Promises as Proposals in Joint Practical Deliberation

Brendan de Kenessey

Abstract: This paper argues that promises are proposals in joint practical deliberation, the activity of deciding together what to do. More precisely: to promise to ϕ is to propose (in a particular way) to decide together with your addressee(s) that you will ϕ. I defend this *deliberative theory* by showing that the activity of joint practical deliberation naturally gives rise to a speech act with exactly the same properties as promises. A certain kind of proposal to make a joint decision regarding one's own actions turns out to have the very same normative effects, under the very same conditions, as a promise. I submit that this cannot be a coincidence: we should conclude that promises and the relevant kind of proposals in joint practical deliberation are one and the same.

Joint practical deliberation is the activity of deciding together what to do. When friends discuss where to meet for coffee, when doctors debate whether to perform a risky surgery, when partners talk about when to have a third child, they are engaged in joint practical deliberation. They are trying to arrive at a single plan of action that expresses and governs all of their wills.

This paper will have much to say about joint practical deliberation – but that is not its primary point. The aim of this paper is to show how, by investigating joint practical deliberation, we can uncover a new and appealing way of thinking about *promises*.

To set up the main idea, we need to begin from two premises about joint practical deliberation. First, it is possible to make joint decisions not just about what *we* will do, but also about what *you* will do or what *I* will do. For example: if you and I are writing a paper together, we might jointly decide that *I* will write the introduction, leaving it quite open what you will do. Second, if we do in fact engage in joint practical deliberation, then we must have some speech acts by means of which we *propose* joint decisions, putting them forward for collective evaluation.

Putting these premises together, we can infer that there must exist speech acts by means of which we can propose joint decisions about what *you* will do or about what *I* will do. When we ask

what these speech acts would look like, we make a surprising discovery: one of these speech acts has *exactly the same properties as promising*. A certain kind of proposal to make a joint decision regarding one's own actions turns out to have the same normative effects, under the same conditions, as a promise. (Or so I will argue). I submit that this cannot be a coincidence: we should conclude that promises and the relevant kind of proposals in joint practical deliberation are one and the same. That is, we should accept

> **The Deliberative Theory of Promises:** For S to promise A that she will ϕ *just
> is* for S to propose to A, by means of the propose-and-challenge method,[1] that they
> make a joint decision to the effect that she will ϕ.

I am not the first to propose this view. Margaret Gilbert has defended a "*joint decision account*" of promises: for one person to make a promise to another is for them jointly to commit themselves, by an appropriate, explicit process, to the decision that one of them ('the promisor') is to perform one or more specified actions" (Gilbert 2011: 99). Though there are some differences in the details of our accounts, Gilbert's main idea is the same as my own: the result of a valid promise to ϕ is a joint decision between speaker and addressee to the effect that the speaker will ϕ. However, the *arguments* Gilbert offers for this view are very different from the argument I will give in this paper. Gilbert's main argument for the joint decision account is that it vindicates a purported analytic connection between promises and obligation: that the proposition that I have promised to ϕ entails that I am obligated to ϕ (2011: 100). But this thesis is controversial. I myself am convinced neither that the thesis is true nor that the deliberative theory entails it (see §2.1). By tying the deliberative theory so tightly to this contentious claim, Gilbert undersells its appeal.[2]

This paper offers a different kind of argument for the deliberative theory. We should accept the deliberative theory, I shall argue, not because it vindicates some controversial or surprising thesis about promises, but because it provides a powerful explanation of the various features that promises are widely agreed to have.

To see what I have in mind, imagine you are a nineteenth-century physicist developing a theory of molecular motion. As you do your calculations, you come to a surprising realization: your

theory predicts that the *mean kinetic energy* of the particles that compose a substance will be perfectly correlated with that substance's *temperature*. Just as increasing a gas's pressure will heat the gas up, so too it will increase the gas's mean kinetic energy. Just as decreasing the temperature of water will make it freeze, so too will decreasing its mean kinetic energy. What could explain these observations? The best explanation seems to be that temperature and mean kinetic energy are the same property. After all, it is unlikely that there are two separate properties, temperature and mean kinetic energy, which coincidentally happen to have the very same characteristics. And by accepting their identity, you gain a great deal of explanatory power. You can now explain temperature's various quirks and qualities by showing how they can be derived from the laws of molecular motion (cf. Block & Stalnaker 1999: 23-24).

My argument for the deliberative theory of promises takes the same form. I begin by setting out an independently motivated theory of joint practical deliberation. Then I show how this theory predicts that a certain kind of proposal within joint deliberation will have exactly the same properties as promises. Like promises, these proposals obligate the speaker to her addressee to perform the proposed action. They give the addressee the power to release the speaker from her obligation, just as a promise does. And, as with promises, these proposals only have their force if they secure uptake from their addressee and are not elicited by means of coercion or deception. The best explanation of these observations seems to be that promises and the relevant kind of proposals in joint practical deliberation are the same speech act. After all, it is unlikely that there are two separate speech acts, promises and these proposals in joint deliberation, which coincidentally happen to have the very same characteristics. And by accepting their identity, we gain a great deal of explanatory power. We can now explain promises' various quirks and qualities by showing how they can be derived from the laws of joint practical deliberation.

The paper proceeds as follows. Section 1 outlines the theory of joint practical deliberation on which the deliberative theory is based. Section 2 shows how, given this theory, we can predict that a certain kind of proposal in joint practical deliberation will have all the signature features of promises. Section 3 considers an objection. Section 4 compares the deliberative theory to three

3

alternative theories of promising: conventionalism, the expectation theory, and the normative powers theory. Section 5 concludes by arguing that anyone who denies the deliberative theory is committed to the unpalatable conclusion that promises are redundant.

1. Joint practical deliberation

Promising doesn't *seem* like proposing a joint decision. When one imagines joint practical deliberation, what comes to mind is a multilateral, collaborative process of deciding on an action we will perform together. But when I make a promise, I seem to be unilaterally binding myself to an action that I will perform alone. The two appear starkly different. However, this appearance is generated by an overly narrow conception of joint practical deliberation. When we appreciate the full variety of ways in which we can make decisions together, it becomes clear that there is ample space for promises within the activity of joint deliberation. To see this, we need to set promises aside for a while and investigate joint practical deliberation in its own right.[3]

1.1. The structure of joint practical deliberation

We begin from the following premise: joint practical deliberation is an activity of the very same kind as *individual* practical deliberation, the activity of deciding what to do on one's own. Like making lasagna or going on a walk, practical deliberation is the kind of activity that can be performed both together and alone. This suggests that joint and individual practical deliberation will have the same basic structure. This structural analogy between joint and individual practical deliberation is the central motivating idea for the account that follows.[4]

What is the structure of individual practical deliberation? Suppose I am deciding where to get lunch. I am faced with a set of *options*, the potential actions I am choosing between: to go to Restaurant A, Restaurant B, or Restaurant C. My aim is to make a *decision* that narrows these options down to one (say, to go to Restaurant B). However, my decision will (typically) not be arbitrary – it will be based on *reasons*. For example, the fact that Restaurant A sources its food

4

locally might give me some reason to go there, while the fact that I'm craving a burrito might give me more reason to go to Restaurant B, which makes a mean burrito.

Thus individual practical deliberation is a process of making a *decision* about which of one's *options* to perform on the basis of one's *reasons*. Pursuing our analogy, then, we can see joint practical deliberation as a process of making a *joint decision* about which of our *shared options* to perform on the basis of our *shared reasons*. Let me elaborate on each of these elements.

Our *shared options* are the possible actions we choose between when deciding together what to do. If we are deciding where to meet for lunch, our shared options might include *meeting at Restaurant A* and *meeting at Restaurant B*. I will say more about shared options shortly (§1.2).

A *joint decision* is the conclusion of joint practical deliberation – an answer to the question of what we are to do. A joint decision will take one of our shared options as its content (say, to meet at Restaurant B). In the individual case, what makes a representation of my ϕing a *decision* to ϕ – rather than, say, a prediction that I will ϕ – is its distinctive effect on my thought and behavior. If I decide to ϕ, then I will be disposed to take the means I believe to be necessary to ϕing and avoid performing actions I believe to be incompatible with ϕing (Bratman 1987: 15-17). Moreover, I will be disposed to take it for granted that I will ϕ in my subsequent deliberation, no longer treating not-ϕing as a live option. Similarly, for a representation of a shared option to count as a joint *decision*, it must have implications for our thought and behavior. If we jointly decide that I will ϕ, I should be disposed to take the means I believe to be necessary to ϕing and to avoid performing actions I believe to be incompatible with ϕing. And I should be disposed to take it for granted that I will ϕ in my subsequent deliberation, no longer treating not-ϕing as a live option. A joint decision, then, is a representation of a shared option that commits each of us to deliberating and acting within its bounds.

