

Reflection, Disagreement, and Context

Edward S. Hinchman

University of Wisconsin-Milwaukee

[Final **corrected** draft of 1 November 2010]

How far, if at all, do our intrapersonal and our interpersonal epistemic obligations run in parallel? How are we epistemically obligated to weigh diverging opinions – a change of mind – that we expect we'll have in the future? Do those obligations resemble whatever obligations we may have to be responsive to the opinions of peers who disagree with us?

This paper recommends that we treat these questions as addressing the *stability* of doxastic commitment in the two dimensions. If we think of belief as paradigmatically the product of doxastic deliberation, as some philosophers now do,¹ then we can view forming a belief as bringing doxastic deliberation to a proper conclusion, thereby generating a properly stable commitment. And we can make our questions more specific: Does a doxastic stance that fails to do justice to expected future opinion manifest a properly stable orientation as it moves forward into that future? Does a doxastic stance that fails to do justice to interpersonal disagreement manifest a properly stable orientation as it moves outward into the social give-and-take of reasons?² How far, if at all, do these species of doxastic stability run in parallel?

In the background lies an analogy between doxastic and practical commitment. It is through being 'stable,' as we'll use the term, that an intention rationally resists being reconsidered or abandoned – despite foreseeable 'temptations' – in the interval between formation and follow-through.³ While there is no doxastic analogue of following through (that

is, acting) on an intention, there is a doxastic analogue of the way in which intention is constrained by how the agent can project a stable ongoing commitment. The issue posed by stability on both sides of the analogy is not, as we'll see, the need to *keep* your commitment stable but the need to *project* a rationally stable future in order to form the commitment in the first place.

The challenge posed by the need for practical stability is best revealed by cases with the structure of Gregory Kavka's Toxin Puzzle.⁴ In Kavka's case, you will be rewarded for forming an intention to drink a mild toxin but will have the reward in hand well before the time comes to drink. You foresee that you'll rationally abandon the intention in that interval. So it seems you cannot form the intention – simply because you cannot regard that intention as relevantly stable. Though you would gladly drink the toxin to get the reward – if that were the offer – you cannot get the reward for merely forming an intention to drink the toxin, because you can foresee that you will rationally abandon that intention. We'll develop a doxastic analogue of the Toxin Puzzle, where you foresee that you will rationally abandon a commitment by changing your mind in a different respect – by undergoing a shift in the context in which you doxastically deliberate. One aim of the paper is to explain how the shift in context that we'll describe can undermine doxastic stability.

We'll consider how expecting or projecting such a stability-undermining shift in context bears on both the intrapersonal and the interpersonal dimensions of belief formation. We'll see how an appreciation of this species of context-sensitivity undermines some common assumptions about both the diachronic dimension of doxastic commitment and the epistemic significance of disagreement.

Our central case is thus designed to reveal how an aspect of doxastic context can bear on doxastic stability. We'll first consider the case, then the parallel with Kavka's toxin case.

The case unfolds in two stages, which we'll consider separately before putting them together. In stage one, it is 4 pm on a Friday, and you are wondering if your bank is open Saturdays because you need to deposit a check for some not terribly important purpose (say, because you suspect that the person who wrote the check would like the money drawn from his account sooner rather than later).⁵ You consult your memory, find an apparent recollection of having once been in your bank on a Saturday, and on that basis form the belief that your bank is open Saturdays. In light of that belief, you decide you don't need to inconvenience yourself with a trip to the bank today, since you can wait till tomorrow to run the errand. So stage one goes, at least, when we consider it separately from stage two.

Stage two picks up where stage one ends. As you conclude your stage-one deliberation, treating your evidence as sufficient in your low-stakes doxastic context, an eccentric but reliable courier makes you a credible offer. He reliably promises to deposit your check tomorrow himself, but only if you retain or soon regain the deliberative stance that you reached at the end of stage one. You must do so, he stipulates, without any confusion or lack of awareness of what he's asking you to do. In particular, the courier asks you to continue to treat your doxastic context as the sort of low-stakes context in which you formed your belief in stage one.

An odd offer, but why not simply persist in your belief and take advantage of it? The trick is that this courier is not only eccentric but notoriously finicky (in a way we'll consider

presently). Anticipating his finicky reaction complicates your situation in a way that ensures that your evidence no longer suffices to close your deliberation. Your new predicament defines a higher-stakes doxastic context.

Anticipating this shift in context complicates your situation all the way back to stage one. Consider how your predicament unfolds in the two-stage case. If you do not conclude, in stage one, that your bank is open Saturdays, the courier will (you believe) do nothing. But if you draw that conclusion, the courier will (you believe) press his offer, and anticipating that offer – even now, before you’ve drawn any conclusion – puts you in a delicate position. Refusing the offer would (you believe) insult him, generating a nasty confrontation that you’d very much prefer to avoid. But accepting the offer would (you believe) generate an equally nasty result if your belief that your bank is open Saturdays should prove false. If you accept, you thus expect that it will become a *very* important question to you whether your bank is open Saturdays. At the very least, you see, you’ll need to call the bank to check their hours. This disposition to inquire shows that you’ll no longer count as believing that the bank is open Saturdays.⁶ Remember: we’re treating deliberative closure as the key to belief formation, not dispositions to bet or more broadly behave. Though your betting dispositions may remain unchanged in the transition from stage one to stage two, your disposition to reopen deliberation marks a key doxastic difference. (We’ll examine this aspect of case in section II.)

What’s interesting about the case, for our purposes, is that merely expecting the courier to make his offer, should you conclude that your bank is open Saturdays in stage one, has the force of thrusting you into that stage-two deliberative context. Merely expecting the doxastic standard to rise, should you draw that conclusion, has the force of raising the standard on its own.⁷ This is not a case in which you expect to undergo a change of mind in the future. We’ll get to those

cases, but that diachronic element is not yet in play. In this basic version of the courier case, you expect merely that your doxastic context will change *if* you draw a certain conclusion. This is not a change in what you believe, since you do not yet have a belief on the matter. The point is that the mere expectation of this change – a change not in belief but in doxastic context – has the force of changing your doxastic context on its own. Our question will be how this feature of the case bears on cases with the diachronic and interpersonal elements on which we'll focus. After considering some objections to the basic case in section II, and some broader features of it in subsequent sections, we'll coin a diachronic courier case in section VI and an interpersonal courier case in section VIII.

