The Moral Restriction on Practical Identities:
A Critique of Harry G. Frankfurt

Abstract

According to Harry G. Frankfurt, "someone may wholeheartedly love what is […] bad, or what is evil", and such a person would thereby have a reason to act badly, or to act evilly. I hope to establish, contra *Frankfurt, that there is a moral restriction on the practical identities a person may feasibly adopt. That is to say, a person cannot reflectively identify herself with a way of being that she simultaneously takes to be immoral. My argument is grounded in Christine M. Korsgaard's argument, presented in* The Sources of Normativity*, which is in turn grounded in Immanuel Kant's argument for the 'Formula of Humanity'. My argument for a moral restriction on practical identities supplements Korsgaard's by drawing on a discrete area in philosophy concerned with interpersonal understanding. In particular, I propose that a moral restriction on practical identities can be secured when we take a certain approach to interpersonal understanding; namely, that advocated by R.M. Gordon and Jane Heal – what I, following Heal, call the Co-cognitive approach. Co-cognition takes a unique approach to the intelligibility both of others and of oneself, so that even reflectively finding* oneself *intelligible involves the points of views of others. One's own practical identity, therefore, necessarily takes into account the points of view of others, amounting to a moral restriction on the practical identities a person may plausibly adopt.*

Contents

iii

1   Introduction

In 1971, in his seminal article 'Freedom of the Will and the Concept of a Person', Harry G. Frankfurt proposed that what it is to be a person is to have second-order volitions – that is, to take a stance as to which of one's more basic, first-order desires are the desires that are actually effective in moving one to action. This hierarchical account of the self allowed Frankfurt to retain the autonomy of the person even in the face of determinism.[1] The crucial idea is that a person's 'real self' is not simply to be equated with the way she actually behaves; rather, who a person *really* is is a matter of how she identifies herself. And, if a person manages to act in the way she really wants to act, in a way with which she identifies, then she acts freely. I take it that this is quite a natural – indeed, a very old – idea. It strikes me, for instance, as the very same idea expressed thousands of years ago by Paul in his letters to the Romans:

> I do not understand what I do. For what I want to do I do not do, but what I hate I do. [...] For I have the desire to do what is good, but I cannot carry it out. [...] Now if I do what I do not want to do, it is no longer I who do it, but it is sin living in me that does it.[2]

Paul goes on, in verse 25, to describe himself as a "slave to sin", indicating that his evil deeds are not performed freely, as an expression of his real self, but as a slave to something alien.

Frankfurt's more recent work has moved on from the theme of the freedom of the will to caring. According to Frankfurt, it is what we care about that gives us reason to act one way rather than another. A student, for instance, has a reason to work hard; an athlete has a reason to have a gym membership; a businesswoman has a reason to not miss her board meeting; and so on. I take this more recent work of Frankfurt's to be an extension of his earlier work on free will, and, as I see it, still very much informed by the concern to preserve the autonomy of the person: the student has a reason to work hard because *she* wants to do well; the athlete has a reason to have a gym membership because *she* cares about keeping fit; the business woman has a reason to go to the board meeting because *she* chose her career; *etc*. However, Frankfurt's zealous defence of the person's autonomy leads him to claim that not even *morality* can tell us what to do, independently of what we care about. In his recent book, *The Reasons of Love*, Frankfurt is quite explicit that "someone may wholeheartedly love what is evaluationally nondescript, or what is bad, or what is evil".[3] However, despite the appeal of the hierarchical account, I find this implication of the Frankfurtian view utterly unacceptable.

---

[1] Frankfurt 1971, p.20: "My conception of the freedom of the will appears to be neutral with regard to the problem of determinism."
[2] Romans 7:15-20
[3] Frankfurt 2004, p.98

As I will argue, I don't think Frankfurt necessarily needs to adopt this radically amoral position, even if he takes as much care to preserve the person's autonomy. Indeed, I think the proper working out of the Frankfurtian 'real self' account actually *entails* a moral restriction. I do not think it is true that a person can care about – and, therefore, have reasons to act in accordance with – what is 'evaluationally nondescript, or what is bad or what is evil'. Here, I take my cue from Christine M. Korsgaard, particularly her arguments in *The Sources of Normativity*.