Finally, shared reasons. Practical deliberation, whether individual or joint, is a *rational* process. Decisions are subject to normative standards: some decisions we ought to make, others we ought not. Akratic exceptions aside, when we deliberate about what to do, we try to make the decision we ought to make. In other words, we try to make a decision that is sufficiently supported

by our *normative reasons*. Similarly, when we deliberate together about what to do, we try to make a joint decision that is sufficiently supported by our normative reasons. Let us define two or more persons' *shared reasons* as the set of normative reasons that determine what joint decisions they ought to make. Whenever one extols the benefits of one's favored joint plan, or objects that a proposed joint decision is unfair or foolish, one is appealing to shared reasons.[5]

So: two or more agents engage in joint practical deliberation when they try to make a joint decision about which of several shared options to perform. If the agents are jointly *deliberating*, then they will try to make a joint decision that is justified by their shared reasons. When they make a joint decision, the agents will take this decision to constrain their subsequent deliberation and action in many of the same ways as their own individual decisions do.

I said earlier that this activity of joint practical deliberation is more flexible than it seems. It is now time to defend this claim. I submit that we should expand our commonsense conception of joint deliberation in two ways. First, we should make room not just for joint decisions about what *we* will do, but also joint decisions about what *you* will do or what *I* will do (§1.2). Second, we must distinguish between two different *methods* of making joint decisions, one of which has eluded most theorists' attention (§1.3).

1.2. Joint decisions about individual actions

The most obvious examples of joint deliberation concern joint action: where *we* will meet for lunch, whether *we* will have a child, whom *we* should hire. But on reflection, there is no barrier to our jointly deliberating about what *you* or *I* will do alone. On a camping trip, we might decide that *you* will build the fire, since you know how. Founding a startup, we might decide that *I* will call potential investors, since I am more extroverted than you are. Appreciating this point is the first step towards the deliberative theory, since promises concern only the speaker's actions.

The possibility of joint decisions about individual actions is entailed by a plausible account of shared options. In the individual case, a person's options are the potential actions she might perform. This suggests an idea: in parallel with the popular possible worlds framework for modeling

6

theoretical deliberation (e.g., Stalnaker 1984), we can model practical deliberation using a set of *possible actions* (cf. Gibbard 2003: Ch. 3). On this model, each of a person's options is a set of possible actions she could perform: for example, my option of *going to Restaurant B* is the set of possible actions in which I go to Restaurant B.

This model can be easily extended to shared options. Instead of considering possible actions one person can perform, we can consider the possible *combinations of actions* multiple persons could compatibly undertake. This will include not only paradigm joint activities, but also *any* possible combinations of the participants' individual actions, such as the combination in which I teach a class in Paris while you deliver a baby in Los Angeles. Then we can model our shared options as sets of possible combinations of our actions. So, the shared option of *our going to Restaurant B* is the set of possible action combinations in which we both go to Restaurant B.

This model makes it clear how we can make a joint decision about what *I* will do: we can decide upon a shared option that substantively constrains only my actions. Consider the set of possible combinations of our actions in which *I read your favorite book*. Jointly deciding on this set would significantly constrain my actions, excluding from consideration any possible combinations of our actions in which I don't read your favorite book. However, it would leave your actions largely unconstrained: any action you could perform that is compatible with my reading your favorite book is part of one of the possible combinations in this set. Thus it seems that nothing of substance is missing if we describe our joint decision simply as a decision *that I will read your favorite book*, omitting the clause 'and you will not prevent me from doing so.'

To deny the possibility of joint decisions like this, one would have to offer an alternative account of shared options that excludes options that constrain only one person's behavior while remaining flexible enough to capture the full range of possible joint activities. I don't see a principled way of doing this; nor do I see what would motivate the undertaking. The most natural model of shared options allows for joint decisions about what you or I will do alone.

1.3. Two methods of joint deliberation

Though we have outlined the structure of joint practical deliberation, we have said little about how this activity is actually performed. How do we go about deciding together what to do?

To count as an activity of joint practical deliberation, a process must meet two criteria. First, for us to count as engaged in joint practical *deliberation* (rather than some arational decision-making process), the process by which we make joint decisions must be guided and constrained by our shared reasons. Second, for us to count as engaged in *joint* practical deliberation (rather than each deliberating separately), the process by which we make joint decisions must be a *joint activity*: meaning, at minimum, that it must involve all of us playing some part.

One procedure that meets these two criteria is what I call the *propose-and-ratify* method. Propose-and-ratify proceeds in two steps. First, a speaker proposes a potential joint decision that she takes to be justified by the shared reasons. Second, the addressees judge whether the proposed joint decision is justified. If their judgment is favorable, the addressees voice their acceptance. A proposed joint decision comes into force just in case all participants explicitly accept it.

Propose-and-ratify is what first comes to mind when one thinks of joint decision-making. I say, "Shall we go to Restaurant B?" You reply, "Sure!" And a joint decision is made. But propose-and-ratify is not the only way we could go about making joint decisions. Note that, in propose-and-ratify, the default result is for a proposal to be *rejected*: if the addressees do not explicitly accept the proposal, no joint decision occurs. But we could also adopt a procedure in which the default result is for a proposal to be *accepted*. Call this the *propose-and-challenge* method.

Propose-and-challenge also begins with a speaker proposing a joint decision that she takes to be justified. But then, instead of explicitly ratifying the proposal if they think it is justified, the addressees are expected to *challenge* the proposal if they think it is *un*justified. A proposed joint decision comes into force just in case there have been no successful challenges to it (meaning either that no challenges have been raised, or that all challenges have been successfully rebutted). The absence of successful challenges signals that every participant takes the proposed joint decision to be justified by the shared reasons. Thus propose-and-challenge meets our first criterion. It meets

the second as well: each participant plays a role in the propose-and-challenge process by being prepared to raise any objections that come to mind. Even when a decision is proposed and then accepted with no further words, it is made jointly by all, since the addressees' silent acceptance signals that the proposal has passed a collective process of evaluation.[6]

Propose-and-ratify, with its call-and-response structure, is more easily recognizable as a method of joint deliberation. But propose-and-challenge seems to me just as worthy of the name. Propose-and-challenge is no less joint, and no less deliberative, than its more explicit twin. When we decide together what to do, we have a choice as to how to proceed: should the default be for a proposal to be accepted, or rejected? To exclude propose-and-challenge, one would need to provide some reason to think that we must *always* treat rejection as the default result.

Recognizing the existence of the propose-and-challenge method is the second step towards the deliberative theory. I noted earlier that promising appears to be more unilateral than the paradigm case of joint practical deliberation. That is because the paradigm case is deliberation via propose-and-ratify. Promises appear more unilateral because they are a case of deliberation via propose-and-challenge.

One more point before we move on. Just as we can make a new joint decision by means of propose-and-challenge or propose-and-ratify, we can also *retract* an already standing joint decision using these methods. Retracting a joint decision works the same way as making a new one. First a speaker proposes to retract a standing joint decision; then, the addressees assess whether the retraction proposal is justified and signal their acceptance, either explicitly (as in propose-and-ratify) or implicitly (as in propose-and-challenge). When a joint decision is retracted, the shared options will expand to include some of the possible combinations of actions that the decision had previously ruled out.

1.4. From joint practical deliberation to promises

One obvious difference between joint and individual practical deliberation is that, while individual deliberation can be performed in solitary thought, joint deliberation needs to happen in

9

conversation: we need to *communicate* with one another to deliberate together. This is why both of our deliberative methods start with the speech act of proposing a joint decision. Given how often we need to decide together what to do, this speech act should be quite familiar – so familiar, in fact, that we should expect to find a word for it in ordinary language. And since there are multiple ways of proposing joint decisions, we might expect our language to distinguish between them, giving different names to different types of proposals. This suggests that *several* recognizable everyday speech acts may be proposals in joint practical deliberation.

I submit that English distinguishes between such proposals on the basis of their *content* and their *intended method of evaluation*. First, we distinguish between proposals to make joint decisions concerning only the speaker's actions ("I will ϕ"), those concerning only the addressees' actions ("you will ϕ"), and those concerning both the speaker and addressees' actions ("we will ϕ"). Second, we distinguish between proposals that are meant to be evaluated by propose-and-challenge and those meant to be evaluated by propose-and-ratify. My conjecture is that our everyday speech act categories map on to these distinctions as follows:

	Propose-and-challenge	**Propose-and-ratify**
"I will ϕ"	Promises	Offers
"You will ϕ"	Demands	Requests
"We will ϕ"	Agreements[7]	

And so we return to the deliberative theory: promises are proposals to make joint decisions of the form "I will ϕ" by means of the propose-and-challenge method. They are distinguished from demands, requests, and agreements by their content: promises concern the speaker's actions, while demands and requests concern the addressee's actions, and agreements concern the actions of both speaker and addressee. Promises are distinguished from offers by their intended method of

evaluation: promises employ the propose-and-challenge method, while offers use the propose-and-ratify method. This captures the fact that offers require explicit acceptance to come into effect, while promises are accepted more implicitly. If I offer to drive you home, and you do not explicitly accept my offer, then I am under no obligation to drive you home. But if I *promise* to drive you home, you don't need to say anything: so long as you do not object, I am bound (cf. Robins 1984: 101; Darwall 2011: 273).