II

Why not leave the case described in terms of your *degree* of belief? As we noted, your betting dispositions may remain unchanged across the stages. If we conceive of your doxastic state as exhausted by such dispositions, then we cannot coin a doxastic analogue of Kavka's toxin case, since that case depends on the fact that the formation of an intention to φ at t is not merely the onset of a disposition to φ at t . If it were, then your expectation that you'll be irrational if you don't change your mind before t would have no bearing on whether you can now form the intention to φ at t – since this 'change of mind' would merely amount to the onset of a different disposition, and it is not a condition on the possibility of undergoing any disposition, just as such, that you not later undergo a different disposition. To get a toxin case – that is, a case that would plausibly manifest a constraint on forming an intention – we must assume that the target intention to φ at t is not merely a disposition to φ at t . (This would not rule out reducing it to a

more complex dispositional state.) Since our aim is to pursue a parallel between intrapersonal practical and doxastic commitment, we must conceptualize belief as possessing a comparably commissive dimension. (This too would not rule out a more complex dispositional reduction.)

So ask instead: how might an appeal to degrees of belief give our courier something to work with? If we rest with saying that your degree of belief remains unchanged as the doxastic standard goes up, then we give him no doxastic difference to exploit. But we can coin a notion of *degreed doxastic commitment* that the courier could exploit. A degree of doxastic commitment would correspond to a degree of doxastic-deliberative closure – which on the deliberative model amounts to a degree of belief.⁸ We can then describe the courier's intervention as follows. In stage one, you believe that your bank is open Saturdays to degree n . In stage two, once the courier makes his offer, you hold that belief to a degree less than n – because you see that more is now at stake. To say that you believe only to degree m , short of n , is to say that *to that extent* – n minus m – you have reopened the matter and are now redeliberating. (We could elaborate this in several ways.⁹) The courier will deposit your check if but only if you keep your mind closed to at least degree n – that is, to the degree that defined stage one. We thus get the same problem as before: since (as you believe) his offer will raise the doxastic standard, if your degree of belief remains n , then (as you believe) your degree of belief should be lower than n .

We can use this degreed notion of commitment to address some worries that may arise about the case as formulated in section I. The case uses a notion of context originally coined to codify intuitions about knowledge attributions: you're more willing to say that S knows that p when there's less at stake (in competing versions, for S, for you, or for one assessing the attribution¹⁰). As described, the case assumes that an analogous consideration of context plays a

role in closing deliberation in the way of forming a belief, as if forming a belief involves a self-attribution of knowledge. But can't you form a belief that *p* without thinking that you know that *p*? And can't you believe that *p* while still deliberating whether *p*? In each instance, the degreed notion of commitment shows us how. Though we began by assuming that one achieves doxastic-deliberative closure whether *p* if and only if one believes that *p* (assuming a positive resolution), we can suspend that assumption and describe the courier's intervention entirely in terms of degrees of deliberative closure. We thus leave open the possibility that you believe that *p* while still deliberating whether *p*. Since insofar as you are deliberating whether *p*, you do not think you know that *p*, we also leave open the possibility that you believe without thinking you know.

Degree of doxastic-deliberative closure is not, of course, what philosophers typically mean by 'degree of belief.' What's important is merely that the courier's intervention trades on *some shift* from stage one to stage two in whether or how strongly you hold the belief. Having acknowledged the possibility of describing the shift as one of degree, let's for simplicity revert to an undegreed concept of belief – till we reach the subject of disagreement, when we'll need degrees of belief to make sense of the issue.

III

The apparent moral of the courier case is that you cannot accept the courier's offer because you cannot reasonably expect that your doxastic commitment will be stable. To get a better sense of how the case bears on the stability of your doxastic commitment, let's now see how it parallels Kavka's toxin case. We'll consider the stability of commitment directly in section IV.

It does not change the force of Kavka's original case to elaborate it in two stages as follows. In the first stage, imagine that you form an intention to drink a mild toxin tomorrow as a stunt for a local radio station, which will pay you a hundred dollars for thus making yourself mildly ill. Forming such an intention is perfectly intelligible: since you will get the payoff only if you drink, and you prefer payoff-plus-bellyache to all other available outcomes, you can reasonably expect that your intention will remain rationally stable till you follow through on it. To say that you regard the intention as rationally stable is not to say that you believe that you *will* follow through on it but that you believe that it would be *irrational* for you not to follow through on it.¹¹ Imagine that it's somehow part of the stunt that you have to form the intention to drink the previous day and persist in it to follow-through. Now imagine, in the second stage, that Kavka's eccentric billionaire proposes to reward you with a million dollars merely for forming the intention. And imagine that with this million dollars in your account you would prefer not to give yourself a bellyache for a mere additional hundred.

Let's assume that Kavka's treatment of his original case is correct: you cannot get the million dollars merely for forming the intention because you foresee that you it would be irrational to follow through on that intention.¹² It follows that you cannot even get the hundred-dollar payoff. For consider how the full two-stage case will look to you as you enter stage one. You would like to get the hundred dollars that the radio station is offering for drinking the toxin, but you can see that as soon as you form the intention to drink you'll enter stage two: the billionaire now proposes to reward you merely for persisting in that intention, and you'll view that additional hundred as not worth the trouble of giving yourself a bellyache. Foreseeing this, you cannot form the intention to drink the toxin even in the way that is perfectly intelligible when we consider only stage one. Foreseeing the billionaire's proposal prevents you from

obtaining even the smaller payoff. (You might think you could simply refuse that proposal, but remember: he proposes to reward you for an intention that you already have. How could it be rational to refuse such an offer?)