Korsgaard's account resembles Frankfurt's in many ways. I describe in detail the similar elements of their accounts in the following chapter, Chapter 2. However, there is a deep difference between the two philosophers' views: as I will put it, Korsgaard imposes a moral restriction on what a person may care about – or, to use her term, a person's 'practical identities' – whereas Frankfurt does not. Frankfurt, does, however, impose two restrictions on practical identities, which will be the subject of Chapter 3. As I will argue, however, these two Frankfurtian restrictions alone are insufficient as they are 'person-particular': dependent solely on the particular person's aggregate of practical identities, and not serving to connect her with anything else, thereby rendering her insular. In Chapter 4, I go on to examine Korsgaard's argument for a moral restriction on practical identities, by which the reasons I may possibly have are limited by the reasons you (may possibly) have. This is just the restriction we need in order to overcome Frankfurtian insularity, because it would serve precisely to connect a person with others (by way of their reasons) and with the world (by way of value). This Korsgaardian argument for a moral restriction is based on Immanuel Kant's argument for the 'Formula of Humanity' – which, as I will show, I take Frankfurt to be as much committed to as Korsgaard. However, as we will see, the Kantian argument isn't quite successful, and in Chapter 5 I explore two ways in which it might be supplemented: the first is the Analogical argument, which I (following Korsgaard) will find unacceptable; the second is Korsgaard's own solution, which relies on Wittgenstein's private language argument. However, I also find Korsgaard's proposed solution to be problematic because I think its accompanying view of the public nature of reasons yields too demanding a view of morality, in which your reasons have normative force not only for *you*, but for me too.

So, in the final chapter, Chapter 6, I try to develop my own solution, which I base on a version of Simulation theory – namely, Co-cognition – which is part of a discrete area of philosophy concerned with interpersonal understanding. I take it that these two fields have something to say to each other, since real self theories are concerned with what it is to do something for a reason, and because much of what our understanding of others – when it comes to the field of interpersonal understanding – consists in our seeing their reasons for acting as they do. Interpersonal understanding – as the name suggests – is concerned with bringing people together, in a certain sense. As I see it, a Co-cognitive account is just what we need to overcome

the insularity of the Frankfurtian account, and to establish a moral restriction on practical identities.

My argument, basically, will be that individuals who embrace immoral practical identities are (insofar as they embrace those identities) unintelligible: we can't understand their actions because we can't co-cognize with them. But the consequences of this kind of unintelligibility go deep: it's not just that we don't 'get it', how someone could do something so despicable – but that such an individual can't make sense of her *own* actions. Insofar as an individual can't make sense of her own actions, they *aren't* actions, where actions are expressions of a person's real self. No person can constitute her real self as immoral, because such a practical identity would serve to erode, rather than constitute, her personhood. That is to say, there is a moral restriction on practical identities.

2   Frankfurt and Korsgaard: similarities and differences

As I see it, Frankfurt and Korsgaard both provide accounts of the forces that shape our lives as persons. Korsgaard begins her seminal work, *The Sources of Normativity*, with the following, rather poetic, description of these shaping forces:

> It is the most striking fact about human life that […] (w)e think of ways that things could be better, more perfect, and so of course different, than they are; and of ways that we ourselves could be better, more perfect, and so of course different, than we are. Why should this be so? Where do we get these ideas […] of a world different from our own [that] call out to us, telling us that things should be like them rather than the way they are, and that we should make them so[?][4]

These forces are not purely motivational – non-person animals and 'wantons'[5] are also subject to those mechanical pulls and pushes, so there is nothing 'strikingly human' about them. Rather, the forces that define our lives as persons' are *reasons*. Frankfurt gives these forces the special name, 'reasons of love'[6] (or, sometimes, 'commands of love'[7]) which, as we will see, is one way in which his account is to be distinguished from Korsgaard's. However, there are also many close similarities between Frankfurt's view and Korsgaard's, and my main task in this chapter will be to outline them. For these purposes, I will speak generally of 'reasons', by which I will mean to include both Frankfurt's and Korsgaard's notions of those shaping forces. There are, however, significant differences between the two philosophers' view, which will emerge towards the end of this chapter, and which will be taken up in detail in subsequent ones.

