereby promise to do X," A is morally obligated to do X. It is B's awareness that A acknowledges this practice-based obligation to do X that assures him that A will in fact do X.

This avoids the initial circularity in Scanlon's account: that A's *assuring* B that A will do X is what is supposed to trigger an obligation to do X under F, the very triggering of which is nevertheless supposed to be what assures B of it in the first place. Instead, A's *uttering the formula*, "*I promise to do X*," which is not yet A's assuring B that A will do X, results in A's incurring a practice-based obligation to do X, which only then assures B that A will do X. Furthermore, this account overcomes the problems with Scanlon's response to the circularity objection, which involved A's appealing to his commitment to D to assure B that A intends to do X. According to Scanlon's response, as we saw, A must attempt to assure B that A will do X on the grounds that A has a prior or nmp reason to do X. This means that in a situation in which A does not take himself to have a prior or nmp reason to do X, A cannot assure B without impermissible deception. According to the present account, by contrast, A attempts to assure B not on the grounds that he has a prior or nmp reason, but instead on the grounds that his communicative act gives him a moral practice-based obligation to do X. This account thus makes it possible for A to assure B, without deception, in situations in which A has no prior or nmp reason: in precisely the situations, that is, in which promising is most needed.

For simplicity, we have supposed that the performance that engages the policy of doing X is the utterance of some specific formula, such as, "I hereby promise to do X." A more plausible account of our actual practices, however, might view the performance as asserting that one will do X, or announcing an intention to do X, in circumstances in which it is clear to both parties that the hearer wants to be assured that the speaker will do X. For our purposes, it does not matter which we choose. What is essential to such social practices, and allows them to overcome the circularity problem, is not that the policy of doing X is engaged by some stock formula, but instead that it is engaged by something less than the production of assurance.

One might wonder why this is. Why is Scanlon's Principle F engaged only when one has produced assurance, whereas the policy of fidelity is engaged whenever one performs an appropriate communicative act, even if it fails to produce assurance? The reason is that the policy is an instrumental device for the purpose of creating assurance, which is of particular value in cases, such as that of Hume's farmers, in which assurance cannot be created by appealing to any prior or nmp reason. The only restrictions on the selection of the communicative act that engages the policy are that it be fair to participants and that it permit the policy to serve its purpose. The circularity problem shows that the policy could not serve its purpose—to provide a means of producing assurance—if the performance required that one have already succeeded in producing assurance. Scanlon's F, by contrast, is not meant to specify a policy general adherence to which would constitute a collective, instrumental good. Instead, F is meant to explain how one wrongs particular people by treating them in certain ways. By appealing to the value of assurance, F explains how one wrongs another when one violates assurance about one's future behavior that one has voluntarily and intentionally led that person to have. One can wrong another in this way only when one has actually assured the other about one's future behavior. This is, to repeat, why F requires the production of assurance.

How troubling should this concession to the practice view be to those who, like the authors, agree with Scanlon's criticism of it? The criticism, recall, was that the practice view fails to capture our intuitive sense of what makes promise breaking wrong. The present concession, however, need not foreclose the solution to this problem that Scanlon's appeal to the value of assurance seemed to offer. For once A has promised B to do

X, and conveyed to B that A has a moral practice-based reason to do X, the conditions of F are satisfied. Now A has a *second* obligation to do X, an obligation under F. That is, he has an obligation backed by the value of assurance.

Consider the following hybrid account of promising. There are three ways of triggering an obligation under F. First, A can lead B to believe that A will do X simply because A has a compelling prior reason to do X, which A succeeds in communicating to B in some way or other. This is a reason to do X that A has independent of his communicative act. In some cases, this reason might involve some established social practice other than promising. In some cases, it might be a moral reason, that is, be motivating only because of A's sense of duty. Yet it need not involve another social practice or engage A's sense of duty, and in the familiar cases, it does not (recall our opera example above). Second, A can lead B to believe that A will do X because A has an nmp reason to do X. In this case, the reason to which A appeals is one that A has only because of his communicative action, but that does not engage A's sense of duty. Finally, A can promise. In promising, A leads B to believe that A will do X because A has a moral practice-based reason to do X, such as a duty not to free-ride. That is, A engages in some performance, such as uttering the formula, "I hereby promise to do X," which, because of a social practice of promising, makes it wrong for A not to do X. This is a reason to do X that A acquires only by promising to do X, and it motivates A by engaging his sense of duty.

By appealing to a prior or nmp reason, or by promising, A can lead B to believe that A will do X and trigger an obligation under F. In both cases, A acquires an additional reason to do X, namely, that failing to do so would violate F. In all of these cases, therefore, A has two reasons to do X. In the nonpromising case, A has a prior or nmp reason to do X and is obligated by F to do it, whereas in the promising case, A has a moral practice-based reason to do X and is obligated by F to do it. Distinctive of the promising case are the following two features: first, that both reasons obtain only because A has promised, and second, that both reasons rely only on A's sense of duty. When A fails to perform in the promising case, he commits two wrongs, whereas in the nonpromising case, he commits only one.²⁶ In both

26. One qualification: If A assures B by appealing to an nmp reason, and if the non-moral practice in question is just, then A may also incur a moral obligation, under the principle of fairness, to follow the rules of the practice, i.e., to keep his word. In such a

cases, he commits the wrong of violating F; in the promising case, he commits the additional wrong of exploiting or undermining the social practice.