We now have a toxin case that parallels our new courier case. If we think of the intention in the former as a practical commitment and the belief in the latter as a doxastic commitment, we can describe them in parallel terms as follows. In each case, you cannot form the commitment in stage one while looking ahead as you normally would to stage two. The reason you cannot do so, as we might say, is that when you ‘project’ yourself ahead from stage one to stage two – that is, in the normal way projectively imagine yourself having retained or reformed the commitment – you can see that you’ll have made a mistake. In the toxin case, you’ll have an intention that you can see you should abandon. In the courier case, you’ll have a belief for which you can see that your evidence does not in context suffice. These are different mistakes, of course – a difference that derives from the differences between intention and belief, and more broadly between practical and theoretical reason. But we can understand each mistake in terms of your anticipating an instability in the commitment.

IV

What then is the stability of commitment in this sense? We’ve been using the term to mark whatever it is that goes missing through the bizarre intervention in each of these cases, but what exactly is this? We’ve described what goes missing as your ability to maintain the commitment as you ‘projectively’ imagine your future – that is, as you imagine the future self of yours that will (in the toxin case) act on the intention or (in the courier case) maintain the belief in an

evidentially more demanding context. To understand what makes a commitment stable or unstable in these forward-looking dimensions, we must understand the nature of this projection.

We can see what's distinctive of the projection in a courier case by considering how you might fail to project. Imagine, altering the case in a crucial respect, that you expect to be unnerved by the courier's offer in a way that you now regard as neurotic or perverse. You expect that when the courier makes his offer – assuming you form the belief in stage one – you'll come to be under an impression that your doxastic context has changed, but you expect that this impression will embody a misinterpretation of your doxastic context. You expect that your doxastic context will not in fact have changed, so you expect that you'll have no basis for reopening deliberation. Here you can form the belief in stage one even as you look ahead to stage two. We can explain why by noting that, though you expect that your impression of your doxastic context will change in stage two, you do not project that future from within your stage-one deliberation.

Can we construct an altered toxin case along these lines? We can, but then we imagine you violating the terms of the billionaire's offer, since we imagine you confused about how your options will have changed. For simplicity, let's work from Kavka's original one-stage case. Assuming the billionaire relevantly reliable, your options will have changed before the time at which you're to drink the toxin: once the money hits your bank account, you can have a million dollars without the bellyache. Of course, if you expect you'll not recognize that this option has become available, then you can form the intention to drink – since you prefer money plus bellyache to the other options available to you as you form the intention. Kavka's billionaire explicitly disallows relying on your own confusion in that way. Note that the courier could not stipulate a parallel condition on his offer in our new case. There is nothing in the *content* of the

courier's offer that ensures that receiving it will have the effect in the second stage of the case as originally described. What changes your doxastic context in a courier case is not the content of the offer but the courier's act of making it – against the background assumption that he's finicky, etc.

This difference reveals something important about the nature of deliberative projection. When you consider whether to φ at t , you must deliberate – at least, in part – from the practical perspective you expect you'll occupy at t . Benefits and costs of forming the intention now – before t – can weigh with you, but you're principally deciding what to do at t , not what to do now. Toxin cases make this vivid, but the point applies to mundane cases as well. The benefits and costs that matter most are those that would accrue by virtue of your φ ing at t , not merely by virtue of your deciding or intending to φ at t . In the doxastic case, by contrast, the requirement that you attend to any particular future context derives from an expectation that your doxastic perspective will be different then in ways that ought to matter to you now. 'Projecting' into a future context, in our sense, just is treating the perspective you expect you'll occupy in that context as imposing obligations on your present deliberation. So what matters in a doxastic case is whether you actually project. In a practical case, it may be that you ought to project, even if you don't, and it is this obligation that reveals the bearing of your future perspective on your present deliberation. But in a doxastic case, the bearing of any future perspective you expect you'll occupy depends on whether you actually project into that perspective. We'll revisit this point in sections VI and VII.

Let's now see how courier cases differ from Ulysses cases, which are the cases that epistemologists have tended to cite both in exploring the analogy between intention and belief and in arguing that stability in belief formation possesses a diachronic dimension.¹³ This will help clarify the species of stability at issue.

The case of Ulysses and the Sirens illustrates vividly the diachronic element in intention formation. Since we lack space to pursue the question, let's concede up front that you cannot form an intention to ϕ at some future time t if you expect that you'll give in to a temptation to abandon that intention before t .¹⁴ Ulysses' determination to sail past the Sirens cannot, therefore, add up to his simply forming an intention to do so. Even a neurotechnologically sophisticated Ulysses who'd (as it were) had his *brain states* 'tied to the mast' to ensure compliance would not thereby count as having formed an intention to sail past the sirens. It seems plausible enough that such 'precommitment' would not count as an intention.¹⁵ But it is much harder to see why you must give a similar standing to diachronic stability in belief formation.

Bas van Fraassen has argued that diachronic stability does play such a role in belief formation. He argues for what he calls the Reflection Principle, which says that your current degree of belief in a proposition p should fall within the range of the degrees of belief that you expect you'll have in p in the future.¹⁶ Since we've resolved for now to work with the notion of belief *simpliciter* – without degrees – we can say that Reflection requires that you believe what you believe your future self will believe. A number of critics have pressed counterexamples to Reflection.¹⁷ Our present aim is not to adjudicate this debate but to see why the Ulysses cases used as a frame by both van Fraassen and his critics are misleading.¹⁸

Note the differences between Ulysses' predicament and your predicament in the toxin case. Ulysses cannot form an intention to sail past the sirens because he expects that he will not thereby succeed in guiding his conduct. You cannot form or retain an intention to drink the toxin, by contrast, not simply because you do not expect that you'll thereby succeed in guiding your conduct but because you cannot rationally draw the deliberative conclusion that you ought to drink the toxin. It is crucial to his case that Ulysses has no such problem. He gets himself tied to that mast because he has drawn the deliberative conclusion that he ought to sail past the sirens but does not expect that he'll be able to implement his judgment in the normal intention-mediated way. This expectation prevents Ulysses from forming the intention – but not, as in a toxin case, because he cannot draw the deliberative conclusion that would inform it.