The above table outstrips the ambitions of this paper. I will not attempt here to defend the claim that all of these phenomena are best understood in terms of joint practical deliberation. Instead, I will focus on defending the proposed analysis of promises:

> **The Deliberative Theory of Promises:** For S to promise A that she will ϕ *just
> is* for S to propose to A, by means of the propose-and-challenge method, that they
> make a joint decision to the effect that she will ϕ.

2. Similarities between promises and proposals in joint practical deliberation

My thesis is that promises are proposals to make joint decisions regarding one's own actions using the propose-and-challenge method. While we evaluate this thesis, however, it will be helpful to have a separate term for the relevant kind of proposal, so as to avoid begging the question regarding its relation to promises. Let us coin the term *I-proposal* to refer to a proposal to make a joint decision regarding one's own actions ("*I* will ϕ") by means of propose-and-challenge. This allows us to state the deliberative theory more succinctly: promises are I-proposals. My argument for this thesis is founded on the claim that I-proposals have the same properties as promises. Here I defend this claim, surveying five of promises' most important features and arguing that I-proposals have each of these features as well.

2.1. Bindingness

Promises create obligations. When I promise you that I will read your paper, I thereby become *obligated* to read your paper. However, while promises typically generate obligations, this

connection is defeasible (*pace* Gilbert 2011: see Introduction). It is sometimes permissible to break a promise: if I promise to come to your talk, but then come across an injured person who urgently needs to be taken to the hospital, I am plausibly permitted to miss your talk and help the person instead. More carefully, we can say that when I promise to φ under normal conditions,[8] I come to have a strong *pro tanto* reason to φ that typically, though defeasibly, obligates me to φ.

Promises don't just create obligations *simpliciter*. They create *directed* obligations: obligations that are owed *to the promisee* in particular. When I promise you that I will read your paper, I become obligated *to you* to do so. My failing to read your paper would wrong *you*. You thus have special standing to demand that I fulfill my promise and resent me if I break it (Gilbert 2004; Darwall 2011).

Our case judgments indicate that joint decisions, too, create obligations, and that these obligations are also directed. Suppose that you and I jointly decide to read each other's papers. You carefully read and write comments on my paper, and come to our meeting to find that I have not even glanced at yours. You would rightly feel wronged. Not only have I done wrong; I have wronged *you*. So, when we make a joint decision under normal conditions, we each become obligated to the other to do our part in seeing it through. As with promises, this obligation can be overridden – if enough is at stake, I may be permitted to violate our joint plan – but even when it is, the fact that we made a joint decision is a strong *pro tanto* reason in favor of my carrying it out.

If you grant that our joint decision is binding in the example above, then it is hard to deny that joint decisions made by means of I-proposals are too. The fact that a joint decision concerns only the speaker's actions, or that it was proposed by means of the propose-and-challenge method, does not seem to diminish its moral relevance. So, if we jointly decide just that *I* will read your paper, and we decide this by means of propose-and-challenge, this joint decision will similarly obligate me to you to read your paper. Thus I submit that I-proposals, like promises, (defeasibly) generate obligations that are owed to their addressees.

(You might be worrying that the obligation to abide by a joint decision seems *weaker* than the obligation to keep a promise. This objection is important enough to get its own section (§3); until then, let us set it aside.)

I will leave it as an intuitive datum that joint decisions are binding; I won't try to provide a deeper explanation of why this is so. This may be surprising. Many papers on promising take their primary question to be, "Why should we keep our promises?" The deliberative theory offers only a partial answer to this question. It says that we should keep our promises because we should abide by our joint decisions – and stops there. As we are about to see, the deliberative theory's explanatory power lies elsewhere. If we grant the premise that joint decisions are binding, then by combining this premise with the deliberative theory, we can explain many of promises' other features: their effects on deliberation, their uptake and validity conditions, and so on. The question that the deliberative theory helps us answer is thus not "Why are promises binding *at all* (rather than not)?" but instead "Why do promises bind *in this particular way* (rather than some other)?" In other words, the deliberative theory aims to explain not the *existence* of promissory obligation, but its *shape*. Of course, if we were to discover some deep moral argument that explained the obligation to abide by joint decisions, this would only augment the deliberative theory's power. But my argument in this paper does not depend on the availability of such an account. All it requires is the intuitive premise that, when you and I make a joint decision, I typically become obligated to you to do my part in carrying it out.[9]

2.2. Exclusion

Most agree that promissory obligations can be overridden if enough is at stake. But some have argued that it is still inappropriate for a promisor to think of the fact that she promised to ϕ as *just another consideration* to be weighed up when deciding whether to do so (Raz 1977; Hart 1982: 255; Robins 1984; Owens 2012: 89-91). Why? Because a promise to ϕ should *settle the question* of whether to ϕ for the promisor. To weigh up reasons for and against keeping one's promise is to open a question that one should be treating as closed.

The attitude one should take towards actions one has *promised* to do is tellingly similar to the attitude theorists of intention have argued one should take towards actions one *intends* to do (Bratman 1987; Holton 2004). Like promises, intentions settle deliberative questions, excluding options incompatible with the intended action from further consideration. Moreover, intentions resist reconsideration: having decided to φ, a rational agent should be disposed not to reopen the question of whether to φ unless her situation changes in some unexpected way.

Given the analogy we have pursued between joint and individual deliberation, we can expect these observations to apply to joint intentions as well. So, just as I should be disposed to exclude options incompatible with my standing decisions from my deliberation, I should similarly be disposed to exclude options incompatible with our standing *joint* decisions from my deliberation. And just as I should not be over-ready to reconsider my intentions, I should not be over-ready to reconsider our joint decisions. If I fail to have these dispositions, this means that I am not taking our joint decisions as seriously as I take my own individual intentions. Thus we can predict that a successful I-proposal, by bringing a joint decision into effect, will typically make it inappropriate for me to seriously deliberate about whether to comply with that joint decision. In other words, I-proposals, like promises, will have exclusionary force.

The deliberative theory thus enables us to *explain* the exclusionary force of promises by appeal to the exclusionary force of intentions (cf. Raz 1975: 69-70; Robins 1984). I should not think of my promise as just another consideration to be weighed up in deliberation because, in doing so, I fail to treat the joint decision my promise brought into force as what it is: a *decision*.

2.3. Uptake

Many have observed that promises require *uptake* from the promisee to be binding (Thomson 1990: 297; Darwall 2011: 268; Owens 2012: 224-226; Roth 2016: 89-92). If I make a promise that you neither hear nor acknowledge, then my promise misfires, failing to obligate me at all. But it is hard to say exactly what kind of uptake promises require. For promises do not seem to require explicit affirmation from both parties in the way agreements, offers, and requests do.

Promises seem *unilateral* in a way that offers or agreements are not. On the other hand, the promisee must recognize a promise for it to be binding, and can block a promise by refusing it.

Notably, both of these claims characterize I-proposals as well. Any proposed joint decision comes into force only if its addressees recognize it and signal their acceptance. But in the propose-and-challenge method, a proposal need not be *explicitly* accepted to come into effect. All that is required is for the addressee to recognize it and refrain from challenging it in good faith. (I say more about good faith in §2.5). Thus I-proposals require the same kind of uptake that promises do: recognition and the absence of objection. If we take on the deliberative theory, then, we can explain why promises seem unilateral: because they use propose-and-challenge.

This should help to assuage one *prima facie* worry about the deliberative theory. One objection I hear often is: "But promising doesn't seem like making a *proposal*." True enough: on the ordinary use of 'proposal', a proposal must be explicitly accepted to come into force. Still, I call promises 'proposals' because they *do* require their addressees' acceptance to come into force; this acceptance is just usually tacit. Compare assertion: on a popular account, an assertion is "a proposal to change the context by adding [its] content to the information presupposed" (Stalnaker 1999: 10). But assertions, like promises, do not need *explicit* acceptance to come into effect – they only need their addressees to recognize them and refrain from objecting. So, promises are proposals in the same technical sense that assertions are.

2.4. Release

Suppose I promise my partner Dee Dee that I'll wash the dishes tonight. But before I pick up the sponge, Dee Dee says, "On second thought, don't worry about doing the dishes. You should relax instead." Dee Dee has *released* me from my promise: I am no longer obligated to do the dishes. The moral situation is much like it would have been if I had never promised at all. In general, promisees have the power to release promisors from their promissory obligations.

The addressee of an I-proposal has a similar power: he can propose to retract the joint decision the I-proposal brought into effect. When we retract a joint decision, its normative effects

15

are reversed: our reason to abide by our standing joint decisions no longer gives us reason to abide by *this* decision, as it no longer stands. So, the addressee of an I-proposal can release the speaker from her obligation by proposing to retract the joint decision the I-proposal brought into force.