2.1   Caring and practical identity

Both Frankfurt and Korsgaard agree that reasons exist because persons exist. As it is often put, reasons are 'mind-dependent': without any minds, or persons, there would be no such things as reasons. This is because, on both accounts, reasons are the product of a person's taking a particular view of herself: in Frankfurt's terms, they are the product of what the person cares about;[8] in Korsgaard's, they are the product of her practical identity.[9] Korsgaard, for one, is aware of the large resemblance between her and Frankfurt's accounts, remarking at one point,

---

[4] Korsgaard 1996, p.1
[5] See Frankfurt 1971, p.11.
[6] This is the title of his (2004) book.
[7] Frankfurt 2004, p.29
[8] Frankfurt 2004, p.50
[9] Korsgaard 1996, p.101

"The affinity of my account with Frankfurt's should be obvious."[10] For instance, compare the following quote from Frankfurt, concerning caring:

> the self-awareness that is characteristic of human beings makes us susceptible to an inner division in which we separate from and objectify ourselves. This puts us in a position to assess the motivating forces by which we happen to be impelled, and to determine which of them to accept and which to resist. When various motivating forces within us conflict, we are generally not passive or neutral with regard to how the conflict is to be settled.[11]

with this very similar one from Korsgaard, concerning practical identity:

> For our capacity to turn our attention on to our own mental activities is also a capacity to distance ourselves from them, and to call them into question. […] I desire, and I find myself with a powerful impulse to act. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I act? Is this desire really a *reason* to act [?][12]

Strictly speaking, these two philosophers do mean comparatively different things by 'caring about' and 'having the practical identity of' (to be discussed in due course), but, once again, there is clearly a large degree of overlap. As such, I will use the terms interchangeably. As I will put it, the difference between Frankfurt's and Korsgaard's views is not that they mean different things by 'caring' and 'practical identity', but that Korsgaard places a restriction on practical identity which Frankfurt does not.

On this point, I should also note that, although I will tend to speak in the singular of 'a person's practical identity' or 'the object of a person's care', any one person will generally embrace many practical identities and care about various things.[13] Our lives, therefore, abound with reasons.

2.2   Descriptive identity

The significance of introducing the notion of practical identity (caring about something) is that a person's practical identity may diverge from the way she actually is. Thus, Korsgaard defines practical identity as

> not a theoretical [conception of one's identity], a view about what as a matter of inescapable scientific fact you are. It is better understood as a description under which you value yourself, a description under which you find your life to be worth living and your actions to be worth undertaking.[14]

I will call this 'theoretical' view of oneself a *descriptive identity*, to be contrasted with one's practical identity. The difference between these two kinds of identities is, roughly, the difference

---

[10] Korsgaard 1996, fn.8, p.99
In a later work, Korsgaard briefly but explicitly lists the similarities she sees between her and Frankfurt's accounts. (See 2006, p.55)
[11] Frankfurt 2004, p.18
[12] Korsgaard 1996, p.93
[13] Korsgaard 1996, p.101
[14] Korsgaard 1996, p.101

between what a person is actually like, and how she would like to be. Consider, for instance, Frankfurt's unwilling addict:[15] this person knows what he's like, knows he's a drug addict – that is his descriptive identity. However, this is a descriptive identity he *rejects*; it is not a description of himself 'under which he finds his life to be worth living'. His practical identity – that description of himself which he *embraces* – is, rather, that of a sober person. Of course, though, there need not be a discrepancy between our descriptive and practical identities, and for those of us who are happily free, our descriptive and our practical identities will coincide.

2.3   Real selves

When a person's descriptive identity conforms to her practical identity, her actions – how she 'actually is' – will be an expression of herself. This is because a person's practical identity serves to constitute her 'real self'.[16] Since a person will generally have many practical identities, her real self is, strictly speaking, the aggregate of all these practical identities.

I take this to be quite a familiar idea, now well established in the literature, and expressed in commonsense analyses of life. It is not uncommon, for instance, to hear a person say that she will not allow her disability/her past experiences/other people's perceptions of her to 'define her'. I take it that what such a person thereby does is she *identifies* herself with a way of being *despite* her descriptive identity, how she happens to be. And her so identifying serves to carve out who she 'really is' – her real self. Consider again the unwilling addict, who does not let his addiction 'define him': because he rejects his descriptive identity as an addict, though his drug-taking behaviour issues from his own body – in Korsgaard's words, though he is the 'location of a causally effective desire'[17] for drugs – his acting on this desire is not an expression of who he really is. It is in this sense that the unwilling addict "may meaningfully make the analytically puzzling statements that the force moving him to take the drug is a force other than his own, and that it is not of his own free will but rather against his will that this force moves him to take it".[18] The unwilling addict's real self is repressed and inhibited in action.