There are thus two dimensions of stability in intention formation. Toxin cases directly violate a stability condition on drawing the deliberative conclusion that you ought to perform the action and thereby – indirectly – prevent you from forming the intention to perform it. You cannot form the intention in a toxin case only because you cannot rationally form an intention to perform an action that you do not judge you ought to perform. Ulysses cases, by contrast, do not violate that stability condition on drawing a deliberative conclusion. Rather, they directly violate a stability condition on forming the intention to perform an action that you *do* judge you ought to perform. This second stability condition is robustly diachronic: you cannot form the intention to ϕ at some future time if you expect that you will not at that time follow through on the intention. But the first stability condition is not diachronic in this way, since it makes no reference to your *expectations* about what you *will* do.

In a doxastic case we have only the first sort of stability condition, the stability condition on drawing the deliberative conclusion. There is no robustly diachronic condition, for the simple

reason that doxastic deliberation is focused on the present context, not on the possibly future context of action. In the courier case, your doxastic context crucially includes anticipated follow-through on your doxastic verdict in practical deliberation about when to deposit the check. But you are not constrained by expectations about how you will go on to reason from this verdict. Because doxastic deliberation is necessarily present-directed (in this broad way), belief does not perform the specifically diachronic function that characterizes intention. Because it does not play that role, belief is not constrained by the diachronic element in Ulysses cases. At least, we cannot derive such a constraint from reflection on an analogy with such cases.

VI

Epistemologists therefore go wrong in focusing on Ulysses cases. We've begun to see how they go wrong, but we can refine the diagnosis. Though the stability condition revealed by the original courier case is not diachronic, we can construct diachronic courier cases. Since some aspects of diachronic courier cases resemble Ulysses cases, we can thereby provide an alternative treatment of the intuitions that attract epistemologists to Ulysses cases. We can also vindicate a restricted – but only that restricted – application of the Reflection Principle.

To see how a diachronic courier case works, we need to distinguish it from a synchronic case that it closely resembles. In this other case, a synchronic case with future-tensed content, you are deliberating whether p , for some future-tensed p , but in a way that amounts to wondering whether p *for present purposes*. Say you are wondering whether your bank will be open on a specific Saturday in the future. You first conclude that it will be open, on a basis similar to stage one of the original courier case. Then, in stage two, the courier intervenes, reliably promising to

deposit the check the preceding Friday, but only if you *now* keep your deliberation closed. On this understanding of his intervention, the courier trades on facts about your present deliberative context. Because in stage two you have active concern for your future predicament – say, because you’re beginning to plan for it – it seems you must reopen deliberation. But if you didn’t care about your future predicament – in which you’ll have to cope with the finicky courier’s reaction should your belief about the bank’s hours prove false – it is not clear why you should reopen deliberation.¹⁹ In this sort of case, everything depends on whether you feel that concern, and rationality does not (without further assumptions) require that you feel it.

We’re looking for a future-tensed case which, like the original case, doesn’t depend on a rationally optional attitude of concern. To get such a case – a truly diachronic courier case – we need to imagine you not merely wondering about your bank’s hours on that future Saturday but projecting yourself into the doxastic deliberation that you expect your future Friday self to conduct about those hours. We can see that there is an important difference between the synchronic and the diachronic cases by considering how the courier intervenes in the diachronic case. This future-oriented courier does not care about your present deliberations; he cares only about your deliberations on the Friday in question. This courier promises to deposit your check provided you keep deliberation closed *on that future Friday*. So the case generates instability not directly in your present deliberative predicament but in a predicament you expect you’ll encounter in the future.

Unlike the synchronic case with future-tensed content, this diachronic courier case does reveal something important about stability in belief formation. When you project an upward shift in your doxastic standard, looking ahead to a future deliberation whether *p* that you expect to undertake, your projection forces open your current deliberation whether *p*. This observation

accords with van Fraassen's Reflection Principle: if *in this projective way* you expect that you will not believe that *p* in the future, you should not now believe that *p*. But (as we saw in section V) the way you hold this expectation is not like the way you hold the parallel expectation in a doxastic Ulysses case. In stage two of a diachronic courier case, you do not merely expect that you will refrain from concluding that *p* in the future. You deliberately refrain from drawing that conclusion now through projection into your future predicament.

VII

To see better how projection can make this difference, imagine a third case alongside the synchronic and the diachronic courier cases, a kind of Ulysses case adapted to our emphasis on the role of doxastic context. In this third case, the issue for you at stage two is that you expect you'll hold yourself to a different doxastic standard on that future Friday – but without projecting yourself into that doxastic context. Setting aside courier cases, it is easy to imagine how you might expect that you'll undergo such a shift in your conception of how to weigh available evidence. You expect you'll grow conservative as you get older. Or you expect you'll have a very different take on your life post-parenthood, no longer believing, as you now do, that the suburbs are for sell-outs. You do not expect that this change of mind will reflect your acquisition of better evidence. In the parenthood case, you do not think you'll get better evidence about suburbia. You merely think you'll weigh the evidence you now have differently.²⁰

We do not tend to think that your expectation in this third sort of case – your expectation of a shift in how you'll weigh the evidence – has any bearing on how you should weigh that evidence now. And the reason is that you are not (as we're imagining these cases) deliberately

projecting yourself into the predicament of your future self. If, by contrast, you do thus project yourself, then your expectation bears on your current deliberation in the way described by the Reflection Principle. This vindicates a restricted application of the Reflection Principle, but also diagnoses why the principle is not true more generally. As the third sort of case makes clear, projecting yourself into the predicament of a future self about whose beliefs you have expectations is by no means mandatory.