But this seems to miss an essential detail. *Any* participant in joint practical deliberation can propose to retract a joint decision, including the person who originally proposed it. The power to retract joint decisions thus seems to be symmetrical. In contrast, the power to release someone from a promise is *asymmetrical*. A promisor cannot release *herself* from her promise; only the promisee can release her. How can the deliberative theory explain this asymmetry?

The answer is that both promisor and promisee *can* propose to retract the joint decision brought into force by a promise, but their proposals should be evaluated using different methods. It is good deliberative practice to evaluate the initial speaker's proposals to retract an I-proposal by the propose-and-ratify method, while evaluating the addressee's retraction proposals by the propose-and-challenge method. This helps correct for both parties' natural biases. The speaker is likely to be overeager to retract the joint decision brought into effect by her I-proposal, since the costs of fulfilling it will fall primarily on her. So, it makes sense to evaluate the speaker's retraction proposals conservatively, requiring the addressee's explicit acceptance for them to succeed. In contrast, the addressee may be reluctant to retract the joint decision brought into force by an I-proposal, since often he will stand to benefit from the I-proposal's fulfillment at little cost to himself. And so it makes sense to evaluate the addressee's retraction proposals using the more liberal propose-and-challenge method. We should expect our practice of joint practical deliberation to reflect these considerations, incorporating a norm on which we evaluate retraction proposals by different methods depending on who made them.

On the deliberative theory, then, what is unique about the promisee's power of release is not that she alone can propose to retract the promise, but that only her proposals are evaluated by means of propose-and-challenge. When Dee Dee proposes to retract our joint decision ("don't worry about doing the dishes tonight"), I need not explicitly accept her proposal for the decision to be retracted.[10] But suppose instead I turn to Dee Dee and say, "I'm pretty tired. Would you

mind if I left the dishes until tomorrow?" I thereby propose to retract the decision brought into force by my earlier promise. But notably, my proposal will only succeed with Dee Dee's explicit blessing. If she says, "Sure, that's fine," our joint decision is then retracted, and I am released from my promise. But if she rejects my proposal ("I'm sorry, but the dishes need to be done tonight"), my promise remains in force. This is just what we should expect if promisors' retraction proposals are evaluated using the propose-and-ratify method (cf. Gilbert 2011: 99-100).

This asymmetry in deliberative methods guarantees that the speaker of an I-proposal will be released from her resulting obligation *only if* her addressee explicitly agrees to her release: either by proposing to retract the decision himself, or by explicitly accepting the speaker's proposal to do so. No such power is given to the speaker: the addressee does not require the speaker's explicit acceptance to release her. Thus just as promises give promisees the power of release, I-proposals give their addressees a tellingly similar power to reverse their effects.

2.5. Invalidation by coercion and deception

Sometimes, promises don't bind at all. A person promises to ϕ, and yet, due to unusual circumstances, her promise fails to generate any reason for her to ϕ (at least in the usual way). Such promises are called *invalid*.[11] The paradigm cases of invalid promises are those elicited by coercion and by deception, as in these colorful examples from Judith Jarvis Thomson:

> **BANK**: "[Bert] is not among the more efficient extortionists: he holds a gun to [Anna]'s head and says 'GIVE ME YOUR WORD THAT YOU WILL GO TO YOUR BANK AND FETCH ME BACK A THOUSAND DOLLARS OR I'LL SHOOT YOU!' [Anna] says 'Yessir'" (Thomson 1990: 310).

> **VAN GOGH**: Diego offers to sell Cora his Van Gogh for $1000, and Cora promises to go to the bank and get $1000 to pay him straightaway. On the way, Cora learns that the painting is forged, and Diego knew this was so (paraphrase of Thomson 1990: 312-313).

The coercion in BANK and deception in VAN GOGH somehow make the promises misfire: neither promisee would be wronged if the promisor did not give him the promised money.

I claim that I-proposals induced by coercion and deception also misfire: they fail to result in the making of *any joint decisions at all*. To see why, note first that it is possible for us to *appear* to make a joint decision without actually making one. Consider two actors who pretend to make a joint decision in a play. If one were to hold the other to their 'joint decision' afterwards, she would clearly be confused – no joint decision was made.

Sometimes, the pretense is less public. One person may think she is engaged in joint deliberation while the other is merely pretending to deliberate with her. To count as deliberating with you, I must evaluate potential joint decisions on the basis of our shared reasons and propose or accept only those joint decisions that I take to be justified by those reasons. Suppose instead that I ignore our shared reasons and instead simply propose and accept joint decisions based on whether their acceptance will suit my purposes. If I behave in this way, I am not deliberating with you; I am only *pretending* to do so.

Let us call this kind of pretending *deliberating in bad faith*, and its opposite – genuine participation in joint deliberation – deliberating in *good faith*. When I deliberate in bad faith, I am not jointly deliberating at all: I am just pretending to do so. This failure to participate robs our interaction of the mutuality that is essential to joint decision-making. The idea of a joint decision implies a meeting of the minds; but if I am deliberating in bad faith, no meeting of the minds occurs. We are not deciding *together*, and so we cannot *decide* anything together. My bad faith thus renders our joint deliberation impotent: so long as I am deliberating in bad faith, we cannot make any joint decisions, though we may falsely appear to do so (cf. Ripstein 2009: 131).

How do we tell whether someone is deliberating in bad faith? Here's a heuristic: if a person knows that a joint decision is unjustified, and yet she intentionally refrains from challenging it, then that probably means that she is deliberating in bad faith. This is what happens in our cases of coerced and deceived promises.

Start with deception. In VAN GOGH, Cora promises to pay Diego $1000 for his painting. On the deliberative theory, Cora's promise is an I-proposal: a proposal to jointly decide that she will go fetch $1000 in payment for Diego's painting. Cora only makes this proposal because she

believes that the painting is a genuine Van Gogh; but Diego knows it to be a forgery. So, the reason Cora takes to justify jointly deciding that she will fetch Diego $1000 – namely, that he will give her a Van Gogh in return – is not a *shared* reason for Diego and Cora, since Diego does not believe it to be true. Without the premise that the painting is a Van Gogh, Cora's proposed joint decision is not justified: a forged Van Gogh is not worth $1000 to Cora.

Diego knows all of this. So, if he were deliberating with Cora in good faith, Diego would challenge her promise as follows: "You are only promising to pay me $1000 for this painting because you believe it is a genuine Van Gogh. But I lied to you – the painting is a forgery. Given that the painting is a forgery, you shouldn't pay me $1000 for it, and so I can't accept your promise to do so." Of course, our duplicitous Diego is unlikely to have this sudden change of heart. But that's the point: the fact that Diego does *not* raise this challenge shows that he is not participating in the propose-and-challenge process in good faith, and this is what invalidates Cora's I-proposal. The proposal fails to bring a joint decision into force because it is not accepted by means of good faith joint deliberation.

Coerced promises are invalid because they are premised on the promisee's illegitimate threat. In BANK, Bert's threat leaves Anna with a choice between promising Bert $1000 and being shot. Understandably, Anna opts for the former. But Anna's promise is premised on an unwarranted restriction of her options. Anna can rightfully demand that Bert not shoot her *whether or not* she promises him $1000. Thus her choice should not be between *promising Bert $1000* and *being shot by Bert*, but instead between *promising Bert $1000* and *not promising Bert $1000*, with Bert's shooting her squarely off the table. In the terms of the deliberative theory, Anna's I-proposal is justified only relative to an *unjustified* set of shared options – namely, the set of shared options that excludes Anna's retaining both her money and her life (compare Frick 2016: 237-238).

Bert is in a position to know all of this. So, if he were deliberating with Anna in good faith, Bert would challenge her promise as follows: "You're only promising to go fetch me $1000 because I threatened to shoot you if you didn't. But I should never have threatened to shoot you in the first place – doing so would be wrong. So let's agree that I won't shoot you, and then you can decide

whether or not to promise me $1000 on your own." Of course, our belligerent Bert won't raise this challenge unless he has a miraculous change of heart. But the fact that Bert *doesn't* raise this challenge, which he knows to be justified, shows that he is not deliberating with Anna in good faith. So the proposed joint decision that Anna will go fetch Bert $1000 does not come into force, as it was not jointly made. This is why Anna's promise is invalid.

3. An objection: are joint decisions weaker than promises?

Suppose the arguments so far have convinced you that promises and joint decisions have a shared pattern of normative effects. Still, you might have a nagging feeling that promises are somehow more *forceful* than joint decisions. To make a promise is to undertake a solemn moral obligation, which would be a serious wrong to break. But making a joint decision seems to be a less serious matter. To make this objection vivid, consider two examples:

> **PLANNED WALK:** One Saturday morning, I propose to my partner Dee Dee that we take a walk in the afternoon. She agrees. Later, I tell Dee Dee, "Sorry, I can't go walking after all – I need to finish writing this section."

> **PROMISED WALK:** Another Saturday, I propose an afternoon walk to Dee Dee; she agrees. Then, remembering my tendency to abandon our walking plans, Dee Dee says: "Okay, but I want to know that you won't change your mind this time. Do you *promise* you'll go on a walk with me this afternoon?" I say, "Yes, I promise."