Contrast this with a happier case, though – say, that of someone with the practical identity of a parent. Although she would love to lie in on her day off, instead of having to get up early to make her son lunch and take him to school, she rejects this desire, and identifies herself instead with the desire, weak though it may be, to get out of bed. That is to say, she adopts the practical

---

[15] See Frankfurt 1971, p.12.
[16] I borrow this term from Wolf (1990).
[17] This is Korsgaard's (1996, p.228) phrase.
[18] Frankfurt 1971, p.13; see also Frankfurt 1988, p.87.

identity of a parent, and thereby constitutes her real self: *she* is not the sort of person who chooses lazing around in bed over getting her son to school on time. As Korsgaard puts it,

> When you deliberate, it is as if there were something over and above all of your desires, something which is *you*, and which *chooses* which desire to act on. This means that the principle or law by which you determine your actions is one that you regard as expressive of *yourself*. […] Your reasons express your identity, your nature[19]

And, if this parent manages to dig deep and conjure up the motivation to roll out of bed and do what needs to be done for her son, her getting up early on her day off is therefore an expression of her real self as a good parent. When our actions conform to our practical identities – that is, when we act according to our reasons – our actions are expressions of our real selves.

2.4   Practical identity and personhood

An individual, such as the unwilling addict or the loving parent, described above, who constitutes her real self by way of her practical identities thereby comes to have an identity that is not reducible to her descriptive identity[20] – in short, she comes to have identity as a *person*. Any behaving system (a dog, the weather, a kitchen mixer) has a descriptive identity, a way in which its behaviour can be described (excited, windy, rhythmical), so persons, being behaving systems themselves, also have descriptive identities – but what sets us apart is that we, unlike the lower animals and inanimate things, can have an opinion on our descriptive identities: we may embrace or reject them. And it is this embracing or rejecting of a descriptive identity – that is, adopting a practical identity – which carves out one's real self, and gives one identity not just as a mechanistically driven behaving system or wanton, but as a person, and as a particular person. And, since reasons are engendered by practical identities, it follows that only persons are the sorts of creature subject to the shaping force of reasons.

2.5   Reflection

As I mentioned, what sets us apart as persons is that we have opinions about which descriptive identities we want to be true of ourselves. As Frankfurt puts it, "(w)hen various motivating forces within us conflict" – such as the parent's desire to sleep in *versus* her desire to get her son to school on time – "we are generally not passive or neutral with regard to how the conflict is to be settled."[21] And this taking a stance on our descriptive identities requires that we be able to think about how we descriptively are – that is, it requires a capacity for reflective thought. Thus

---

[19] Korsgaard 1996, pp.100-101
[20] This is easier to see in the case of the unwilling addict, where his practical and descriptive identities diverge.
[21] Frankfurt 2004, p.18

By its very nature, caring manifests and depends upon our distinctive capacity to have thoughts, desires, and attitudes that are *about* our own attitudes, desires and thoughts. In other words, it depends upon the fact that the human mind is *reflexive*.[22]

Korsgaard places the same emphasis on reflection – in fact, two entire chapters of *The Sources of Normativity* are dedicated to its discussion.[23] In a later work, she makes her concurrence with Frankfurt on this point quite explicit:

I agree with Frankfurt's view that the distinguishing feature of human life is a form of self-consciousness – namely, our capacity to take our own mental states and activities as the object of our attention. Like Frankfurt, I think that this form of self-consciousness is the source of the distinctively human tendency to self-assessment and the resulting capacity for normative self-government. We also agree that this kind of self-consciousness is the source of normativity, or anyway makes normativity possible, and is the source of freedom of the will.[24]

As we will see, this reliance on reflective self-consciousness has important implications for these two philosophers' views. In the end, I think it is this centrality of reflection to both Korsgaard and Frankfurt's accounts, and the particular Co-cognitive view I advocate, that secures the conclusion, *contra* Frankfurt, that there is a moral restriction on practical identities. This, though, must wait till Chapter 6.

