

# Integrity Over Time: Korsgaard and the Unity Criterion

By Maximilian de Gaynesford

Estates have wings, and hang in Fortune's pow'r  
Loose on the point of ev'ry wav'ring Hour;  
Ready, by force, or of their own accord,  
By sale, at least by death, to change their Lord.

Alexander Pope, *Imitation of Horace*, 2<sup>nd</sup> Epistle of Second Book, 248–51.

## 1. Introduction

**M**OST PREVIOUS DISCUSSIONS OF PERSONAL INTEGRITY HAVE FOCUSED ON THE questions:

- (i) What is the nature of integrity?  
and  
(ii) What is necessarily involved in the integrity of a person?

Question (i) can seem intractable, asked straight out like that; the entry for “integrity” in the OED is certainly discouragingly various and contradictory. Question (ii) has seemed more hopeful, and it is generally expected that, in answering it, we can answer (i).

Discussions of (ii) reveal almost as many conceptions of integrity as there are contributors to the debate. But faced with the leading writers, and allowing for permutations and sub-divisions with further contrastive possibilities, the following four options seem to capture the basic positions:

- (a) *The Unity Criterion*: Integrity is essentially a matter of being united in agency; being integrated, intact, whole, undivided, or harmonious .

Advocates of the *Unity Criterion* include Christine Korsgaard, John Cottingham, James Griffin, Gabrielle Taylor, and Valerie Tiberius.<sup>1</sup>

- (b) *The Authenticity Criterion*: Integrity is essentially a matter of being wholehearted; being true to one's (“true”) self; identifying with

*Maximilian de Gaynesford, formerly Fellow of Lincoln College Oxford, is Professor of Philosophy at the University of Reading. His interests centre on the philosophy of mind and language, and include Wittgenstein, Heidegger, and philosophical aspects of poetry and film. He is the author of I: The Meaning of the First Person Term (2006, OUP), Hilary Putnam (2006), and John McDowell (2004).*

---

one's actions, or motivations, or projects.

Advocates of the *Authenticity Criterion* include Bernard Williams, Harry Frankfurt, and Martin Hollis.<sup>2</sup>

- (c) *The Constancy Criterion*: Integrity is essentially a matter of remaining consistent, steadfast, resolute, singleminded, devoted.

Advocates of the *Constancy Criterion* include John Rawls, Alasdair MacIntyre, Michael Slote, and Lynne McFall.<sup>3</sup>

- (d) *The Incorruptibility Criterion*: Integrity is essentially a matter of remaining pure, innocent, decent, or free from taint.

Advocates of the *Incorruptibility Criterion* include Sissela Bok.<sup>4</sup>

In arguing for competing conceptions, and trying to resolve the issues arising, discussions of what is necessarily involved in the integrity of a person have been curiously self-limiting. This is because they have been based around cases that restrict themselves to a single choice on a single occasion. For example, "Jim and the Indians" and "George and the Laboratory" restrict themselves to a single situation and a single decision: whether or not to shoot a person and whether or not to accept a job, respectively. In so doing, they make time figure in an incidental way only; it is merely that in the course of which the relevant persons have developed the character and attitudes relevant to making their choices. It would not matter to the way these cases are usually discussed if—somehow—the person acquired both instantaneously.

I want to discuss the possibility that—sometimes at least—we need to take temporal relations into account in determining when and whether persons act with integrity. It may be, for example, that whether or not a person is to be considered as acting with integrity depends essentially on what is the case, or on what is predicted to be the case, at two or more different times.

Parfit's case of the "Russian Nobleman" provides a useful device for investigating this issue, and particularly our tendency to be gripped by inconsistent and contradictory ideas about what integrity over time is and what is valuable about it. The case turns on relations between actual and predicted situations with different temporal locations:

*The Nineteenth Century Russian.* In several years, a young Russian will inherit vast estates. Because he has socialist ideals, he intends, now, to give the land to the peasants. But he knows that in time his ideals may fade. To guard against this possibility, he does two things. He first signs a legal document, which will automatically give away the land, and which can only be revoked by his wife's consent. He then says to his wife, "Promise me that, if I ever change my mind, and ask you to revoke this document, you will not consent." He adds, "I regard my ideals as essential to me. If I lose these ideals, I want you to think that I cease to exist. I want you to regard your husband then, not as me, the man who asks you for this promise, but only as his corrupted later self. Promise me that you would not do what he asks." [ . . . ] The young man's ideals fade, and in middle age he asks his wife to revoke the document. Though she promised him to refuse, he declares that

---

he now releases her from her commitment . . . [She believes] that she is not released.<sup>5</sup>

Discussions of this case have tended not to frame it in the most perspicuous way. Parfit told the story, but then concentrated on the wife's position and actions alone. Those who have taken up the story—Williams, Elster, Velleman, Korsgaard, Brink, Lenman, Holton—have concentrated instead on the husband's position and actions alone. One idea, whose plausibility I hope to demonstrate, is that the relevant issues for either partner fail to come into focus unless the positions and actions of both are treated with equal centrality, equal attentiveness.

The story is particularly useful for our purposes because it is far from offering a paradigm of integrity-behaviour, one which most could be expected to agree about. There may be something worth admiring in the married couple's actions, for example, but it is really not obvious and self-evident that they manifest integrity. Some would say they do; some would say they do not. In getting clear about the reasons for disagreement, we might begin to get clear about integrity over time.

## 2. Integrity Criteria and Time

CAN THIS PAIR BE REGARDED AS ACTING WITH INTEGRITY?

Here are some arguments, for and against. They are not meant to offer an exhaustive representation of the options available, but to locate features that discriminate between different criteria of integrity in their application to cases essentially involving time.

### (A) *The Authenticity Criterion*

Those who believe in the *Authenticity Criterion* may easily argue that the husband *does* act with integrity, both when young and when middle-aged.

(A.i) When young, he has socialist ideals that generate projects with which he identifies himself. He is so wholehearted about this that he not only contracts to give away his estates when they come to him in the future, but extracts a promise from his wife to ensure that, come what may, the contract is not revoked.

(A.ii) When "middle-aged and cynical," he lacks ideals; the self he now identifies with looks with scorn on the naiveté of his youth. But he may nevertheless continue to act with integrity if this is also a self to which he can be true—by striving against doing what he no longer believes in, for example, or at least by not pretending he has ideals when he lacks them. The youth recognizes in himself this durable capacity for identifying wholeheartedly with whatever he takes himself to be. The point of extracting the promise from his wife, after all, was precisely to prevent what he regarded as the evil effects of combining middle-age with authenticity.

We have less information about the wife—despite the fact that Parfit raised the case to discuss how *her* actions might be defended. But this gives us leeway that can be dialectically useful. Those who believe in *any one* of the four criteria can easily argue either way. This is chiefly because of two factors which the story leaves undetermined: the identity of the persons over time, and the reasons for the wife's willingness to make her promise at one time and her desire to keep it at another.

(A.iii) Those who believe in the *Authenticity Criterion* can argue either way because what matters to them — that a person act in a wholehearted manner, one that counts as true to themselves — may, but need not, apply to the wife. That one should identify with what one does and thinks does not tell us *what* one should think and do. Evidently it is open to the wife to think differently about this. (This creates scope for too many versions of the story to discuss adequately here. I shall stick with a couple of variants throughout what follows; they give us enough shape to work on, without drowning us in possibilities.)