Granting that joint decisions generate obligations, the obligations Dee Dee and I undertake to one another in PLANNED WALK still appear fairly weak. If either of us simply forgot to make time for the walk, it wouldn't be a big deal. It also appears to be rather easy to get out of our obligations: telling Dee Dee that I'm busy writing seems to be enough to get me off the hook. In contrast, it *would* be a big deal if, after I promise to go on a walk with Dee Dee in PROMISED WALK, I failed to do so. And I can't get out of this obligation simply by informing Dee Dee that I need to do my work instead; I am bound unless she explicitly releases me from my promise. So, my obligation in

PROMISED WALK is stronger than my obligation in PLANNED WALK in at least two ways: it is more important that I fulfill it, and it is more difficult to extract myself from it.

Prima facie, this contrast appears flatly incompatible with the deliberative theory. The deliberative theory seems to imply that the obligations in PLANNED WALK and PROMISED WALK should be identical. After all, it holds that there is no difference between my promising Dee Dee that I will go on a walk with her and my proposing to jointly decide that I will do so. Worse, the deliberative theory doesn't seem able to even *make sense* of the interaction in PROMISED WALK. For Dee Dee asks me to promise to do something that we've *already* jointly decided that I will do. On the deliberative theory, this request seems redundant, perhaps even incoherent. While if promising is something different from proposing a joint decision, Dee Dee's request makes perfect sense. Thus it seems the only way to explain the contrast between our examples is to accept that promises are not proposals to make joint decisions.

Not so fast! Appearances notwithstanding, the deliberative theory has the resources to explain the contrast between PLANNED WALK and PROMISED WALK. The trick is to recognize that some joint decisions generate stronger and stickier obligations than others. What makes the obligation in PROMISED WALK stronger than the obligation in PLANNED WALK, I claim, is that the joint decision generated by my promise in PROMISED WALK is *more resolute* and *harder to retract* than the joint decision Dee Dee and I make in PLANNED WALK. Let me elaborate on these two differences.

Begin with resolve. Consider the difference between casually planning to make spaghetti for dinner and solemnly deciding (say, as a New Year's resolution) to go to the gym three times every week. The latter decision is more *resolute* than the former: one is more committed to seeing it through. I might change my dinner plans on a whim; but I will be reluctant to give up my resolution to go to the gym, and disappointed in myself if I do. Similarly, joint decisions can be more or less resolute. Some joint decisions, like a plan to go for a walk this afternoon, are fairly casual; others, like the decision to get married, are more serious. The more resolute our joint decision, the more reluctant we will be to reconsider it, and the more disappointed we will be in ourselves (and each

other) if we fail to see it through. Resolve thus understood is a psychological concept. It is a feature of the attitude we take towards our (joint or individual) decisions. But it has a direct normative upshot: plausibly, more resolute joint decisions generate stronger reasons to see them through.

A second way in which joint decisions can differ is in their *retraction conditions*. As I argued in §2.4, the joint decisions brought into force by promises have *asymmetric* retraction conditions. While the promisee can propose to retract the joint decision using propose-and-challenge, the promisor can only propose retraction by means of propose-and-ratify. By design, this asymmetry of methods makes it more difficult for the promisor to get out of her obligation, but relatively easy for the promisee to release her. In contrast, most joint decisions that concern *both* parties' actions – i.e., agreements – will have symmetrical retraction conditions. A casual agreement might allow both of us to propose retraction using propose-and-challenge; a more serious agreement might require us both to propose retraction using propose-and-ratify.

Here is how I suggest we understand the contrast between our cases. In PLANNED WALK, the joint decision Dee Dee and I make to go on a walk is low in resolve and easy to retract. Since our decision is not very resolute, it generates a relatively weak obligation to see it through – which is why it is not a big deal if we fail to do so. And since there's no strong reason to avoid reconsidering our plan, it seems appropriate for either Dee Dee or I to propose to retract our joint decision using propose-and-challenge. I am doing exactly this when I say "Sorry, I can't go walking after all" – and since Dee Dee does not object, our plan is thereby withdrawn.

In PROMISED WALK, our initial joint decision (before Dee Dee asks me to promise) is just like the decision in PLANNED WALK. But Dee Dee realizes that so long as our plan is casual and easy to retract, I am likely to abandon it. So she asks me to promise. The effect of my promise is to bring into force a joint decision that has the same content as our prior decision (that we will go on a walk), but that is more resolute and harder to retract. Since the resulting decision is more resolute, my obligation to go on a walk with Dee Dee is stronger, and can't be overridden by (e.g.) my desire to finish some writing task. And our joint decision now has the asymmetric retraction conditions characteristic of promises. I now need Dee Dee's explicit permission to be released from my

obligation; I can't get off the hook just by saying "Sorry, I can't go." So, when Dee Dee asks me to promise, she is really asking me to re-propose our joint decision with greater resolve and asymmetric retraction conditions. By endowing our joint decision with these features, my promise generates an obligation that is stronger and stickier than the one I had before.

Zooming out from these examples, I propose the following hypothesis: in general, joint decisions made by means of promises will tend to be more resolute and difficult to retract than joint decisions made via agreements. This is because promises and agreements tend to serve different social functions. Since agreements concern multiple agents' actions, their main function is to help us plan and coordinate joint actions. Since a joint decision doesn't need to be high in resolve or resistant to retraction to serve this function, it is common to make agreements that are low in resolve and easy to retract (such as the agreement in PLANNED WALK).

In contrast, since promises concern only one agent's actions, they don't automatically earn their keep by enabling coordination. Instead, the primary function of a promise seems to be to *commit* the speaker to a course of action – often, as many have emphasized, to assure the addressee that it will be done (see §4.2 below). This committing function will be best served by a joint decision that has high resolve (so as to strengthen the speaker's reason to perform the action) and asymmetric retraction conditions (so as to make it difficult for her to back out). And so we can predict that promises will usually result in joint decisions that are high in resolve and difficult to retract (such as the decision in PROMISED WALK).

The worry with which we began was that promising is a more serious matter than joint decision-making. I've suggested that this objection stems from a failure to appreciate the versatility of joint decisions. Some joint decisions, especially those that play a purely coordinative function, are casual and easy to retract. Others – particularly those that, like promises, aim to establish commitments – are more serious and difficult to take back. Recognizing these dimensions of variability is the key to defusing the objection. Joint decisions are flexible enough to ground both the weak and revisable obligations of everyday plans and the solemn and stubborn obligation of a promise.

4. Other theories of promising

We are now in a position to compare the deliberative theory to alternative theories of promising. This section considers the three most prominent theories in the literature: conventionalism, the expectation theory, and the normative powers theory.

4.1. Conventionalism

Conventionalism is the view that promising is a move in a conventional social practice. Following John Rawls, we can understand a *practice* as "any form of activity specified by a system of rules which defines offices, roles, moves, penalties, defenses, and so on, and which gives the activity its structure" (1955: 3). Examples include baseball, voting, weddings, jury trials, and queuing in a line. The conventionalist locates promising within a 'promising practice' that has two roles (promisor and promisee), two moves (promising and releasing), and one rule: if a promisor promises a promisee that she will ϕ, then the promisor must ϕ unless (i) the promisee releases her from the promise or (ii) certain invalidating conditions (e.g. coercion or deception) obtain (Jones 1966: 288; Rawls 1971: 344-345; Kolodny and Wallace 2003: 120). This allows the conventionalist to ground the obligation to keep promises in a wider obligation to abide by the rules of valuable practices (Rawls 1971: 342-50).

After a long reign as orthodoxy,[12] conventionalism has fallen out of favor, largely due to the objections raised by T. M. Scanlon in his agenda-setting "Promises and Practices" (1990).[13] This raises a worry for the deliberative theory. Like conventionalism, the deliberative theory also holds that promising is a move within a social practice: the practice of joint practical deliberation. So we must consider whether the objections that toppled conventionalism also cause problems for the deliberative theory. I will discuss the two objections that have been most influential.

The first objection claims that conventionalism cannot explain why promises generate obligations that are owed *to the promisee* in particular (Scanlon 1990: 221; Darwall 2011: 263-264; Owens 2012: 135). For, the argument goes, conventionalism says that promise-breaking is wrong

because it violates the rules of a certain practice. But breaking the rules of a practice plausibly wrongs *everyone* who participates in that practice (perhaps except oneself). So, if the relevant practice is that of promising generally, then breaking a promise seems to wrong *everybody* who has ever made a promise, not just the promisee.

This argument hinges on the assumption that there is a single promising practice in which all promising agents participate. The conventionalist should reject this assumption. Consider the practice of playing chess. Each pair of chess players has their own distinct game of chess; they are not all playing a single enormous game. So, if I cheat in our game of chess, I *do* wrong everyone else who participates in this practice – but since you are the only other player in our game, this means that I wrong you and you alone. To assume that breaking the rules of chess must wrong every chess-player that ever lived is to conflate *types* of practices with *token* instances of a practice type. Every chess player participates in the type of practice, chess, in which we are engaged; but they do not all participate in our token game of chess.