The contrast between the third sort of case and those discussed in the previous section reveals how intrapersonal projection amounts to taking responsibility for an aspect of your future. Projecting yourself into a future doxastic context, in the respect we're discussing, is not a matter of deliberating 'for' that future self – as if you expect that you will need to rely on your present deliberation because you won't then have time to think matters through. To say that you deliberate 'for' another is to say that you do not take full responsibility for the deliberation, insofar as you do not regard it as speaking for your own point of view. When you project yourself in the way we're investigating, you do not cease deliberating 'for' yourself: you do not cease regarding the question before you as what to believe in your present context. When you project yourself you take deliberative responsibility for this future perspective, not as an alternative to but as a continuation of the perspective from which you are now deliberating.

As we've just seen, you need not take this attitude toward all future perspectives that you expect you'll occupy. As a younger person, for example, you need not take *deliberative* responsibility for the conservatism that you expect will strike at middle age. On the other hand, we might conjecture that you must take this attitude toward at least some of these future perspectives. (The attitude in question – deliberative projection, in our sense – seems fundamental to thinking of yourself as a temporally extended or articulated being. We cannot, of

course, pursue that question here.)

VIII

This application of the Reflection Principle – restricted to cases of intrapersonal projection – codifies an intrapersonal epistemic norm. Does the norm have an interpersonal analogue?

Some philosophers have argued that when an epistemic peer disagrees with you, you should give the peer’s judgment an authority over your beliefs more or less equal to the authority of your own judgment.²¹ By an ‘epistemic peer’ we mean someone whom you believe to be as good a judge as you are on the subject matter of your disagreement and who does not relevantly differ from you in the evidence she possesses.²² These philosophers argue that your attitude toward your peer should be like your attitude toward your future self when you expect that you will non-pathologically change your mind – but without the presumption, created by the temporal asymmetry, that inquiry is likely to have yielded better evidence. You should adjust your degree of belief accordingly, perhaps ‘splitting the difference’ with your peer.²³

Let’s reconstruct this issue within our framework by developing an interpersonal courier case. Such a case would mimic the projective structure of the diachronic courier case. If you are not projecting yourself into your peer’s doxastic predicament, then the case involves that other person in the way that a merely synchronic case might involve the future: the interpersonal element figures only in the content of your solitary deliberation rather than informing the deliberation itself. We must therefore stipulate that you are projecting yourself into your peer’s doxastic predicament. As in the intrapersonal case, this is a matter of deliberating not ‘for’ the

peer but in a way that is sensitive to and takes appropriate responsibility for differences between your peer's doxastic context and your own, narrowly defined. To project yourself into your peer's doxastic predicament is to deliberate in a way that is sensitive to what you expect would matter *to her*, if she were rational and apprised of all *your* evidence. (We'll elaborate this conception of interpersonal projection in section XI.)

In stage one of this courier case, you deliberate whether your peer's bank is open Saturdays as you imagine she would in light of her non-pressing need to deposit a check, reaching an affirmative conclusion. In stage two the courier intervenes, offering to deposit the check your peer needs deposited if but only if you keep this deliberation closed – as you continue to project yourself into her predicament. Because you can see that the intervention raises the stakes *for her*, you cannot keep deliberation closed. This contrasts with another case that we might imagine, a non-interpersonal courier case with interpersonal content, in which you deliberate without interpersonal projection – in light of what's at stake *for you* – but where the deliberation reflects your recognition of what your peer needs in the two stages. How you deliberate that case depends on how much you care about her needs. As in the synchronic case with future-tensed content that we considered in section VI, that case would not directly engage our issue.

There are two points at which this parallel between intrapersonal and interpersonal breaks down. First, there is a key respect in which the projections must differ. In the diachronic case, the courier's intervention targets your future deliberation and therefore bears on your current deliberation only insofar as you expect to undergo it. A precise parallel between the diachronic and the interpersonal would therefore require that the interpersonal case locate the courier's intervention entirely in your peer's deliberation. But that would generate an uninteresting case in

which you merely believed that *someone else* would undergo a courier case. Your projection in the diachronic case targets your own future deliberation, but your projection in the interpersonal case does not in the same way target your peer's deliberation. It rather serves as a stand-in for her deliberation. It is this stand-in deliberation that the courier engages in the interpersonal case.

That difference entails a second difference. In both cases, when you draw the conclusion that p , where p is the proposition in question, you aren't thinking that that is what the other party – whether future self or interlocutor – *will* conclude but that that is what the other party rationally *ought* to conclude. In the diachronic case this can count as a deliberation at once whether p and whether to believe that p because the one who would believe that p on this basis – your own future self – can simply inherit the conclusion and therefore the commitment in the medium of memory. Not so, obviously, your interlocutor. Whereas deliberative projection in the diachronic case is a way of making up your mind, deliberative projection in the interpersonal case does not yet involve a commitment to the conclusion in question and so is not a way of simply 'making up' your – or anyone's – mind.

IX

These differences between the intrapersonal and the interpersonal show that we have not yet captured the phenomenon of disagreement. Projecting yourself into an expected future deliberation whether p can influence your current deliberation whether p because that expected deliberation is *yours* – that is, one for which you prospectively take responsibility. If projecting yourself into your peer's deliberation whether p could influence your own deliberation whether p – and we have so far seen no reason why it should – that would not be because this peer actually

disagrees with you, since the peer's actual as opposed to projected deliberation and judgment are not yet in view. To make room for them we must move beyond your projection into her deliberative space and ask how she might look back at you from there. But to consider her perspective on you when she disagrees with you, we cannot simply ask how she might project herself into your deliberative space. We are not imagining her deliberating on your behalf – as if you could simply follow through on her conclusion directly. We are imagining her in position to influence you through interlocution. To gauge the nature of this influence, we must ask how she might address you. What is the form of address distinctive of disagreement?

This is a question about the nature of informative assertion. Is an informative assertion one that expresses the speaker's beliefs in *her* context? There is a good reason to think not. Exploiting the asymmetries of doxastic context under discussion, we can easily imagine how a speaker might inform someone that *p* without having concluded that *p* in her own context – that is, without having formed the belief that *p*. When a speaker provides her interlocutor with testimony, perhaps disagreeing with him, she is typically showing him what conclusion she would draw in *his* doxastic context, a context that she may regard as different from her own.