(A.iv) Suppose we regard the middle-aged husband and wife as the same persons as those who were party to the original promising, and suppose the wife made her promise in a wholehearted way, just because her youthful husband begged her to. We could argue that she acted with integrity then and continues to do so in refusing her middle-aged husband. This might be because what she *now* identifies herself with differs. Perhaps she *keeps* the promise because she has acquired the socialist ideals her husband sloughed off. This need not conflict with the integrity claim because consistency is not essential to this criterion. Nor need it conflict with the personal identity claim. (There is a two-way independence between identity and “identify with”: one can “identify with” conflicting types of belief and action at different times and remain the same person; equally, people can “identify with” the same types of belief and action at the same time and remain different persons.)

(A.v) Suppose we think that the identity of persons over time is such that qualitative change (of a sort consistent with this story) might destroy numerical identity: that the person of the early husband (or wife) might cease to exist; that the resultant, middle-aged husband (or wife) might be a different person. (Suppose, for example, we think that the identity of persons just consists in physical and/or psychological continuity, and that, due to lack of such continuity, the middle-aged couple are different persons from those party to the original promising.) If this failure of continuity is partly due to the fact that neither of them is now wholehearted about *anything*, then we can conclude that the wife lacks integrity *whatever* she does. If this is because the wife is now wholehearted about something else, we could argue she is acting with integrity: in refusing to revoke the document, she is being true to the way she now conceives herself to be.

#### (B) *The Constancy Criterion*

Those who believe in the *Constancy Criterion* may easily argue that the husband does *not* act with integrity, either when young or when middle-aged — at least if his forecast about his recidivism turns out to be accurate and it really is a case of allowing his ideals to fade away, that is as opposed to exposing them as false and replacing them with new ones.

(B.i) What matters to *Constancy* theorists is that he fails to remain consistent or steadfast. He surrenders his socialist ideals instead of being sufficiently devoted to them to stay resolute. Indeed, they will consider the device by which he binds his future self as evidence of this lack of integrity. Even whilst governed by his early ideals, he was lacking in single-mindedness, and sufficiently aware of this that it was necessary to construct the binding device. He *never was* a person of

integrity.

As regards the wife, those who believe in the *Constancy Criterion* can argue either way because what matters to them—that a person remain consistent, resolute—may, but need not, apply to her. Evidently such steadfastness requires, amongst other things, honoring the promises to which one is bound. But it is another question whether the wife is bound by her promise in her husband's middle-age.

(B.ii) Those who think that the identity of persons over time is such that qualitative change (of a sort consistent with this story) might destroy numerical identity can deny that there is sufficient continuity to identify the middle-aged husband with the person to whom she made her promise. Thus the promise she made is not one from which her middle-aged husband could release her. So in keeping that promise, she would be constant, steadfast, and thus acting with integrity. (This is the view Parfit takes.)

(B.iii) If we reject this view of the identity of persons, or insist that there is insufficient discontinuity between the youth and the middle-aged man to mark a change of persons, we could argue that the wife is *not* acting in a consistent or resolute way. Suppose she makes the promise, not because she has socialist ideals herself, but because she does what her husband *begs* her to do, whenever possible—say because she knows that when he *begs* it is not an everyday desire he wants satisfied, but a matter essential to him. Since the middle-aged man is the same as the youth, he has the power to release her from her promise, and he *begs* her to regard herself as so released. In agreeing to do what he *begs* at one time, because he *begs* it, and then refusing to do what he *begs* at another time, when it is evidently in her power to do what he asks, she fails to be constant, consistent, and hence lacks integrity.

(B.iv) Suppose we alter (B.iii) in this one respect: the wife makes the promise, not just because her youthful husband *begs* her to, but because she shares his ideals—both factors are necessary and jointly sufficient explanations for her willingness to promise. Then her middle-age husband's request creates a conflict that those who believe in the *Constancy Criterion* also face. They could argue that there is no way for her to be constant in her commitment to *both* her socialist ideals *and* doing what her husband *begs*—hence it is possible that there is no way for her to act with integrity.

### (C) *The Incorruptibility Criterion*

Those who believe in the *Incorruptibility Criterion* may equally easily argue that the husband—either when young or when middle-aged—does *not* act with integrity (at least if his forecast about his recidivism turns out to be accurate).

(C.i) What matters to *Incorruptibility* theorists is that in middle age he turns cynical and to that extent fails to remain pure or innocent. He surrenders his ideals altogether, rather than convincing himself of their falsity and replacing them with new ones. He may have been decent and free from taint at one time, but if he never possessed the resources to *remain* so, and knew that he lacked them (his recourse to the binding device is evidence of this), then he *never was* a person of integrity.

As regards the wife, those who believe in the *Incorruptibility Criterion*

can argue either way because what matters to them—that a person remain pure, innocent, decent, free from taint—may, but need not, apply to her. There is a significant constraint here: that innocence cannot be recovered once lost. If the reasons for which the person who first made the promise were corrupt, then that person cannot be acting with integrity whether or not they later decide to keep it.

(C.ii) Suppose the wife is not corrupt at the point her middle-aged husband asks her to give up her promise. Nevertheless, she might refuse out of a desire to drive him to despair and suicide. We could then argue that the act is corrupt and blocks the possibility of her acting in the future with integrity.

(C.iii) Suppose the wife is corrupt at the point of making the promise but is not the same person when asked to renounce it. She may latterly have taken on socialist ideals, for example—if renouncing such ideals can create a new person, then presumably acquiring them can also. If her reasons are decent, we can argue that she is then acting with integrity.

#### (D) *The Unity Criterion*

As regards the husband, those who believe in the *Unity Criterion* can argue either way. This is because time can be made to figure differently in their appreciation. What matters to them is that a person of integrity manifests unity of agency, integration, or intactness. It is another question whether being integrated requires *remaining* so—for some appreciable period (there are, presumably, no fixed criteria for this).

(D.i) Those who do not make integration depend on *remaining* integrated focus on whether a person can be considered united and intact in their action at a particular time. It is plausible to suppose that, on these grounds, both the young man and the middle-aged man act with integrity. There need be nothing disunited about their exercises of agency on each occasion. Evidently, there is considerable disunity between the young and middle-aged husband. But since what they believe and do here is separated significantly in time, it is irrelevant that what they believe and do differs markedly.

(D.ii) Those who *do* make integration depend on *remaining* integrated for some appreciable period *could* regard both as acting with integrity. But this might require holding a revised view of the identity of persons. For suppose we think that the identity of persons over time is such that qualitative change (of a sort consistent with this story) might destroy numerical identity, and we agree with the young man that there is insufficient continuity to identify the person he is with the person of the middle-aged man. Then we cannot regard the differences between the two as undermining the integration of either one. The young man manifests unity of agency because he remains integrated *so long as that person exists*; likewise the middle-aged man. In other words, we would have a reason that is not *ad hoc* to deny that the “appreciable time” over which the young husband would have to *remain* integrated to count as satisfying this interpretation of the *Unity* criterion must extend to that point in the future when he inherits the estates and must decide what to do.