The conventionalist should say that promising, like chess, is (usually) a two-person game. (If a promise has two addressees, it is a move within a three-person practice, and so on). When I break my promise to you, I am violating the rules of our two-person practice of making promises to one another. My violation *does* wrong everyone else who participates in this practice – but since you are the only other participant in our token promising practice, this means that I wrong you and you alone. The deliberative theory can co-opt this response: breaking my promise wrongs you in particular because it violates the norms of our two-person practice of joint practical deliberation.

The second objection charges that conventionalism makes the existence of promising too *contingent*. Conventionalism entails that our ability to make promises depends upon the prior existence of a social practice of promising. Since we could conceivably fail to have a practice of promising, conventionalism seems to imply that we could lack the ability to make promises to one another. Yet many have thought that morally competent persons *necessarily* have the power to make binding promises (Locke 1689/1980: 13; Raz 1977: 214-215; Scanlon 1990: 201; Thomson 1990: 303-304; Shiffrin 2008).

However, the conventionalist is only vulnerable to this objection if she bundles together two claims that can and should be separated: the claim that promising is a move within a social practice, and the claim that this practice is *optional*, in the sense that social life would be possible without it. It is easy to run these claims together, because it is natural to assume that *all* social practices must be optional. But this assumption is simply false.

Consider *assertion*. Assertion has all the signature features of a move in a social practice. It is governed by a rule that dictates which assertions are permissible – the much-discussed 'norm of assertion' (Williamson 2000). Beyond this norm, assertions are subject to many other 'maxims' determining when they are relevant or cooperative (Grice 1989). And, as in all practices, what assertions are permissible in a conversation depends upon what other 'moves' have already been made (Lewis 1979). No wonder philosophers studying assertion, just like those studying promises, have felt the irresistible pull of the analogy with baseball (compare Rawls 1955 with Lewis 1979).

Yet few are tempted to say that the assertion practice is socially optional. The ability to share information in the way assertion makes possible seems to be a precondition for cooperative society. And the norms governing assertion are widely thought to be "derived from general principles of rationality, cooperation and/or cognition" (Korta & Perry 2015). Thus though assertion is a move in a practice, this practice is not *conventional* in the Lewisian sense of the term: it is not one of *multiple different* possible solutions to a coordination problem (Lewis 1969/2002: 70). The practice of assertion, with just the rules it has, is likely the unique solution to the problem of sharing information between persons.

Taking assertion as our model, then, we can make room for a *nonconventionalist practice theory* of promising. Such a theory would agree with the conventionalist that promising is a move in a practice, but depart from classical conventionalism in claiming that this practice is not socially optional. The deliberative theory occupies this relatively uncharted area of logical space. Like conventionalism, the deliberative theory claims that promises are a move within a practice. But unlike conventionalism, the deliberative theory does not take the power to promise to be "activated only by the happenstance that a social convention of promising has developed" (Shiffrin 2008:

483). For the practice within which it locates promises – joint practical deliberation – serves a function essential to social life: it enables us to plan and act together on terms that all can accept. It is not a contingent fact that morally competent persons engage in joint practical deliberation; and thus it is no more contingent that we have the power to promise.

4.2. The expectation theory

As an alternative to conventionalism, Scanlon proposed what has come to be known as the *expectation theory* of promising (1990; 1998).[14] On this view, the obligation to keep our promises derives from a practice-independent obligation to not disappoint expectations we have deliberately created in others. When I promise you that I will φ, I give you *assurance* that I will φ, leading you to form a strong expectation that I will do so; having deliberately raised this expectation, it would be wrong for me to violate it.

The expectation theory has substantial intuitive appeal. One of the main reasons why we make promises is to give others assurances that they can rely upon. If you are worried about whether I will actually show up to our meeting, I can reassure you by promising you that I will. The fact that I promised to come to our meeting justifies you in having a particularly strong expectation that I will do so, which you can then rely upon in planning your future conduct.

However, the expectation theory doesn't just say that one important function of promises is to provide assurance; it says that *all it is* to promise is to give someone assurance. And as the substantial literature in response to Scanlon's proposal has revealed, there is reason to doubt this stronger claim. The obligations one undertakes by deliberately raising expectations in others seem importantly different from the obligations one undertakes in making a promise. If I inform you of my future plans to φ, but then change my mind, I can discharge my obligations to you by simply warning you that I won't φ, and perhaps compensating you for any losses you incurred by relying on my testimony. In contrast, if I promise you that I'll φ, merely warning you that I've changed my mind and compensating your losses is not enough: I am obligated to φ, pure and simple (Owens 2006: 53-56; Southwood and Friedrich 2009: 266-271). This difference does not seem to be

accounted for by supposing that promises give rise to especially strong expectations. For I can tell you that I will ϕ with as much confidence as I like, and yet still coherently add the disclaimer, "but I don't promise to ϕ" (Raz 1972: 99; 1977: 216; Robins 1984: 9; Owens 2006: 60; Darwall 2011: 267). In such a case, it seems that my disclaimer succeeds: I have no promissory obligation to ϕ. This seems to show that I can knowingly and voluntarily lead you to form a very strong expectation that I will ϕ without thereby promising to do so.

Whether these objections are fatal for the expectation theory is, of course, a matter of debate. I won't contribute to that debate here. Rather than try to refute the expectation theory, my aim is just to highlight the appeal of my proposed alternative. To that end, I will show how the deliberative theory can capture both the observation that motivates the expectation theory – that promises provide assurance – and the observation that has motivated some to reject it – that promises are distinct from testimony.

Begin with the assurance-providing role of promises. Our jointly deciding that I will ϕ gives you a particularly strong basis for forming an expectation that I will ϕ. First, it (normally) puts me under an obligation to ϕ, one that I owe *to you* in particular (§2.1). Our joint decision thus generates a new, and weighty, reason for me to ϕ. The more resolute our joint decision, the stronger this reason will be – which is why joint decisions that aim to provide assurance will typically be quite resolute (see §3). Second, the norms of joint practical deliberation require me to treat our decision that I will ϕ as I do my own intentions, excluding the possibility of not ϕing from my future deliberation and not being over-ready to reconsider this plan (§2.2). Finally, since joint decisions cannot be retracted unilaterally, I cannot simply change my mind about whether to ϕ without violating my obligation to you – even if I give you fair warning. If I want to retract our joint decision, I must propose to do so by means of propose-and-ratify, thereby giving you the final say over whether our joint decision remains in force (§2.4). These factors combined make joint decisions a powerful tool for giving others assurances about what we will do.

Crucially, however, the deliberative theory clearly distinguishes promises from testimony regarding one's future actions. Proposing to jointly decide that I will ϕ is different from telling you

or otherwise leading you to believe that I will ϕ. I can coherently tell you that I will ϕ while explicitly denying that I am proposing to jointly decide that I do so.

The deliberative theory sees the distinction between promises and testimony regarding one's future actions as the interpersonal analogue of the distinction between *intention* and *belief* regarding one's future actions. Consider the difference between my *intending* to finish this paper by May and my *believing* that I will do so. Suppose that, midway through April, I notice that I am not on track to finish the paper before May. If I *intend* to finish the paper by May, I am committed to making this happen, and so need to change my behavior – e.g., working longer hours. While if I merely *believe* that I will finish the paper by May, I need only adjust my beliefs to reflect my new evidence – e.g., coming to believe that I'll finish the paper in June.

Promises and testimony are exactly parallel. If I promise you that I will finish the introduction to our coauthored paper by May, the resulting joint intention commits me to doing what is required to finish the introduction by May. In contrast, if I merely *tell* you that I will finish the introduction by May – clarifying that this is a prediction, not a promise – then the result of my testimony is not a joint intention, but instead something like a shared belief. So if it begins to look like I won't finish by May, then all I need to do is inform you that the introduction is taking longer than expected. The different obligations attending promises and testimony are thus a direct consequence of intention and belief's different directions of fit: we change the world to fit our intentions, while we change our beliefs to fit the world. Promises bind us to *act*, not just to warn, because they result in *decisions*, not just in beliefs.

4.3. The normative powers theory

As both conventionalism and the expectation theory have come under criticism, a third view has gained prominence: the *normative powers theory* (Raz 1972; 1977; Owens 2006; 2012; Shiffrin 2008; 2011). The normative powers theory rejects the idea that promissory obligation needs to be explained by appeal to more general moral principles such as those governing our participation in social practices or our care for others' expectations. Instead, the obligation to keep

promises is "morally fundamental" and "derive[s] directly from the expression of an individual's will to be bound" (Shiffrin 2008: 482).

The normative powers theorist offers an alternative strategy for explaining promissory obligation. We are obligated to keep our promises, she argues, because it is *in our interests* to have the ability to obligate ourselves by making promises. The ability to bind ourselves by promising is essential to human social life. If we lacked this capacity, we would be much worse off. On the normative powers view, the fact that having the ability to make binding promises is in our interests is sufficient to make it the case that we *do* have this ability. Morality conspires to give us the tools that we need to live together, and one of these tools is the power to promise.