Recall the two ways in which, as Scanlon convincingly argues, the pure practice view fails to capture our intuitive understanding of what makes promise breaking wrong. First, the practice view cannot make sense of the fact that nonperformance in cases that do not involve promising seems to involve the same kind of wrong as breaking a promise. Second, the practice view cannot account for our conviction that the wrongness of breaking promises is to be explained by reference to the reasons that the promisee in particular would have to object to the promiser's failure to perform. By identifying the wrong of promise breaking with the wrong of exploiting or undermining the practice, the pure practice view cannot account for our belief that the obligation to keep a promise is owed primarily to the individual promisee, rather than being an obligation that is owed collectively to all those who contribute to or benefit from the practice as a whole.

The hybrid account appears to honor these intuitions. First, failure to perform in the nonpromising case is wrong in one of the two ways in which it is wrong in the promising case. It is wrong in virtue of violating F. Second, the explanation of the wrong of violating F, which appeals to the value of assurance, draws on the reasons that the promisee, in particular, would have to object to the promiser's failure to perform. Therefore, the hybrid account can do justice to our conviction that one of the two obligations to keep a promise is owed to the promisee in particular.

Of course, the hybrid account implies that breaking a promise is wrong in an *additional* way, in which failure to perform in the nonpromising case is not. And it implies that there is *one* obligation to keep a promise that is to be explained in terms of the wrongness of exploiting or undermining the practice, and that is thus owed to everyone who plays their part in or stands to gain from the practice. On reflection, this aspect of the hybrid account seems faithful to our intuitive understanding of the obligations incurred by promisers. Although the more significant wrong

case, A actually has three reasons to do X: (i) the nmp reason (e.g., his honor as a Reindeer), (ii) a duty of fairness, and (iii) a duty under F. This still differs from a promise in two respects. First, A intends to assure B primarily by appealing to his nmp reason, not to his sense of duty. Second, the social practice that A invokes is not the social practice of promising; it is not a practice of agreement making that is primarily sustained by its participants' sense of impartial duty.

in breaking a promise is clearly the wrong of violating the assurance given the promisee, we often do, in ordinary thought, recognize the wrong of exploiting the practice. It is familiar enough for a bystander to a broken promise (especially a promise of a kind that she typically keeps at some cost to herself) to complain: "What if everyone did that?" What she means, when her thought is fully spelled out, is roughly: "What gives you the right to enjoy the benefits of the practice, but to excuse yourself from making the same contributions to it that everyone else makes, contributions without which the benefits that you enjoy would not be available?"²⁷ True, if asked, "Is there one wrong in breaking a promise, or two?" most people would probably answer, "One." But it is not clear how seriously we ought to take this. As we have seen, it is remarkably difficult to understand precisely how promises function to generate obligations of fulfillment, and to distinguish clearly the different moral principles that promises are capable of implicating. In light of these difficulties, we should not attach too much significance to unreflected verdicts about the number of wrongs that are involved when promises are broken.

Scanlon's latest discussion suggests that he is willing to countenance some role for practice-based considerations in reconstructing the obligation to keep one's promises. "It could well be true that a person who promises to do something commonly incurs two kinds of obligation to do it: one of these an obligation to abide by and support a useful practice, the other an obligation of the kind Principle F describes. (The first of these obligations would in fact support the second, since the practice would provide a way of creating expectations of the kind that are central to Principle E.)"²⁸ If this is correct, then Scanlon can have no principled objection to an account that draws on the conventional character of promising in reconstructing the moral obligations that are incurred by

27. The hybrid account implies that the profligate pal, described in n. 14 above, has a practice-based moral obligation to keep his promise to repay the loan, but not an obligation under F to do so. On reflection, this seems accurate. One cannot say that the profligate pal wrongs you personally by violating the assurance he has given you, for the straightforward reason that he has not succeeded in assuring you. (He may wrong you personally, however, by violating principle D, by not taking due care to avoid leading you to expect falsely that he will pay you back. Perhaps he ought to take a sober look at his past behavior and conclude that he is unlikely to make good.) Nevertheless, he has free-ridden, albeit to limited effect, on a social practice that exists only because of the contributions of others, a practice whose direct and indirect benefits he has enjoyed on many other occasions.

28. "Reasons, Responsibility, and Reliance," pp. 523–24.

promisers. Moreover, if our argument is on the right lines, a satisfactory theory of the obligation to keep promises *must* make reference to the moral implications of the fact that promising is a practice. Only by doing so can we capture the distinctive way in which uttering the formula, “I hereby promise to do X,” gives rise to a moral obligation of fulfillment under the more general principle of fidelity.