2.6   Possible descriptive identities

However, this emphasis on *reflection* does, I think, run the risk of implying that a person's practical identity can only take the form of either her embracing or her rejecting her descriptive identity – that is, her taking a stand on *her* descriptive identity, how she *actually* is. That the practical identities available to a person are confined to her taking a stand on her actual descriptive identity is suggested by descriptions such as the following:

I desire, and I *find myself with* a powerful impulse to act. But I back up and bring that impulse into view and then I have a certain distance. Now the impulse doesn't dominate me and now I have a problem. Shall I act?[25]

And this confinement is more than just suggested by Frankfurt in the following:

Now, when the statement that *A* wants to *X* is used [to denote *A*'s reflectively embracing *X*], it does entail that *A* already has a desire to *X*. […] It is only if he does want to *X* that he can coherently want the desire to *X* not merely to be one of his desires but, more decisively, to be his will.[26]

---

[22] Frankfurt 2004, p.17; see also Frankfurt 1971, pp.7&13; and1988, p.83
[23] Korsgaard 1996, Lec.2&3, pp.49-130
[24] Korsgaard 2006, p.55, referencing Frankfurt 2006, pp.4-5
[25] Korsgaard 1996, p.93 (emphasis added)
Similarly: "When an impulse – say a desire – *presents itself to us*, we ask whether it could be a reason" (1996, p.113), and: "When we *experience* a desire or an impulse, we consider whether to treat it as a reason" (p.140) (emphases added).
[26] Frankfurt 1971, pp.9-10
A person's 'will', on the Frankfurtian account (p.8), is the desire that is causally effective in moving her to action (or would be, if she had freedom of action).

However, Frankfurt does go on to consider, in a footnote to this passage, a counterexample to his claim, in which "*A* admires *B* so fulsomely that, even though he does not know what *B* wants to do, he wants to be effectively moved by whatever desire effectively moves *B*; without knowing what *B*'s will is, in other words, *A* wants his own will to be the same"[27] – that is, *A* adopts a practical identity that is *not* a direct embracing or rejecting response to his own *actual* descriptive identity. Furthermore, Frankfurt does allow that it is possible for a person to 'bring it about' that she cares for something[28] (i.e., to want to have a descriptive identity that is not currently true of herself), and that "(i)t is generally possible for us to imagine ourselves loving things other than those that we do love".[29] Thus, I see no reason why practical identities ought to be confined in this way, and it strikes me as quite plausible that we might formulate practical identities also in response to, not actual, but *possible* descriptive identities, possible ways of being. Indeed, I think it's oftentimes the practical identity that comes first ("That's the kind of person I want to be") and, if one is lucky, the descriptive identity that follows.

The same applies to Korsgaard. I think her references to 'finding oneself with an impulse', impulses' 'presenting themselves' to one, and 'experiencing an impulse' as the precipitants of practical identity formation is just a convenient manner of speaking. She endorses the Kantian denial that "all of our desires simply spring up in us, like mushrooms in the back yard, to be approached with caution", adding that "our first-order desires and impulses do not all derive immediately from instinct, like those of the other animals, but rather arise from a complex interplay of instinct and reason."[30] This I take to imply that, intuitively quite correctly, Korsgaard also allows that a person can adopt a practical identity that is not a simple embracing or rejecting response to an actual descriptive identity of hers – one that has 'sprung up' in her, as the product of instinct – but that our practical identities may also take as their objects conceivably *possible* ways of being.

2.6.1 Thinking about possible ways of being

If this is the case, there are two implications we must note, though. Firstly, this means we must also credit persons with a capacity to think about the non-actual – possible ways of being – over and above our capacity for reflection. However, I take our ability to think hypothetically to be quite uncontroversial,[31] so this is hardly an undesirable consequence of allowing for possible descriptive identities. Indeed, I take it that our reflecting just *is* our thinking hypothetically about ourselves – or, as I will put it towards the end of this project, our *co-cognizing* with ourselves.

---

[27] Frankfurt 1971, fn.4, p.10,
[28] Frankfurt 2004, p.44
[29] Frankfurt 2004, p.48
[30] Korsgaard 1996, pp.238-239
[31] See Heal 1996b/2003, p.65; and 1998a/2003, pp.36-37.