(D.iii) If, on the other hand, we reject this view of the identity of persons, or retain it while insisting that the discontinuity between the attitudes of the

youth and the middle-aged man is *not* sufficient to mark the end of one person and the beginning of another, we can argue that there is a lack of integrity here. This might be the view most *Unity* theorists would take.

(D.iv) A weaker version of this view distinguishes the youth from the middle-aged man. The youth's adherence to his ideals enables him to manifest unity of action for a sufficient time to count as remaining integrated. At some point, he loses intactness and with it integrity, turning into the middle-aged man.

(D.v) A strong version of this view denies that either the youth or the middle-aged man manifest integrity. Certain pieces of evidence—such as the youth's dependence on a device to bind himself in the future—reveal that he never was intact; not even when he adhered to his ideals.

As regards the wife, those who believe in the *Unity Criterion* can argue either way because what matters to them—that a person be united in agency, integrated, whole—may, but need not, apply to her.

(D.vi) Those who do not make integration depend on *remaining* integrated can argue straightforwardly that both the young and the middle-aged woman act with integrity. What they believe and do is significantly separated in time, so it is irrelevant that what they believe and do differs markedly.

(D.vii) Those who *do* make integration depend on *remaining* integrated for some appreciable period *could* regard both as acting with integrity. They would not need to take a particular view of the identity of persons to do so.

(D.viii) Suppose both the husband and the wife remain the same persons throughout. We could regard the unity of the wife as manifest in her willingness to keep her promise, to sustain the course of action she entered upon in making that promise, despite her husband's recidivism.

(D.ix) Suppose that the husband and the wife have become different persons. We could not then base the integratedness of the wife on the fact that she is the same *person* as the one who made the promise. But we could argue that her unity of agency is manifest in the fact that she bases her integratedness as the person she is *now* around a commitment to carrying out the projects to which another person (her 'earlier self') was committed.

### 3. The Unity Criterion

I HAVE GIVEN A FIRST DESCRIPTION OF SEVERAL DIFFERENT VIEWS ABOUT INTEGRITY AND time. These views make different claims about ordinary people and ordinary lives, even though we discussed them in terms of a fictional case.

Which—if any—of the criteria of integrity are correct? To answer this, we need to know more about what these criteria entail.

In the remainder of this paper, I shall look more closely at the *Unity Criterion*. This is not because it is the *dominating* view of integrity. There is no such thing. (That is why the so-called "Integrity Objection" seems intractable.) It is simply because it seems to me the most *interesting* view to examine. As we have just seen, the *Unity Criterion* appears—at first glance, anyway—to be the most "open" of the alternatives. It is natural enough for *Authenticity* theorists to argue that the husband *does* act with integrity, and equally natural for *Constancy* and *Incorruptibility* theorists to argue that he does *not*. Not so for *Unity* theorists;

it is unclear what, if anything, they would find it “most natural” to say about the husband. Because the *Unity Criterion* makes the case more troublesome, it is more interesting to consider it. It may also be more *fruitful*.

I have suggested that it would not be implausible for a *Unity* theorist to regard the husband as acting with integrity. This is so even if such a theorist took the view that to *be* integrated and united in agency requires *remaining* so for some appreciable period.

Korsgaard endorses the *Unity Criterion*.<sup>8</sup> But she argues in the diametrically opposed way.<sup>9</sup> Call the point at which the husband decides to ask his wife for her promise *Request Time*. On her view:

(Main Conclusion) The youthful husband could not be regarded as acting with integrity at *Request Time*.

This is so even if we *deny* that to be integrated and united in agency requires remaining so for some appreciable period. For the youthful husband is, even at *Request Time*, “a mere heap of unrelated impulses,” and thus “in a condition of war with himself” — indeed, more precisely, in a condition of “civil war.”<sup>10</sup>

Before examining this view, I should say at once that my intention is to keep the discussion in bounds and focused on the issue which immediately interests us: integrity over time. In effect, this means conceding to Korsgaard aspects of her position which lead away from that focus—and, in particular, granting certain controversial Kant-inspired claims about marriage and reasons for action which would, on other occasions, call for discussion.

Korsgaard offers two separable arguments for her main conclusion. The first argument aims to show that

(I) The youthful husband’s attitude to his wife is such as to make acting with integrity impossible for him at *Request Time*.

Call the argument for (I) the *Marriage Disunity Argument*. The second argument aims to show that

(II) The youthful husband’s attitude to his own future self is such as to make acting with integrity impossible for him at *Request Time*.

Call the argument for (II) the *Life Disunity Argument*.

If they are sound, Korsgaard’s two arguments show independently that, *regardless of what subsequently occurs*, the youthful husband could not be regarded as integrated and united in agency at *Request Time*; that it would be inconsistent with Parfit’s story to suppose he could be; and that since integrity is a matter of being integrated and united in agency, the youthful husband could not be regarded as acting with integrity at *Request Time*.

It may immediately seem that there must be two significant differences between these arguments: that the *Marriage Disunity Argument* concerns a *different* person and the *present* time, whereas the *Life Disunity Argument* concerns the *same* person and a *future* time. But this might be misleading for two reasons.



Distinguishing the arguments in this way might make it seem that considerations of time are significant to the *Life Disunity Argument* and not to the *Marriage Disunity Argument*. This would furnish grounds for an immediate objection: that the *Marriage Disunity Argument* is irrelevant to the issue of integrity over time. But this would be false.

The *Marriage Disunity Argument* turns on the husband's attitude to his wife *as his wife*—that is, as someone whose future he is bound to by the married state. It is precisely his attitude to that future which must make him disunited, on Korsgaard's view. Conversely, the *Life Disunity Argument* turns on the youth's attitude to his future self, that is as someone whose past self he now is. It is precisely his attitude in the present to that self which must make him disunited, in Korsgaard's view. Hence both arguments concern both present and future.

There is a second reason to be concerned about this way of distinguishing the arguments. It might make it seem that, for the *Life Disunity Argument* to work, one would have to think that the youthful husband and the middle-aged husband are the same person. But a revisionist about the identity of persons could deny that they are the same person. Hence such a revisionist would be furnished with an immediate objection: the *Life Disunity Argument* begs the question, assuming precisely what they deny. But this would be false.

It is true that Korsgaard herself does not hold the revisionist view that would imply that a qualitative change (of a sort consistent with this story) might destroy numerical identity. She regards the youthful and middle-aged husband as the same person.<sup>11</sup> But the *Life Disunity Argument* might still go through even if she *did* hold the revisionist view.

What matters for the *Life Disunity Argument* is the youthful husband's attitude towards himself at *Request Time*; it is here that he fails to act with integrity, regardless of what subsequently occurs. It is in part because he regards the future person who will inherit the estates as not himself that he is disunited at *Request Time*. Suppose the youth is correct in his forecast of deep qualitative change. And suppose there were another philosopher, in other respects like Korsgaard, but willing to agree that the youth and the middle-aged husband are not the same person—*Revisionist Korsgaard*. Then the *Marriage* and *Life Disunity Arguments* would not differ in respect of persons: both would be concerned with the youth's attitudes toward what he takes to be—and *is*—a different person. But Revisionist Korsgaard could still apply the *Life Disunity Argument*. This is because the fact that the youth is correct in his forecast does not make him any the less disunited in his agency.