This explanatory strategy has traditionally been combined with a particular hypothesis about the interests that promises serve. Most normative powers theorists are united in rejecting the expectation theorist's idea that the function of promises is to provide assurance (Raz 1972: 99; Owens 2006). Instead, they argue, the purpose of promising is to grant others *authority* over one's actions (Owens 2006). When I promise you that I will ϕ, I give you the authority to demand that I will ϕ and to rebuke me if I fail to do so. Moreover, I give you the authority to determine whether I will continue to be bound to ϕ by deciding whether to release me from my promise. The power to transfer authority in this way plays an important role in making valuable forms of human relationship possible (Shiffrin 2008). These benefits of transferring authority, in turn, ground the normative power to promise.

Understood this way, the normative powers theory is compatible with the deliberative theory. The two could be adopted together. On this combined 'deliberative normative powers theory', the normative power to promise arises from the wider normative power to make binding joint decisions. The obligation to abide by joint decisions, in turn, would be explained by appeal to our interest in being able to morally bind ourselves to the results of our joint deliberation. One (perhaps *the*) interest that the power to make binding joint decisions serves is our interest in giving others authority over our actions. When we jointly decide that I will ϕ, I give you the authority to demand that I will ϕ and to rebuke me if I fail to do so. Moreover, I give you the authority to

determine whether I will continue to be bound to φ by deciding whether to agree to retract our joint decision. Thus, coupling the normative powers theory with the deliberative theory, we can hold that promises transfer authority *by* bringing a joint decision into force.

Of course, the normative powers theorist could also deny the deliberative theory. But in doing so, she would deny herself the explanatory resources that the deliberative theory offers. As we have seen, once we grant the bindingness of joint decisions, the deliberative theory gives us promises' other features for free: their uptake requirements, validity conditions, release conditions, and exclusionary effects on deliberation. If she denies the deliberative theory, however, the normative powers theorist must instead write each of these features directly into her description of the normative power of promising. Then she must show why our interests most favor having *this exact kind* of normative power, with precisely the uptake conditions, validity conditions, release conditions, and effects on deliberation that promising has. Perhaps this challenge can be met. (See Owens 2012 for a formidable attempt). But if she instead adopts the deliberative theory, the normative powers theorist's task becomes much easier. All she needs to do is show why it is in our interest to invest the results of our already-existing practice of joint practical deliberation with the force of moral obligation. Once we have that, the rest of promises' features will follow.

5. Conclusion: the redundancy argument

Here's what I've argued so far. If we are able to decide together what to do, then we must have speech acts by means of which we propose joint decisions. One of the speech acts we need is a proposal to make a joint decision about what the speaker will do by means of the propose-and-challenge method: what I have called an 'I-proposal'. Given plausible assumptions about the structure of joint practical deliberation, we can draw a number of conclusions about the properties such proposals will have. Together, these conclusions imply that I-proposals have *exactly the same properties as promises*. When made and accepted in good faith, I-proposals have the same normative effects as promises: they give the speaker a strong *pro tanto* reason to perform the relevant action, typically obligating the speaker *to* the addressee to do so (§2.1) and making it inappropriate for her

31

to seriously consider doing otherwise (§2.2). For I-proposals to bind in this way, however, certain conditions must obtain: the proposal must be recognized and not challenged by the addressee (§2.3); the addressee must not have released the speaker from her obligation (§2.4); and the addressee must not have deceived or coerced the speaker into making the proposal (§2.5). These are the very same conditions that are required for promises to have their normative force. So an I-proposal has the same normative effects, under the same conditions, as a promise. The best (perhaps the only) way to explain these systematic similarities between promises and I-proposals is to conclude that promises *are* I-proposals.

In offering this abductive argument, I have mostly focused on presenting the deliberative theory's attractions rather than objecting to its opponents. However, we are now in a position to give an argument that challenges alternative theories more directly.

Say you accept that people engage in joint practical deliberation, and that one of the speech acts required for this activity has the very same properties that common sense ascribes to promises. If you accept this much, it is hard to go on and deny that promises *are* these proposals in joint practical deliberation. Doing so would commit you to the existence of two distinct speech acts with indistinguishable properties: I-proposals and promises. This view makes it mysterious why we have promises at all. I can obligate myself to you to ϕ by proposing that we decide together that I will ϕ. So why would I need a second, distinct way of undertaking obligations to you with exactly the same normative upshots? Such a speech act would be redundant, unnecessary, even epiphenomenal. This is the status we must assign to promises if we deny the deliberative theory. But promises are *not* redundant in this way. Therefore, the deliberative theory must be true. Call this *the redundancy argument*.

We could, of course, avoid the argument's conclusion by simply accepting that promises are redundant. But this is deeply unattractive. First, it would call the significance of promising into question: if promises are redundant, why should we care about them? More importantly, accepting promises' redundancy would saddle our theory with new and difficult mysteries. Why do promises exist, if their central functions can all be fulfilled by the more general practice of joint practical

deliberation? If I-proposals are an additional speech act, separate from promises, then why don't we have a word for them in ordinary language? How has this entire category of speech act, at least as normatively significant as promising, escaped the notice of both ordinary language and philosophical inquiry? And why do these two separate speech acts, promises and I-proposals, have such eerily similar properties? Is it just a cosmic coincidence? The only way to avoid these awkward questions is to accept the deliberative theory.

Let's briefly consider how the redundancy argument applies to the three theories of promising we considered in section 4.

Begin with classical *conventionalism*, which says that promises are moves in the social practice of making and keeping promises (§4.1). Once we understand the nature of joint practical deliberation, we can see that all of the functions served by this promising-specific practice could be performed equally well by the wider practice of joint practical deliberation. If we already have the practice of joint practical deliberation, why would we need a separate practice of making and keeping promises? If we do have such a practice, it is redundant.

Now consider the *expectation theory*, on which to promise to ϕ is to offer your promisee a special kind of assurance that you will ϕ (§4.2). This special assurance is based in the promisee's knowledge that you are morally obligated, and specifically obligated *to him*, to ϕ. But if we engage in joint deliberation, I can offer you exactly this kind of assurance by proposing that we jointly decide that I will ϕ. Why then would we need a distinct speech act of promising to provide a second way of offering the same assurance? If we do have such a speech act, it is redundant.

Finally, consider the *normative powers theory* (§4.3). On this view, to promise is to exercise a normative power to obligate oneself by declaration, granting one's promisee the authority to demand that one perform the promised action and to release one from this obligation if he wishes. But I can give you exactly these kinds of authority by simply proposing that we decide together that I will ϕ. Why, then, would we need a *sui generis* normative power of promising, if we can shape the normative landscape in exactly the same way by proposing joint decisions? If we do have such a normative power, it is redundant.

In sum, if one accepts that we engage in joint practical deliberation, and thus that we can propose to make joint decisions regarding our own actions by means of the propose-and-challenge method, it is strange to maintain that promises are a further, distinct speech act. Such a position is comparable to that of a nineteenth-century physicist who, acknowledging that mean kinetic energy exists and has all of the same observable features as temperature, still maintains that temperature is a further, distinct property. It is hard to see what could motivate such a view other than theoretical inertia. It is far more elegant and plausible to conclude that there are not two things here, but one: promises and I-proposals are one and the same.

We thus have a general argument against all alternatives to the deliberative theory: these theories render promising redundant. The only way to escape this argument is to deny that there exist proposals in joint practical deliberation that have the same properties as promises. But that is the claim that the whole of this paper has been occupied with defending. To deny this claim, one must contest the theory of joint practical deliberation from which it follows. So, if the deliberative theory is false, that must be because the account of joint practical deliberation on which it rests contains some hidden flaw. For if we accept that joint practical deliberation exists and has roughly the shape I have described, the conclusion that promises are moves within this activity seems almost inevitable.

---

[1] I explain the propose-and-challenge method in §1.3.

[2] The other arguments Gilbert gives in favor of the joint decision account are that it vindicates (i) the observation that promissory obligations are owed *to the promisee* in particular (2011: 101), and (ii) the idea that promises create obligations by an act of will (102). I agree that (i) is true, and there is likely some interpretation on which (ii) is true as well. But I see no reason why other theories of promises cannot also accommodate these claims.

[3] Constraints of space mean that I won't be able to fully defend the theory of joint practical deliberation presented in this section. (For a more detailed defense, see my [Other Work]). Thus this paper might be read more conservatively as defending a conditional thesis: *if* joint practical deliberation takes the form I sketch here, then promises are moves within this activity. I think the antecedent is plausible enough to make the conditional worth considering.

[4] This approach distinguishes my account from prior work on joint practical deliberation (e.g., Westlund 2009; Laden 2012; Bratman 2014: Ch. 7), which has tended to place less weight on this analogy.