Imagining yourself back in the speaker's role, consider some homely examples in which such a difference in context makes a key difference to your testimony. Imagine you're literally allergic to nuts: ingesting anything with even a hint of nut sends you straight to the ER. Now someone who you know is not even metaphorically allergic to nuts inquires of an item that you're confident doesn't have nuts as a main ingredient, 'Is this nut-free?' Well, *you* wouldn't eat it, because you wouldn't conclude that it is nut-free, given your allergy. But say you know that your interlocutor's practical exigencies fall far short of imposing the doxastic standard that your life-or-death exigencies impose: he merely dislikes the taste of nuts, and not even all that

strongly. Must informative testimony register the doubts that prevent you from drawing this conclusion in *your* context? If you can explain all this complexity to him, that would yield a solution. But imagine you cannot. You ought then, it seems, to tell him what you would judge in his context – that is, in your projection of that context – and not in your own. We can easily imagine a case in which judging in his context – judging projectively – is the only way to be helpfully informative.

If that case seems unlikely to generalize, consider what would follow from, say, your metaphorical ‘allergy’ to the slightest hint of clutter. Do you testimonially stigmatize as ‘messy’ or ‘chaotic’ arrangements that you know most people would find perfectly neat? If, trying not to be a testimonial crank, you mostly reserve such assessments for your private conclusions, does that make you a less than informative testifier?

The virtue of testimonial informativeness does not require such revelations. The norms informing our testimonial practices derive from our epistemic interest in helping each other believe what there is reason to believe, not from our merely psychological interest in prying into each other’s mental states. Testimonial norms are rooted in the public forum, not in the confessional. Informativeness accordingly does not require that you reveal your own beliefs; it can in a given case suffice for informativeness that you are putting your deliberative powers, including your base of evidence, into the service of your interlocutors’ doxastic-deliberative needs.²⁴ Though other norms and virtues inform our practices of assertion, it can suffice for full informativeness that you testify to the conclusion you draw as you project yourself into your interlocutor’s context – given, of course, your actual evidence.

This testimonial norm has direct application to the epistemology of disagreement. The interpersonal analogue of van Fraassen's Reflection Principle claims that you should – more or less (again, we're not addressing this issue) – split the doxastic difference with an epistemic peer who disagrees with you. But what if you and your peer are judging in different doxastic contexts?

We could simply stipulate that by 'X's epistemic peer' we mean, *inter alia*, someone judging in X's doxastic context. But such a restriction could not be motivated by the analogy between the interpersonal and the intrapersonal that we're pursuing. We noted in section VI that a norm of diachronic stability in belief formation requires that you respond appropriately to *projected* shifts in your doxastic context. As we saw, projecting an upward shift in your doxastic standard when you look ahead to your expected deliberation whether p forces open your current deliberation whether p. This shows that the Reflection Principle has at least one important application even if it does not apply when the subject *merely* expects to change his mind, without deliberative projection. So an analogy between intrapersonal and interpersonal applications of the principle could not motivate a restriction to shared doxastic contexts.

Such a restriction would also set aside some of the most interesting cases of disagreement. It is tricky to say generally and precisely just when doxastic contexts differ. But it is clear enough that many disagreements emerge in part from differences in how interlocutors interpret their doxastic contexts – and that some disagreements are even sustained by such differences. Consider the debates over global warming wherein people reach different conclusions because they have different conceptions of what is practically at stake.²⁵ Perhaps

both sides could agree that if there were comparatively little at stake, or lots of time before warming could have any effect, then one should be skeptical about leaping to a conclusion before ‘all the evidence’ is in. But both sides do not accept those practical assessments, and part of the disagreement is one side’s impatience with the other’s refusal to draw a conclusion where – in a context of dire exigencies – less than ‘all’ the evidence nonetheless counts as deliberately sufficient. Many non-scientific cases are like that as well. You and your interlocutor seem to overhear a remark that counts as offensive for only one of you, and the question is: Did she really say *that*? You think yes, not deeming it offensive, but your interlocutor thinks no – simply because he would need more evidence, ruling out possible mishearings, before concluding that someone had said such an offensive thing. We can imagine any such case unfolding between epistemic peers. There is no general reason why such a disagreement in or about context must undermine epistemic peerhood.

On the standard approach to the epistemology of disagreement, the question is how your belief that an epistemic peer disagrees with you whether p should bear on your deliberation whether p . If this peer is judging in a different doxastic context from yours, the natural answer is *not at all* – at least, not directly. That answer is too quick, however: you should not simply dismiss the peer. You need to find a way to let the fact of disagreement engage your deliberative predicament. Even if no norm of informative assertion positively requires that your interlocutor projectively judge in your doxastic context – a stronger thesis than any in play here – the fact that such projection can in a given case suffice for full informativeness provides an incentive for the peer to engage you.

In this way, an account of the epistemology of disagreement should build on an account of testimonial informativeness that acknowledges the point emphasized in section IX. Since

disagreement may depend on a difference in doxastic context, we cannot say how your belief that someone disagrees with you whether p should bear on your own deliberation whether p without understanding how the disagreement would inform testimonial engagement – that is, actual argument – with this interlocutor. How would your peer project a doxastic context on your behalf? How would you reciprocate? Which model would the exchange resemble more closely: our nut-allergy case or our global-warming case? If the former, then disagreement dissipates once the difference in doxastic context is spelled out explicitly. If the latter, then disagreement could perhaps survive such spelling out – the disagreement now registering different interpretations of what a shared doxastic context requires of you and your interlocutor. The epistemic significance of disagreement in such a case turns on whether you and your interlocutor can resolve this more fundamental interpretive question.

XI

The parallel we've uncovered between intrapersonal and interpersonal epistemic obligation rests on a parallel between intrapersonal and interpersonal deliberative projection. We considered intrapersonal projection in section IV. How might we explain projection in the interpersonal case?