Indeed, Revisionist Korsgaard is free to diagnose the youth's prophecy as self-fulfilling: that it is precisely *because* of the attitude he takes towards his future self that the self he turns into is indeed a different person. However, this would be an additional claim, unnecessary for her purposes. The essential claim is that it is because the youth takes the attitude he does towards his future self that he is disunited at *Request Time*. And here Korsgaard and Revisionist Korsgaard are as one.

Beneath Korsgaard's own—scattered—formulations of her position, we can discern a common form to the *Marriage Disunity Argument* and the *Life*

*Disunity Argument.* In Korsgaard's view,

- (a) At *Request Time* the youthful husband simultaneously both tries to commit himself to a certain course of action and does not try to commit himself to that same course of action.
- (b) (Given (a)) The husband is "at war" with himself; engaged in "civil war"; he could not be regarded as integrated or united in agency.
- (c) (Given the *Unity Criterion*) If someone could not be regarded as integrated or united in agency, they could not be regarded as acting with integrity.
- (d) (Given (a-c)) The youthful husband could not be regarded as acting with integrity at *Request Time*—this is Korsgaard's main conclusion.

Evidently what matters for each realization of this form of argument—the *Marriage Disunity Argument* and the *Life Disunity Argument*—is the support they give (a).

Premise (a) evidently involves the youthful husband in a contradiction. The most plausible way to explain it, perhaps, is to say that of *part* of him it is true that he tries to commit himself to a certain course of action, and of *part* of him it is true that he does *not* try to commit himself to that same action. This would imply that he is fundamentally disunited, of course, and in a strong sense. Not only is he in parts, but those parts are in strenuous conflict. And that is precisely how Korsgaard sees it, hence her advocacy of (b).

Two questions are worth raising. The first is whether the evidence of conflict in (a) is indeed strong enough, or of the right sort, to support (b). If one is to count as fundamentally disunited and engaged in "civil war," it may be necessary that one be fully self-conscious and self-reflective about the fact that one is both trying and not trying to do something. The second question is whether the evidence of conflict in (a) is indeed strong enough to support (c) when combined with (b). It may be necessary, for example, that one's self-conflict be at a very deep level if it is to count as undermining one's capacity for integrity. Korsgaard herself does not address either issue, and I shall not pursue them further in what follows.

#### 4. The Marriage Disunity Argument

THE MARRIAGE DISUNITY ARGUMENT AIMS TO SUPPORT (A) BY CLAIMING THAT, AT *REQUEST Time*, the youthful husband both tries and does not try to commit himself in marriage to his wife. The argument proceeds as follows:

1. At *Request Time*, the youthful husband tries to commit himself in marriage to his wife.

Korsgaard assumes this. She takes it on faith, I think, that the youthful husband is sincere and, having entered into the married state, he acts in such a way as to make (1) true.

Korsgaard takes from Kant's *Lectures on Ethics* and the *Metaphysics of Morals* the idea that

2. To commit himself in marriage to his wife means that the husband must try to “unify his will” with hers.

By “unify his will” she means that the husband must make decisions together with her, “deliberate *together*” with her, and “arrive at a shared decision.”<sup>12</sup>

Korsgaard then argues, for reasons we will investigate, that

3. At *Request Time*, the youthful husband does *not* try to “unify his will” with that of his wife.
4. (Given 2 and 3) At *Request Time*, the youthful husband does *not* try to commit himself in marriage to his wife.

And hence

5. (Given 1 and 4) At *Request Time* the youthful husband simultaneously both tries to commit himself to a certain course of action and does not try to commit himself to that same course of action—that is, premise (a) above.

On the face of it, (2) seems a weak requirement on committing oneself in marriage. For two people, A and B, may be described as “deliberating together” even though A has the dominating and subjugating voice throughout the discussion. And they may be described as arriving at a “shared decision,” even though A obtains all she wants without taking what B wants into account, in the sense that both A and B consider themselves bound by the decision, and equally responsible for putting it into practice, and for dealing with the consequences. If this is all that is required for “unity of will,” then even committees with the most tyrannical of chairs can claim to meet it. This would not be such a problem were it not for the danger in which it places Korsgaard’s argument. For if the requirement on “unifying one’s will” is as weak as she makes it, what plausible reason can she offer to deny that the youthful husband satisfies it (that is, (3))?

Korsgaard’s argument for (3) is terse, but I think we can find within it two separable arguments.

The first turns on the idea that the youthful husband is not in a position even to *try to commit himself*:

- (i) To try to “unify his will” with that of his wife requires that the husband be capable of committing himself to his wife.

Korsgaard seems to be dependent here on some version of the claim that genuinely to be counted as trying to do something entails that one must be—or at least take oneself to be—capable of doing it.

- (ii) At *Request Time*, the youthful husband asks his wife to promise not to revoke his contract.
- (iii) Hence (ii) reveals that at *Request Time*, the youthful husband is incapable of committing himself (to anyone or anything).

This may seem an extraordinary claim, and one not required by the story. But on Korsgaard's interpretation, the wife's promise is precisely a device to ensure that, though the husband is incapable of committing himself, a commitment is nevertheless made, and kept, namely by her. She asks, rhetorically, "So what is she supposed to think of his *marriage vows*?"<sup>13</sup>

- (iv) (Given (i) and (iii)) At *Request Time*, the youthful husband cannot try to "unify his will" with that of his wife.
- (v) (Given (v)) At *Request Time*, the youthful husband does not try to 'unify his will' with that of his wife—that is, (3).

Korsgaard's second argument for (3) turns on the position in which the youthful husband puts his wife by asking for her promise.

- (vi) At *Request Time*, the youthful husband requests her promise not to revoke the contract he has made when he inherits his estates.
- (vii) If she does *not* comply with this request, given what it means to him, then it will be impossible for her to "unify her will" with her husband's.

Here also Korsgaard seems to depend on a much stronger requirement on "unifying the will" than she herself offers. For it is evidently possible for her to "deliberate together" with him, and—if she is strong-willed enough—to arrive at a "shared decision," even though that decision entails her not complying with his request. She may, for example, insist that the scheme be abandoned, and obtain her husband's agreement, even though it costs him dear. Or, if she is more accommodating, they may arrive at some compromise.

- (viii) If she *does* comply with his request, she faces a future in which she has to choose between her loyalty to her youthful husband and her loyalty to what has become of him.
- (ix) (Given viii) If she *does* comply with his request, she faces a future in which *whatever* course she takes, she has been forced to will as an independent person.

Korsgaard seems to have the following thought in mind. Suppose the wife carries out the wishes of the youth. Still, she has had to decide to do this independently of him: it was as an independent person that she chose to favour *his* wishes before those of the middle-aged man. The same would apply if she had favoured the middle-aged man instead, or taken any other course of action.

- (x) (Given (ix)) If she *does* comply with his request, it will be impossible for her to "unify her will" with her husband's.