[5] The requirement that joint decisions be based on shared reasons strikes some as unrealistically demanding. Since I do not have the space to address this worry here, it's worth pointing out that one could adopt the deliberative theory while rejecting the idea that joint decisions are based on shared reasons. Most of the deliberative theory's explanations appeal to the 'downstream' effects of making a joint decision, rather than the 'upstream' basis on which joint decisions are made. The only exception is the account of why coerced and deceived promises are invalid (§2.5), which appeals to the link between sensitivity to shared reasons and deliberative good faith.

[6] Though note that in the right context, silence could express a challenge (say, combined with a pointed stare).

[7] Two notes about agreements. First, unlike our other terms, 'agreement' is a success term. While our other terms refer to proposed joint decisions, the term 'agreement' refers to *accepted* joint decisions. If the parties didn't agree to it, it wasn't an agreement. Thus agreements are not speech acts, strictly speaking; *proposals* to make agreements are speech acts. Second, we don't have words in English that distinguish between the propose-and-challenge and propose-and-ratify methods of proposing agreements. (This is why the bottom row of our table has one box where the others have two). But the distinction still applies: consider the difference between "Let's go on a walk!" and "Would you like to go on a walk?" The former uses propose-and-challenge; the latter uses propose-and-ratify.

[8] By 'under normal conditions' I mean that (a) the promisee does not reject the promise and (b) no coercion or deception is involved. Please take this qualification as understood unless context indicates otherwise.

[9] The opponent of the deliberative theory might still claim that joint decisions are binding *because* they involve a mutual exchange of promises. Thus, *contra* the deliberative theory, we should explain joint decisions in terms of promises rather than the other way around. Fortunately, I think it can be shown that joint decisions are *not* promise-exchanges: there is no exchange of two distinct promises that can mimic the normative effects of a single joint decision. I refer the reader to Gilbert (1993), who argues this point in compelling detail.

[10] This account does imply that I could *challenge* Dee Dee's attempt to release me from my promise. This might be surprising: we rarely object to being released from our promises. But sometimes we do. Say you promise your proud but frail grandmother that you will help her move into her new home. Being proud, she tries to release you from your

promise: "You don't have to help, I'll be fine." Since she *does* need your help, you reply: "I insist: I gave you my word and I'm sticking to it!" You thereby challenge your grandmother's proposal to release you from your promise.

[11] I assume that invalid promises are still *promises*: a promise was made; it just failed to generate any reasons.

[12] Influentially defended by David Hume (1739/1978: 516-525), conventionalism was popular through the late 20th century, winning the allegiance of John Rawls (1955; 1971), John Searle (1969), and G. E. M. Anscombe (1978), among others.

[13] Some exceptions to this trend are Deigh (2002), Shockley (2007), and Taylor (2013).

[14] Scanlon was not the first to propose the expectation theory (see Árdal 1968, Narveson 1971, MacCormick 1972, and Thomson 1990: Ch. 12). But Scanlon's is widely regarded as the canonical presentation of the view.

References

Anscombe, G. E. M. 1978. Rules, Rights, and Promises. *Midwest Studies in Philosophy* 3:318-323.

Árdal, Páll S. 1968. "And That's a Promise." *Philosophical Quarterly* 18:225-237.

Block, Ned, and Robert Stalnaker. 1999. Conceptual Analysis, Dualism, and the Explanatory Gap. *Philosophical Review* 108:1-46.

Bratman, Michael E. 1987. *Intentions, Plans, and Practical Reason*. Stanford: CSLI Publications.

Bratman, Michael E. 2014. *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press.

Darwall, Stephen. 2011. Demystifying Promises. In *Promises and Agreements*, edited by Hanoch Sheinman, 255-276. Oxford: Oxford University Press.

Deigh, John. 2002. Promises Under Fire. *Ethics* 112:483-506.

Frick, Johann. 2016. What We Owe to Hypocrites: Contractualism and the Speaker-Relativity of Justification. *Philosophy & Public Affairs* 44:223-265.

Gibbard, Allan. 2003. *Thinking How to Live*. Cambridge: Harvard University Press.

Gilbert, Margaret. 1993. Is an Agreement an Exchange of Promises? *Journal of Philosophy* 90:627-649.

Gilbert, Margaret. 2004. Scanlon on Promissory Obligation: The Problem of Promisees' Rights. *Journal of Philosophy* 101:83-109.

Gilbert, Margaret. 2011. Three Dogmas about Promising. In *Promises and Agreements*, edited by Hanoch Sheinman, 83-108. Oxford: Oxford University Press.

Grice, Paul. 1989. Logic and Conversation. In *Studies in the Way of Words*, 22-40. Cambridge: Harvard University Press.

Hart, H. L. A. 1982. Commands and Authoritative Legal Reasons. In *Essays on Bentham: Jurisprudence and Moral Theory*, 243-268. Oxford: Oxford University Press.

Holton, Richard. 2004. Rational Resolve. *Philosophical Review* 113:507-535.

Hume, David. 1739/1978. *A Treatise of Human Nature*, edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Oxford University Press.

Jones, David H. 1966. Making and Keeping Promises. *Ethics* 76:287-296.

Kolodny, Niko, and R. Jay Wallace. 2003. Promises and Practices Revisited. *Philosophy and Public Affairs* 31:119-154.

Korta, Kepa and John Perry. 2015. Pragmatics. In *The Stanford Encyclopedia of Philosophy* (Summer 2015 Edition), edited by Edward N. Zalta. <http://plato.stanford.edu/archives/sum2015/entries/pragmatics/>.

Laden, Anthony Simon. 2012. *Reasoning: A Social Picture*. Oxford: Oxford University Press.

Lewis, David. 1969/2002. *Convention: A Philosophical Study*. Oxford: Blackwell Publishers Ltd.

Lewis, David. 1979. Scorekeeping in a Language Game. *Journal of Philosophical Logic* 8:339-359.

Locke, John. 1689/1980. *Second Treatise of Government*. Edited by C. B. Macpherson. Indianapolis: Hackett Publishing Company, Inc.

MacCormick, Neil. 1972. Voluntary Obligations and Normative Powers – I. *Proceedings of the Aristotelian Society, Supplementary Volumes* 46:59-78.

Narveson, Jan. 1971. Promising, Expecting, and Utility. *Canadian Journal of Philosophy* 1:207-233.

Owens, David. 2006. A Simple Theory of Promising. *Philosophical Review* 115:51-77.

Owens, David. 2012. *Shaping the Normative Landscape*. Oxford: Oxford University Press.

Rawls, John. 1955. Two Concepts of Rules. *Philosophical Review* 64:3-32.

Rawls, John. 1971. *A Theory of Justice*. Cambridge: Belknap Press.

Raz, Joseph. 1972. Voluntary Obligations and Normative Powers – II. *Proceedings of the Aristotelian Society, Supplementary Volumes* 46:79-102.

Raz, Joseph. 1975. *Practical Reason and Norms*. Oxford: Oxford University Press.

Raz, Joseph. 1977. Promises and Obligations. In *Law, Morality, and Society: Essays in Honour of H. L. A. Hart*, edited by P. M. S. Hacker and Joseph Raz, 210-228. Oxford: Clarendon Press.

Ripstein, Arthur. 2009. *Force and Freedom: Kant's Legal and Political Philosophy*. Cambridge: Harvard University Press.

Robins, Michael H. 1984. *Promising, Intending, and Moral Autonomy*. New York: Cambridge University Press.

Roth, Abraham Sesshu. 2016. Intention, Expectation, and Promissory Obligation. *Ethics* 127:88-115.

Scanlon, T. M. 1990. Promises and Practices. *Philosophy and Public Affairs* 19:199-226.

Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge: Harvard University Press.

Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Shiffrin, Seana Valentine. 2008. Promising, Intimate Relationships, and Conventionalism. *Philosophical Review* 117:481-524.

Shiffrin, Seana Valentine. 2011. Immoral, Conflicting, and Redundant Promises. In *Reasons and Recognition: Essays on the Philosophy of T. M. Scanlon*, edited by R. Jay Wallace, Rahul Kumar, and Samuel Freeman, 155-178. Oxford: Oxford University Press.

Shockley, Kenneth. 2007. On that Peculiar Practice of Promising. *Philosophical Studies* 140:385-399.

Southwood, Nicholas, and Daniel Friedrich. 2009. Promises Beyond Assurance. *Philosophical Studies* 144: 261-280.

Stalnaker, Robert C. 1984. *Inquiry*. Cambridge: The MIT Press.

Stalnaker, Robert C. 1999. *Context and Content: Essays on Intentionality in Speech and Thought*. Oxford: Oxford University Press.

Taylor, Erin. 2013. A New Conventionalist Theory of Promising. *Australasian Journal of Philosophy* 91:667-682.

Thomson, Judith Jarvis. 1990. *The Realm of Rights*. Cambridge: Harvard University Press.

Westlund, Andrea C. 2009. Deciding Together. *Philosophers' Imprint* 9:1-17.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.