Interpersonal deliberative projection plays a key role in transforming a mere disagreement into a possible dispute. As we've seen, interlocutors may disagree on the truth value of a proposition merely because they are judging in different doxastic contexts. Such a disagreement is not yet a dispute. Your prudish interlocutor turns to you in shock: 'Surely she didn't say what she seemed to say?' You begin to reply, 'Yes, she did,' but when it becomes

clear to you how offensive your interlocutor would find that assertion, you back off: ‘Well, maybe we misheard.’ Or it becomes clear that your disagreement with someone over the reality of global warming derives from their assumption that, even if warming is real, God would never let us feel its consequences. You don’t change your position, but you do recognize that you don’t really have a dispute with this person – at least, about *that* question.²⁶ In the first stage of each case there is disagreement, but it is not disagreement that can generate a genuine dispute. A genuine dispute presupposes that parties to it are deliberating and judging in a shared doxastic context – that is, against a broadly shared sense of what’s at stake in settling the question and thus of how much evidence would be required to settle it.

The familiar conception of judgments as ‘moves in the space of reasons’ – which we lack space to examine more fully – must therefore reflect the possibility of dispute, not merely of disagreement. On this conception, when you settle a doxastic question for yourself, you do so against an implicit conception of how someone might not merely disagree with you but dispute you on the question. Even when you remark to yourself that ‘there is nothing else to think’ on the matter, you don’t mean that you cannot imagine someone disputing you. You mean merely that such a person would be obviously irrational. It’s a point about your ability to imagine interlocutors who disagree with you in a shared doxastic context. When you doxastically deliberate, on this conception, you deliberately project yourself into these imagined contexts. Of course, the contexts will often be real, since you’ll often know of people who would dispute you on the question. The point is that thinking of these disagreements as disputes requires thinking of the disputants as sharing – or as close enough to sharing – your doxastic context. But that is just what we mean by interpersonal projection. On this conception of doxastic judgment, interpersonal deliberative projection is fundamental to our capacity to judge.

Again, we lack space to consider further that conception of judgment. But given its familiarity we should not find it surprising that deliberative projection proves fundamental to the epistemology of disagreement. One might put the point by noting that the epistemology of disagreement should strictly be relabeled the epistemology of disputation. Interlocutors who disagree only because they do not share a doxastic context are – epistemologically, at least – failing to engage.

XII

We may conclude that epistemologists who pursue an analogy between intrapersonal and interpersonal versions of the Reflection Principle are indeed onto something, but the insight does not lie where they think they've found it. Whether or how that principle applies depends on a more fundamental question about the roles of projection and context in belief formation.

As we've seen, the real parallel between reflection and disagreement lies in how you take responsibility for your beliefs in these two dimensions. Intrapersonally, you take responsibility for your status as a temporally extended epistemic agent – when you do: again, you needn't always – by deliberating in a way that is informed by any expectation you may have that you'll in future confront a stricter doxastic standard. Interpersonally, you take responsibility for your status as exercising your agency in collaboration with epistemic peers by deliberating in a way that acknowledges the possibility, not of mere disagreement, but of genuine dispute against a shared doxastic standard. We have not explored exactly how the latter acknowledgement should register in your deliberations – whether in terms of 'equal weight' or more complexly.

As toxin cases help us understand the nature of intention, so courier cases help us understand the nature of doxastic commitment. The original courier case helps us grasp the distinction between the context of your deliberative projection and the context – that of your future self – that your projection attempts to engage.²⁷ Diachronic and interpersonal courier cases help us grasp another important distinction: between the context of your projection and the context informing the belief you would form without projection. These elaborations develop the key implication of our courier’s intervention, which is that one can draw a conclusion about what to believe – whether projectively or not – only against an understanding of how that conclusion would bear on relevant practice. The original courier case exploits this sensitivity to context with the assumption that you will not fail to project. We thus assume that the courier’s original intervention immediately forces open your deliberation.

When expecting the shift in context does not force open your deliberation, you may nonetheless confront the question whether to project. Is that future self one for whose doxastic context you’re willing to take responsibility? Is that fellow believer one whom you’re willing to treat as a discussion partner and potential disputant, thereby taking responsibility for your epistemic stake in ensuing disagreements? On your answers to these questions turn key dimensions of interpersonal and interpersonal obligation. The parallel between the dimensions that we’ve charted runs directly through your willingness to take responsibility for an aspect of your personal relations, whether with your own future selves or with potential interlocutors.

¹ See, for example, David Owens, *Reason without Freedom* (London: Routledge, 2000); Richard Moran, *Authority and Estrangement* (Princeton: Princeton University Press, 2001); Nishi Shah and David Velleman, “Doxastic Deliberation,” *Philosophical Review* 114

(October 2005), 497-534. To say that belief is paradigmatically the product of deliberation is not, of course, to say that it necessarily is. There isn't space in this paper to defend the deliberative approach to belief on general grounds. Our aim merely to discover one place it leads.

² The issue about disagreement that we'll pursue is well illustrated by the dialectic among these papers: Thomas Kelly, "The Epistemic Significance of Disagreement," T. S. Gendler and J. Hawthorne (eds), *Oxford Studies in Epistemology*, Volume 1 (Oxford: Oxford University Press, 2005), 167-96; Adam Elga, "Reflection and Disagreement," *Noûs* 41 (2007), 478-502; and Thomas Kelly, "Peer Disagreement and Higher Order Evidence," in R. Feldman and T. Warfield (eds), *Disagreement* (Oxford: Oxford University Press, 2010).

³ For competing views on the 'stability of intention,' see Michael Bratman, "Temptation Revisited," in his *Structures of Agency* (Oxford: Oxford University Press, 2007); and Richard Holton, *Willing, Wanting, Waiting* (Oxford: Oxford University Press, 2009), Chapters 1 and 7.

⁴ "The Toxin Puzzle," *Analysis* 43 (1983), 33-6. The case has generated a considerable literature. Both Bratman and Holton use it to frame their views of the stability of intention (see the previous note).