Presumably Korsgaard does not think that, once the wife's will is "unified" with that of her husband, there is no respect in which she can will in the future as an independent person. That would be crazy. It must be that there are decisions of

*certain kinds* that she cannot make independently whilst being so united (examples she herself offers include decisions about the house and the car), and the resolution to comply with his request is one of them.

- (xi) (Given (vii) and (x)) At *Request Time*, the youthful husband makes it impossible for his wife to “unify her will” with his.
- (xii) To try to “unify his will” with that of his wife requires that the husband avoid making it impossible for his wife to “unify her will” with his.
- (xiii) (Given (xi) and (xii)) At *Request Time*, the youthful husband does not try to “unify his will” with that of his wife—that is (3).

To summarize: Korsgaard takes the view that, at *Request Time*, the youthful husband puts his wife in a position where, *whatever* she does, she cannot decide in union with him; hence he has not tried to “unify his will” with hers and has not tried to be committed in marriage to her—that is, (4). In conjunction with (1), (4) entails (5) which is an instance of (a): that is at *Request Time* the youthful husband simultaneously both tries to commit himself to a certain course of action and does not try to commit himself to that same course of action. And in conjunction with (b)-(c), (a) entails Korsgaard’s main conclusion: that is the youthful husband could not be regarded as acting with integrity at *Request Time*.

The *Marriage Disunity Argument* may be faulted. Korsgaard thinks that the youthful husband’s recourse to his device—requesting the promise from his wife—is incompatible with an attempt to “unify his will” with that of his wife. In response, we might say that the youthful husband’s recourse to this device is not simply *consistent* with the attempt to “unify his will” with that of his wife, but actually *required* by it.

To see why, consider a comment made by Korsgaard. She notes that, by making this promise, the wife “is to hold him, by holding herself, to giving up the estates,” and she asks, rhetorically, “But if she can do this, why can’t he?”<sup>14</sup> But this is a strange complaint. There are many important projects to which one cannot commit oneself without others being committed too. It may be easier for others to be so committed (just as the wife’s role in this scheme is easier, perhaps; it is not she, after all, who stands to inherit the estates. It is only by virtue of marriage that she stands a chance of benefiting from them). In such cases, we accept that those who have it easier may be able to play their role while those who have it harder cannot.

Korsgaard might want to say that it would be better if the youthful husband were stronger in himself and thus had no need of such a device. This may be so. But the fact that he is *not* “stronger in himself” does not in itself make him incapable of “uniting his will” with that of his wife in marriage. Indeed, it can *strengthen* that capacity. For the husband might say, for example, that the device is less a way of having himself bound to a certain course of action in the future than a way of constructing a “unity of will” with his wife in the present—a unity which, he hopes, will flourish through the future, based partly on their joint commitment to a common project of living their life in accord with socialist principles and without dependency on a future inheritance.

Korsgaard could reply that this is false to the story. It makes it seem as if the youthful husband's primary purpose is to "unify his will" with his wife, and requesting this promise from her is just a means of achieving that. But as Parfit tells the story, the youthful husband's primary purpose is evidently to deal with his recidivism. He does not invent or play upon his recidivism to create opportunities for union with his wife. He clearly wishes he was not a recidivist.

We can adapt the objection to meet this response. As Parfit tells the story, the husband's primary purpose is indeed to deal with his recidivism. He obliges himself to act as his youthful self judges best. He could have picked another as the means by which this is achieved, writing them into the contract and extracting a similar promise from them. But he precisely chooses to incorporate his *wife* into the means by which his recidivism is checked. In so doing, he makes his recidivism into an opportunity to "unify his will" with his wife. He does not treat her *merely* as a means in making this request of her. He freely offers her the chance to help him, precisely because he wants to ensure that she continue to have her own voice and role in this matter which touches her, as well as him, so completely.

We can strengthen this objection. The youthful husband might ask how *else* he is to live out the "unity of will" which the Kantian view of marriage enjoins on him unless he involves his wife—via her promise—as a free voice in the device by which he binds his future self. He might continue: it is surely required of him by this Kantian view that his way of dealing with his recidivism is not to hide it from his wife, or attempt to disguise it in some way, but to incorporate her—to the extent the laws of inheritance make possible—as a partner in dealing with it.

Korsgaard can reply that there *is* another option: to live out what marriage under the Kantian conception requires by finding a middle course, one which *both* his present and future self can endorse. Lenman concurs: it would resolve matters if the youthful husband thought of a possible life that is "choiceworthy" from both perspectives and lived it.<sup>15</sup> But here it is we who have recourse to the details of Parfit's story. For it is at least consistent with that story to deny that this *is* an option for the youthful husband.

As Parfit describes him, the man is in a similar position to Luther at the Diet of Worms (at least according to the disreputable way philosophers recall history<sup>16</sup>): renouncing his estates matters to him not just so *much*, but *in such a way*, that he cannot do anything else *but* renounce them. Having reached his position on the issue, there are indefinitely many things he cannot now do. Every *apparent* alternative to renouncing his estates is *actually*—for him—unthinkable. This is a situation, in other words, where he would consider it correct to say, "There is nothing else to think on the matter." His deliberative decision not to do anything else, reached on the basis of considerations that are totally decisive for him, just is the conclusion that he *cannot* do it. This is not because of a deficient capacity on his part, an inability to take another course because of some internal compulsiveness, for example, or because of some external force with which he does not identify. It is because he has, so he believes, such good *reason* to reject any other course. These rational considerations act on him like a force, and one that he cannot bring himself to overcome.<sup>17</sup> Indeed, the fact that doing anything else is simply not an option for him might naturally be taken precisely as a mark of his integrity.

If this is plausible, then the *Unity* theorist with Kantian views appears to be in an awkward position. Someone in similar circumstances to the youthful husband might decide not to renounce their estates in order to preserve their unity. This person might thereby count as a person of integrity. But this person is not like the youthful husband, and in a way that does not reflect badly on the youthful husband. For the decision not to renounce his estates is not an option for him. And the fact that this is not an option for him is plausibly to be taken as a mark of his integrity. Yet, in being unable to live out what marriage under the Kantian conception requires by finding a middle course, one that his future self could also endorse, it seems that he *cannot* be a person of integrity. How could we make this conclusion seem defensible?

One line alone seems available to the *Unity* theorist, who insists that integrity is a matter of unity and integratedness. If the youthful husband is indeed to be a person of integrity, he must defend and promote that unity and integratedness. Hence, in this instance, he *must* choose the middle course. This is so even if it means doing what, at this time, he regards as unthinkable, as that which he cannot do.

But this line would be self-defeating, and in a way that Korsgaard herself would mark out as salient. For if the youthful husband were to promote unity over and against what he feels he has no option but to do, he will be setting himself 'at war' with himself, and thus *lose* whatever unity and integratedness he has. In short, if integrity is a matter of unity, it cannot be consistent with integrity to promote unity over and against what one feels one has no option but to do.

So this is no solution to the awkward position. We have some reason to regard the youthful husband as a person of integrity. But if we retain the *Unity* criterion, at least in the way Korsgaard defends it, we must deny this. Acting with integrity is impossible for him. Now this is, of course, precisely the conclusion Korsgaard draws. But it is not for the reasons she gives. Indeed, it is for reasons that may — perhaps should — make *her* uncomfortable too: namely, it is not possible for the youthful husband to remain a person of integrity *and* care so much and in such a way about something that one course of action and no other is an option for him.