2.6.2 Reasons without motivation

The second implication we must note is that, if we allow practical identities to take as their objects possible descriptive identities, and if reasons are engendered by our practical identities, then this means that a person may have a reason to Φ but have no *motivation* to Φ, because her being a person who Φs is just a possible, and not an actual, descriptive identity for her. Once again, I do not take this to be an undesirable consequence. In fact, I think it's quite an accurate – though somewhat depressing – description of our lives: sometimes we acknowledge that we have a reason to do something, yet we just have *no* motivation to do it. In some cases, this may be a merely temporary 'teething phase': in time, the person may be able to bring it about, by way of her practical identity as someone who Φs, that she has the desire to Φ and is thereby motivated to Φ. But, for some unfortunate persons, this may be a more permanent state; a person "might discover that he does not have and that he does not subsequently develop the feelings, attitudes and interests constitutive of the sort of person which his [practical identity] has committed him to being",[32] and this is just a sad fact of life.

2.7 Volitional necessity

What a person cares about imposes, in Frankfurt's terms, 'volitional necessity' on her;[33] she is constrained to act in certain ways. Most parents, for instance, will attest that they simply cannot help caring about their children, and almost all of us cannot but care about our own survival.[34] Or, consider a person with the practical identity of a vegetarian, and how she might excuse herself from an omnivorous dinner with: "Oh, no, I can't eat meat." Of course, there is an obvious sense in which she *can* eat meat (which her unsympathetic host may well bring up): she's got teeth, and she's got the necessary digestive enzymes. But, because she identifies herself by way of her practical identity as a vegetarian, *she* cannot eat meat – if/when she does eat meat, it is not an expression of her real self, it's not *her* acting in this way; *she* cannot eat meat. She's not physically or biologically constrained; she's *volitionally* constrained to refrain from eating meat. Frankfurt summarises this as follows:

> The [volitional] necessities of a person's will guide and limit his agency. They determine what he may be willing to do, what he cannot help doing, and what he cannot bring himself to do. They determine as well what he may be willing to accept as a reason for acting, what he cannot help considering to be a reason for acting, and what he cannot bring himself to consider a reason for acting. In these ways, they set the boundaries of his practical life; and thus they fix his shape as an active being.[35]

---

[32] Frankfurt 1988, p.85
[33] Frankfurt 1988, p.86; 1999, pp.170-171; and 2004, p.44
[34] Frankfurt 2004, p.27
[35] Frankfurt 2004, p.50

Although she does not use the term 'volitional necessity' Korsgaard makes essentially the same point, by way of a discussion of integrity. Etymologically, the word refers to oneness, wholeness, unity – consider the etymologically related 'integer': a *whole* number, as opposed to a fraction. Yet, today, the term is most commonly used to refer to a person of upstanding character, "someone who lives up to his own standards".[36] But these two uses are not so disparate: somebody who lives up to his own standards – whose descriptive identity conforms to his practical identity – is a whole, undivided person. A person's practical identity has the force of volitional *necessity* because to act against her reasons is to sacrifice her integrity and do damage to her *self*. Consider the volitional necessity expressed by the not unusual utterances, "I couldn't live with myself if I did that"[37] or "I would rather die than…". Of course, the vegetarian, say, who gave in to the temptation of eating meat would not literally suffer bodily death as a result of it – but there would be some kind of suffering: her actions, which would not express her practical identity, would serve instead to stifle and repress her real self. A person *must* act as her practical identity gives her reason to, otherwise she hurts herself. Someone who does not act as she has reason to might describe herself as "torn" or "falling apart" because this is precisely what's happening: she is *disintegrating* by doing damage to her self.

Now, although a person is volitionally *constrained* to act as her practical identity gives her reason to, such constraint is not experienced "as alien or external"[38] to herself, and the person who acts as she has reason to does not do so with the sense that she has been forced into it. The feeling of acting as we are volitionally constrained to is not one of "dispirited passivity or confinement" but, rather, of an "invigorating release and expansion of ourselves",[39] an experience of "liberation and enhancement"[40] and an exercise of autonomy.[41] This is because a person's volitional constraint is the product of her practical identity – by way of which she constitutes herself. So, volitional necessity is *self*-imposed. Being constrained to act as you have reason to is not to be forced into anything you don't want to do – indeed, it is being free to do what you *really* want to do, as captured by your practical identity.

As I will go on to explain, the force of volitional necessity is, in fact, the force of *law*, but a self-imposed law, so that when a person hurts herself in failing to act as she has reason to, what she is actually doing is *violating* the law that she is unto herself. I will take this up in section 4.2 (particularly 4.2.3).