⁵ Stage one echoes an example coined by Keith DeRose to motivate contextualism about knowledge attributions ("Contextualism and Knowledge Attributions," *Philosophy and Phenomenological Research* 52 (December 1992), 913-929).

⁶ We could easily imagine a richer disposition to inquire, if that would make it more vivid that your deliberation is now open. And note one important caveat. Since it would

violate evidentialism to treat a context-setting practical consideration as a reason *within* doxastic deliberation, we must think of it as merely setting a framework for it. (For discussion of this point, see David Owens, *Reason without Freedom*, Chapter 2.)

⁷ Note that if we imagine the case with a lowering of the standard from stage one to stage two it is harder to get the intuitive result we're after, because the mere fact that the stakes have become lower may not on its own assuage your worry that you aren't really entitled to draw the conclusion. This talk of a rising and falling standard is mere shorthand for whatever shifts between the contexts. (For that issue, see Jonathan Schaffer, "What Shifts? Thresholds, Standards, or Alternatives?," in G. Preyer and G. Peter (eds), *Contextualism in Philosophy* (Oxford: Oxford University Press, 2005), 115-130.)

⁸ We can leave it as an open question whether there could be an analogously graded notion of intention. (Note that the present angle on this question is quite unlike Richard Holton's angle in *Willing, Wanting, Waiting*, Chapter 2.)

⁹ One way to elaborate the schematic description is to say that when you treat your evidence as ruling out a range of relevant alternatives to *p*, you count as doxastically committed to *p* to that degree. Note that in stage two of the case you treat your evidence as ruling out a smaller range of relevant alternatives than in stage one – simply because your shift in doxastic context from stage one to stage two includes a shift in which alternatives to 'My bank is open Saturdays' you treat as relevant. (In stage one, for example, you are not worried about the possibility that you're remembering a different bank's hours, but in stage two you are.)

¹⁰ For a discussion of what turns on these differences, see John MacFarlane, "The Assessment Sensitivity of Knowledge Attributions," in Gendler and Hawthorne, *op. cit.*

¹¹ Note well: that attribution of irrationality is from your perspective as you form the intention. You need not expect that you will not undergo a preference reversal and change your mind about the rationality of following through.

¹² Though it is a minority opinion within the now-extensive literature on the Toxin Puzzle, some philosophers argue against this conclusion. See, for example, David Gauthier, “Rethinking the Toxin Puzzle,” in J. L. Coleman and C. W. Morris (eds), *Rational Commitment and Social Justice* (Cambridge: Cambridge University Press, 1998); and Edward McClennen, *Rationality and Dynamic Choice* (Cambridge: Cambridge University Press, 1990), 227-231. Holton frames his account as “broadly consistent with McClennen’s” (*Willing, Wanting, Waiting*, 141, n. 5), but he agrees with Kavka (and Bratman) that you cannot form the intention in Kavka’s case (or any other one-off case, such as our elaboration).

¹³ An analogy between practical and doxastic Ulysses cases was first proposed by Bas van Fraassen in “Belief and the Will,” *The Journal of Philosophy* 81 (May 1984), 235-256, and “Belief and the Problem of Ulysses and the Sirens,” *Philosophical Studies* 77 (1995), 7-37. For objections to his treatment specifically of Ulysses cases, see Richard Foley, *Intellectual Trust in Oneself and Others* (Cambridge: Cambridge University Press, 2001), Chapter 6.

¹⁴ This may be too strong; perhaps you must merely deem it sufficiently probable that you won’t give in to temptation. That won’t matter to the argument that we’re pursuing.

¹⁵ For the concept of precommitment, see Jon Elster, *Ulysses Unbound* (Cambridge: Cambridge University Press, 2000), Chapter I. For a full argument for this claim, see Edward S. Hinchman, “Trust and Diachronic Agency,” *Noûs* 37 (2003), 25-51, section VII.

¹⁶ This is the General Reflection Principle stated in “Belief and the Problem of

Ulysses and the Sirens,” 16. The refinements that figure in Special Reflection won’t matter to our issue. (For Special Reflection, see “Belief and the Problem of Ulysses and the Sirens,” 19. This is the principle defended in “Belief and the Will.”)

¹⁷ See, for example, the work by Foley cited in note 13 above.

¹⁸ A defense of Reflection need not, of course, depend on the analogy.

¹⁹ Note that this is not like the case we considered in section IV in which you regard your future self as *mistaken* about its predicament.

²⁰ This third sort of case shows that a difference in doxastic context need not involve the idea that the doxastic standard goes ‘up’ or ‘down.’ Sometimes doxastic contexts are merely *different*.

²¹ See, for example, the first two papers cited in note 2 above. (Elga’s paper focuses directly on the parallel between the intrapersonal and the interpersonal.) We’ll leave it at “more or less equal” because we won’t consider whether the idea should be formulated in terms of precisely splitting the difference (as opposed, for example, to the more complex norm proposed in the third paper cited in note 2).

²² Somewhat more precisely put, you and the epistemic peer have exactly the same evidence whether p, and you have no reason to doubt the peer’s judgment in reaching a conclusion whether p. This use of ‘epistemic peer’ is compatible with Elga’s refinement of Kelly’s usage in note 21 of “Reflection and Disagreement.”

²³ Such difference-splitting requires that we frame the issue in degrees of belief. Though the need for a degreed notion here does not rest on the worry that we assuaged in section II, that discussion may help assuage the broader worry that our framework simply cannot admit

degrees of belief.

²⁴ For further discussion of such cases, see Edward S. Hinchman, “Assurance and Warrant,” forthcoming in *Philosophers’ Imprint*.

²⁵ Of course, not all the global warming debates are between epistemic peers! But the description we’re considering applies to perhaps some of them.

²⁶ Your interlocutor here is presumably not an epistemic peer. But see the previous note. Behind the simple example lies a more complex phenomenon.

²⁷ As we saw in section IV, you fail to project in the original courier case only when you expect that your future self will misinterpret its context.