There is a way to make the discomfort more pointed, given Korsgaard's Kantian conception of marriage. Suppose that what some person A cares about so much and in such a way is his spouse, or the success of his marriage. Suppose further that there are circumstances in which this makes one course of action and no other an option for him. Is it really acceptable to propose that, this being so, it is impossible for him to act with integrity?

### 5. The Life Disunity Argument

BEFORE ADDRESSING THIS QUESTION, WE SHOULD EXAMINE THE SECOND ARGUMENT Korsgaard offers in support of her main conclusion. The *Life Disunity Argument* aims to support (a) by claiming that, at *Request Time*, the youthful husband both tries and does not try to commit his own future self to giving his inheritance away. The argument proceeds as follows:

6. At *Request Time*, the youthful husband tries to commit his own

future self to giving his inheritance away.

In drawing up his contract and extracting the promise from his wife, he acts in such a way as to make (6) true.

Korsgaard<sup>18</sup> takes from Kant's *Lectures on Ethics* and the *Metaphysics of Morals* the idea that

7. To commit his own future self to giving his inheritance away means that the husband must try to "will a law" for himself right now—that is at *Request Time*.

She explains what she means by "will a law":

8. To try to "will a law" for oneself right now is to accord normative standing to the reasons for which one is acting right now.

She explains what she means by "according normative standing":

9. To accord normative standing to the reasons for which he is acting right now, the husband must accord normative standing to the reasons for which his future self will propose to act.

She then argues that:

10. At *Request Time*, the youthful husband does *not* accord normative standing to that for which his future self will propose to act.

Putting these claims together, she argues that

11. (Given 7-10) At *Request Time*, the youthful husband does not try to commit his own future self to giving his inheritance away

And hence draws the conclusion:

12. (Given 6 and 11) At *Request Time* the youthful husband simultaneously both tries to commit himself to a certain course of action and does not try to commit himself to that same course of action—that is (a).

To summarize: Korsgaard takes the view that, at *Request Time*, the youthful husband puts himself into a particular position with regard to his future self, one which makes him incapable of trying to commit that future self to giving his inheritance away—that is, (11). In conjunction with (6), (11) entails (12), which is an instance of (a): that is at *Request Time* the youthful husband simultaneously both tries to commit himself to a certain course of action and does not try to commit himself to that same course of action. And in conjunction with (b)-(c), (a) entails Korsgaard's main conclusion: the youthful husband could not be regarded as acting with integrity at *Request Time*.

The *Life Disunity Argument* may be faulted. We might start with (9), the



claim that the youthful husband must accord normative standing to the reasons for which his future self will propose to act if he is to accord normative standing to the reasons for which he is acting right now. We might agree but block Korsgaard's inference. For we might deny that the future in which such reasons retain normative standing must include the point at which he inherits the estates. For it may be that, to count as such, reasons cannot have normative standing for an instant. But it does not follow that, to count as such, they could *never* lose that standing in the indeterminate future.

Korsgaard does not respond to this possibility. But she anticipates one way of making it seem legitimate for the youthful husband to deny normative standing to that for which his future self will act: namely, that he thinks his future self will then be acting out of weakness (for example, a lack of self-control) or out of irrationality (for example, a clouded judgement). She accepts that this *would* be a legitimate excuse. So she would endorse a qualification of (9):

- 9\*. To accord normative standing to the reasons for which he is acting right now, the husband must accord normative standing to the reasons for which his future self will propose to act—unless he thinks his future self will be acting out of weakness or irrationality.

But Korsgaard also appeals to the fact that Parfit expressly says the case is not like that of Odysseus having himself bound to the mast when sailing past the Sirens.<sup>19</sup> She takes this to mean that the husband does not think of his future self as acting out of weakness or irrationality.

It may seem strange for Parfit (or Korsgaard) to deny that the youthful husband is like Odysseus. For evidently the cases *are* alike in fundamental respects. In asking his sailors to bind him to the mast, Odysseus does as the youthful husband does in extracting his wife's promise: he uses others to prevent his future self doing what his present self does not want him to do, something that he knows his future self is otherwise liable to do. Moreover, although both Odysseus and the husband put themselves under the control of others, that is only possible because they are in control of themselves—as the captain in the ship and the husband in the marriage, respectively. It is precisely because they are both *in* control in the general case that they are able to put themselves *under* control in this specific instance.

But Korsgaard could reply: the cases nevertheless do differ, and in the relevant respect. For Odysseus anticipates that his future self, if not controlled, will act out of weakness or irrationality. But the youthful husband assumes his future self will act out of a genuine allegiance to "different values" (Korsgaard's phrase).<sup>20</sup> Hence he has no legitimate excuse for denying normative standing to that for which his future self will act.

This may be so. But it seems to produce a contradiction in her argument. For if the youthful husband does assume his future self will act not out of weakness or irrationality but genuine allegiance to "different values", then surely he is according normative standing to that for which his future self will propose to act. Granted, the reasons his future self will be acting on are not reasons with which his present self can *identify*; they are not reasons for which his *present* self would

act. But they are reasons nevertheless, and to be accorded normative standing. This produces a contradiction, of course, because it is the negation of (10).

The point can be expressed more strongly. It is not just that the youthful husband need not *deny* that that for which his future self acts has normative standing. He must *accept* that it has such standing if he is to think of his future self as acting out of genuine allegiance to “different values,” rather than out of weakness or irrationality.

The upshot is a constructive dilemma. Either the youthful husband *does* regard as reasons the “different values” on which his future self will propose to act, or he does *not*. If Korsgaard thinks he *does*, then she must reject (10): the youthful husband thinks of his future self as acting for values that are reasons, just not reasons with which he would now identify. If she thinks he does *not*, then she must acknowledge that the qualification in (9\*) offers the youthful husband a let-out. For then he has a legitimate excuse *not* to accord normative standing to what his future self will be acting on (that is at the time of acquiring the inheritance), while nevertheless according normative standing to that on which he is acting right now (that is at *Request Time*). (If this is the truth of it, then the youthful husband *would* be similar to Odysseus, of course; for Odysseus freely accords normative standing to that on which he is acting when ordering his men to bind him to the mast while denying such standing to what his future self will (try to) act on, when under the sway of the Sirens’ song.) Either way, the youthful husband may be regarded as according normative standing to what he acts on at *Request Time*. If this is correct, then there is no reason to deny that, at this point, he successfully “wills a law” for himself.

Korsgaard may reply that there is middle ground: the youthful husband can think of his future self as acting out of genuine allegiance to “different values” without thereby according normative standing to what he thereby acts out of. How she would make good on that claim I do not know. She says at one point that the youthful husband thinks of that for which his future self will propose to act not as “reasons” but as “facts to contend with, as tools and obstacles.”<sup>21</sup>

But the attempt at exclusive contrast is unpersuasive. Some of the most formidable facts we have to contend with are what others count as reasons—it is precisely their being accounted reasons that can *make* them so formidable. Similarly, the youthful husband may think of his future attitude as a tool or obstacle to be contended with or got around. But that is how it is for him at *Request Time*. It certainly does not preclude his thinking of this future attitude as providing reasons for his future self at the time of inheritance. Indeed, it is precisely because it *will* provide reasons, perhaps, that he strives so hard to prevent his future self from acting on that attitude. Obstacles to one course of action are the more formidable for being reasons for another course of action. In short, at *Request Time*, the youthful husband can think of his future attitude as *both* a reason (for his future self) and an obstacle (for his present self).

So Korsgaard would have difficulty in finding middle ground. But suppose she can. Unfortunately, that would not help her case in the long run. For there would then be another reason to fault her arguments. She would have made her case depend on an interpretation of Parfit’s story that we are by no means obliged to accept.

On Korsgaard's interpretation, the youthful husband "expects to change his mind without a reason."<sup>22</sup> But the story does not require this. Indeed, there are three alternative, more plausible scenarios consistent with the story as told.

First, the youthful husband may expect his future self to lose his socialist ideals, so that he no longer has reason to give away his estates. And *that in itself* enables him to credit his future self with a reason to change his mind. For the *lack* of a reason to act in one way can give one a reason to act in another way.

Second, the youthful husband may credit his future self with *additional* reasons to change his mind. He may consider the likelihood of his future self's having a large family to support, or of the peasants having moved away. These possibilities also enable him to credit his future self with reasons to change his mind. Here, it need not be that he has lost his socialist ideals, it is just that he will have acquired new reasons which *trump* those given by his ideals.

Third, the youthful husband may expect his future self to acquire an allegiance to more conservative views. This would again enable him to credit his future self with a reason to change his mind. Here, he *will* have lost his socialist ideals, but he will have acquired new reasons which *replace* those given by his ideals.

In short, it is a possible—not to mention more plausible—interpretation of the story that the youthful husband expects to change his mind *with* reason, either the reason that he no longer has the reason on which his earlier self acted, or the reason that he now has different reasons on which to act, some which *trump* his older reasons, others which *replace* them.

Korsgaard wants to argue that the youthful husband's recourse to his device—requesting the promise from his wife—is incompatible with an attempt to accord normative standing to the reasons for which he is now acting. In response, we might say that the youthful husband's recourse to this device is not simply *consistent* with the attempt to accord such standing to his reasons for acting at *Request Time*, but part of what maintains them as reasons for him.

To see why, consider a comment made by Korsgaard. She notes that, by asking for his wife's promise, the youthful husband fails to "will a law that he thinks he can commit himself to acting again later on, *come what may*."<sup>23</sup> But this is a strange complaint. The law he wills *is* something he thinks he can commit himself to acting again later on, at least for some indefinitely long stretch of time. Over this time, he thinks he can stand by the contract he has drawn up, by directly resisting all attempts to persuade him to revoke it, for example. He also thinks that there may well come a time when indirect means will be necessary to prevent its being revoked; that is, a time when he will himself move to revoke it, and it is only his wife's promise not to consent to this that prevents its being revoked.

We can strengthen this objection. The youthful husband might ask how *else* he is to accord normative standing to his reasons for acting at *Request Time* unless he deploys a device of the sort he sets up. He might continue: it is surely required of him by the Kantian view of acting for reasons that his way of dealing with his recidivism is not to hide it from himself, or attempt to disguise it in some way, but to ensure that he takes it into account in whatever laws he wills for

himself. The “law” he wills at *Request Time* is that *whoever* comes into the estates, those estates be given to the peasants. And the purpose of the device is precisely to ensure that this law will continue to be acted on, “*come what may.*”

Korsgaard can reply that there *is* another option for the youthful husband: to will a law for himself that *both* his present and future self can endorse. But we have already seen the problem with this claim. It is at least consistent with Parfit’s story—and may be the more plausible interpretation of it—to deny that this middle course *is* an option for the youthful husband. The man may be like Luther at the Diet of Worms. Renouncing his estates matters to him not just *so much*, but *in such a way*, that he cannot do anything else *but* renounce them. And the fact that doing anything else is simply not an option for him might naturally be taken precisely as a mark of his integrity

If this is plausible, then just as before, the *Unity* theorist with Kantian views appears to be in an awkward position. The fact that the decision not to renounce his estates is not an option for him is plausibly to be taken as a mark of integrity. Yet, in being unable to take a course that his future self could also endorse, it seems that he *cannot* give normative standing to that for which he acts (under the Kantian conception of what that requires), and thus *cannot* be a person of integrity (under the *Unity* criterion). How could we make this conclusion seem defensible?

As before, one line alone seems available to the *Unity* theorist. If the youthful husband is indeed to be a person of integrity, he must defend and promote his own unity and integratedness over time. And that in turn means he *must* choose the course that his future self could also endorse. But we saw that this line is self-defeating. For if the youthful husband were to promote unity over and against what he feels he has no option but to do, he will be setting himself “at war” with himself, and thus *lose* whatever unity and integratedness he has.

Thus we are stuck again in the awkward position. We have some reason to regard the youthful husband as a person of integrity. But if we retain the *Unity* criterion, at least in the way Korsgaard defends it, we must deny this. Acting with integrity is impossible for him. Again this is precisely the conclusion Korsgaard herself draws. But it is for reasons that may—perhaps should—make *her* uncomfortable too: that it is not possible for the youthful husband to remain a person of integrity *and* care so much and in such a way about something that one course of action and no other is an option for him.

## 6. Conclusion

I HAVE BEEN DISCUSSING KORSGAARD’S ARGUMENTS AS PART OF A MORE GENERAL INQUIRY into criteria of integrity over time. What should we conclude?

I hope to have shown that this option is, if not untenable, certainly implausible:

- (A) (i) retain the *Unity Criterion* of integrity, (ii) retain Korsgaard’s Kantian views about marriage and reasons for action; and (iii) deny that someone who cares so much and in such a way about something that one course of action and no other is an option for him could be a person of integrity or act with integrity.

Amongst alternative options for which our arguments do provide grounds is this:

- (B) (i) accept that someone who cares so much and in such a way about something that one course of action and no other is an option for him could be a person of integrity and act with integrity; (ii) accept that the youthful husband may be someone who cares so much and in such a way about something that one course of action and no other is an option for him; (iii) renounce the *Unity Criterion*.

The claims (B.i) and (B.iii) could be linked: perhaps we are directed to endorse such a person's claim to integrity precisely *because* we are being guided by adherence, not to the *Unity Criterion*, but to an alternative conception of what integrity is, of what is essential to it. For reasons discussed above, it is the *Authenticity Criterion* that would be the most natural replacement criterion here. I have tried to show in section 2 that it is the one criterion which could be made straightforwardly to support the possibility which (B.i) and (B.ii) hold open: that the youthful husband may be a person of integrity.

There is a third option that could also appeal to our arguments for grounds:

- (C) (i) accept that someone who cares so much and in such a way about something that one course of action and no other is an option for him could be a person of integrity or act with integrity; (ii) accept that the youthful husband may be someone who cares so much and in such a way about something that one course of action and no other is an option for him; (iii) reject those aspects of Korsgaard's interpretation of Parfit's story which block the possibility of (ii); (iv) retain the *Unity Criterion* of integrity.

Does endorsing claim (C.iii) entail rejecting Korsgaard's Kantian views of marriage or of reasons for action? I think not, for reasons discussed above. I have tried to show in sections 4-5 that it would be consistent with these views to regard the youthful husband as committed in marriage to his wife and as according normative standing to that for which his future self will act. Hence Kantian views of marriage and of reasons do not block what (C.i) and (C.ii) hold open: that the youthful husband may be a person of integrity.

Since (C) is the most modest and irenic of the options supported by our arguments, pending further investigation of alternative criteria of integrity over time, it is this option we have most reason to endorse.<sup>24</sup> ◻

### Notes

<sup>1</sup> Christine Korsgaard, *Self-Constitution* (Oxford: Oxford University Press, 2009). John Cottingham, "Ethics and Impartiality," *Philosophical Studies* 43, 1983, pp. 83–99; "Partiality, Favouritism and Morality," *Philosophical Quarterly* 36, 1986, pp. 357–73; "The Ethics of Self-Concern," *Ethics* 101, 1991, pp. 798–817. James Griffin, *Well-Being* (Oxford: Oxford University Press, 1986). Gabrielle Taylor, "Integrity," *Proceedings of the Aristotelian Society*, Supplementary volume, 55, 1981; *Pride, Shame and Guilt* (Oxford: Oxford University Press, 1985) pp. 108–41;

*Deadly Vices* (Oxford: Oxford University Press, 2006). Valerie Tiberius *The Reflective Life: Living wisely within our limits* (Oxford: Oxford University Press, 2008).

<sup>2</sup>Bernard Williams, "A Critique of Consequentialism" in J. J. C. Smart and Bernard Williams *Utilitarianism: For and Against* (Cambridge: Cambridge University Press, 1973) pp. 77–150; see pp. 97–118; "Persons, Character and Morality" (1976) in his *Moral Luck* (Cambridge: Cambridge University Press, 1981) pp. 1–19; "Moral luck" (1976) in his *Moral Luck* pp. 20–39; "Utilitarianism and moral self-indulgence" (1976) in his *Moral Luck*, pp. 40–53; "Replies" in J. E. J. Altham and Ross Harrison eds, *World, Mind and Ethics* (Cambridge: Cambridge University Press, 1995) pp. 185–224. Harry Frankfurt, "Freedom of the will and the concept of a person" (1976) in his *The Importance of what we Care About* (Cambridge: Cambridge University Press, 1988) pp. 11–25; "Identification and externality" (1976) in his *The Importance of what we Care About* pp. 58–68; "The importance of what we care about" (1982) in his *The Importance of what we Care About* pp. 80–94; "Identification and wholeheartedness" (1987) in his *The Importance of what we Care About* pp. 159–76; "Autonomy, necessity and love" (1994) in his *Necessity, Volition and Love* (Cambridge: Cambridge University Press, 1999) pp. 129–41; "On caring" in his *Necessity, Volition and Love* pp. 155–80. Martin Hollis, "A death of one's own," in J. M. Bell and S. Mendus eds, *Philosophy and Medical Welfare* (Cambridge: Cambridge University Press, 1988).

<sup>3</sup>John Rawls, *A Theory of Justice* (Oxford: Oxford University Press, 1972). Alasdair MacIntyre, *After Virtue* (London: Duckworth, 1985). Michael Slote, "Morality not a System of Hypothetical Imperatives," *American Philosophical Quarterly* 19, 1982, pp. 331–340; *Goods and Virtues* (Oxford: Oxford University Press, 1983); "Satisficing Consequentialism, Part 1" *Proceedings of the Aristotelian Society Supplementary Volume* 58, 1984, pp. 139–164; "Morality and Self-Other Asymmetry," *Journal of Philosophy* 81, 1984, pp. 179–192; *Common sense morality and Consequentialism* (London: Routledge, 1985); "Agent-Based Virtue Ethics," *Midwest Studies in Philosophy* 20, 1995, pp. 83–101; *Morals from Motives* (Oxford: Oxford University Press, 2001). Lynne McFall, "Integrity," *Ethics*, 98, 1987, pp. 5–20.

<sup>4</sup>Sissela Bok, *Lying* (Hassocks, Sussex: Harvester Press, 1978).

<sup>5</sup>Parfit, *Reasons and Persons*, (Oxford: Oxford University Press, 1984) pp. 327–8.

<sup>6</sup>I say this rather than "Do this pair act with integrity?" because Parfit's story provides us with a general description only, one which can be realized in interestingly distinct ways—depending on how we fill in various details, it is possible to arrive at different tokens of this general type, and this way of phrasing the question holds open the possibility that we arrive at different answers to the question depending on which such tokens we focus on.

<sup>7</sup>Parfit, *Reasons and Persons*, p. 329.

<sup>8</sup>*Self-Constitution* (Oxford: Oxford University Press, 2009) pp. 7; 19; 176. To capture the views of Plato and Kant on which she depends, Korsgaard speaks of "integrity in the metaphysical sense" and integrity "in the moral sense"; she characterizes the former as "the unity of agency" and the latter as "goodness"; but she makes clear that, for these authors as for herself, this is only a notional distinction: integrity in either sense is in fact "one and the same property."

<sup>9</sup>Korsgaard's full discussion of the case is in *Self-Constitution* chapter 9 *passim*; especially pp. 185–8; 195; 203–4. She has offered a brief version of one aspect of one of her arguments in "Creating the Kingdom of Ends: Reciprocity and responsibility in personal relations," *Creating the Kingdom of Ends* 1996, pp. 188–221; pp. 207–8.

<sup>10</sup>*Self-Constitution* pp. 204; 195.

<sup>11</sup>*Self-Constitution* p. 203: "his future self is just himself."

<sup>12</sup>*Self-Constitution* p. 190.

<sup>13</sup>*Self-Constitution* p. 188. This aspect of the story is reminiscent of a scene in the film *Croupier* (Mike Hodges, 1998): a man says to his girlfriend "I need you to be my conscience," to which she replies "Why; don't you have one of your own?"

<sup>14</sup>*Self-Constitution* p. 187.

<sup>15</sup>"The politics of the self" 2009, ms.

<sup>16</sup>Luther never said "Hier stehe ich, ich kann nicht anders" (it was the addition of a later editor), and his views about the will (formulated in *The Bondage of the Will* 1525) would have precluded his endorsing such a statement as a characterization of his position.

<sup>17</sup>For description of this phenomenon, I have drawn equally on what Harry Frankfurt

characterizes as “volitional necessity” (*The Importance of What We Care About* pp. 85–8), and what Bernard Williams calls “moral incapacity” (“Moral Incapacity” pp. 48ff).

<sup>18</sup> *Self-Constitution* pp. 203–4.

<sup>19</sup> *Reasons and Persons*, p. 328.

<sup>20</sup> *Self-Constitution* p. 185.

<sup>21</sup> *Self-Constitution* p. 195.

<sup>22</sup> *Self-Constitution* p. 203.

<sup>23</sup> *Self-Constitution* p. 203; my emphasis.

<sup>24</sup> I am grateful to John Broome, Brad Hooker, and the Moral Philosophy Seminar at Oxford.