---

[36] Korsgaard 1996, p.102
[37] Korsgaard 1996, p.101
[38] Frankfurt 1988, p.87
[39] Frankfurt 2004, pp.64-65
[40] Frankfurt 2004, p.64
[41] Korsgaard 1996, p.106

2.8   Necessary practical identity

In section 2.4, we noted that only persons are the sorts of creature subject to reasons, because to have a reason one must have adopted a practical identity, and this requires a capacity for reflection. The point, in fact, goes deeper than this: not only must one *be* a person, one must also reflectively *conceive of oneself as* a person. This is a point that both Frankfurt and Korsgaard note: our practical identities – as parents, vegetarians, businesspeople, *etc* – are underpinned by a more fundamental conception of ourselves. That is to say, to have practical identity, one must embrace one's own nature as the sort of creature that has practical identity: "someone who cares about anything also cares about caring".[42] On Frankfurt's account, this amounts to 'loving oneself', and on Korsgaard's, 'valuing one's own humanity'. Since this fundamental conception of oneself as the sort of creature who has practical identity is that the conception of oneself as a person, I will call it the *necessary practical identity of a person*.

I contrast this necessary practical identity, which we must all posses – being a person – with our various *contingent* practical identities – being a parent, a vegetarian, a businesswoman, *etc*. Of course, my being, say, a parent, isn't contingent *for me* – it's *constitutive* of me, my real self; without the practical identity of a parent, I wouldn't be who *I* am. However, my personhood doesn't depend on *this* practical identity: even if I chose to embrace vegetarianism rather than parenthood, I would nevertheless be a person (even though I wouldn't be the *same* person: me). It is in this sense that our contingent practical identities – as parents, as vegetarians, as businesspeople – are contingent. However, my (or any person's) having the fundamental practical identity of a person is necessary,[43] and necessary to all persons. As Korsgaard puts it:

> For unless you are committed to some […] practical identity, you will lose your grip on yourself as having any reasons to do one thing rather than another – and with it, your grip on yourself as having any reason to live and act at all. But *this* reason for conforming to your particular practical identities is not a reason that *springs from* one of those particular identities. It is a reason that springs from your humanity itself, from your having identity simply as a *human being*, a reflective animal who need reasons to act and to live. And so it is a reason you have only if you treat your humanity as a practical, normative, form of identity, that is, if you value yourself as a human being.[44]

So, though you and I may share none of the same contingent practical identities, and thus be subject to none of the same reasons, we do have one thing in common (if both you and I are indeed subject to any reasons at all), and that is our personhood. Just as a particular person

---

[42] Korsgaard 2006, fn.4, p.107
[43] Not only is it necessary that I take *some* stance towards my personhood, by – as Frankfurt's (1971) willing and unwilling addicts respectively do – *either* embracing or rejecting the desire I discover within myself; I must *embrace* my personhood. A person cannot reject her personhood, insofar as she is indeed a person: someone who does will have no contingent practical identities, no real self, and thus no identity as a person as opposed to a mere behaving machine or animal.
[44] Korsgaard 1996, pp.120-121

expresses her contingent practical identity when she acts as she is thereby volitionally constrained to, any possible person expresses her necessary practical identity by adopting contingent practical identities, whatever they may be. In Frankfurt's words: "someone who loves himself displays and demonstrates that love just by loving what he loves."[45] It is in her adoption of contingent practical identities that an individual constitutes herself as a person, and a person volitionally distinct from any other.

2.9   Reasons and value

Frankfurt's claim of a person's contingent practical identities expressing her self-love and the similar Korsgaardian reference to valuing one's humanity are not, however, strictly identical claims, which brings me to this present point. The difference is not so much one of these philosophers' views of what I have called the necessary practical identity of a person as it is of their discrepant views of the role of *value* in practical identity generally.

2.9.1   Loving *versus* valuing

Note that, for Frankfurt, a person's having a contingent practical identity is a matter of her self-*love*. So far, I have had much to say about what is involved with *caring* (having a practical identity), but not *loving* – except for the brief mention that, according to Frankfurt, those forces that shape our lives as persons are the 'reasons of love'. Loving, according to Frankfurt, is a mode of caring,[46] so loving shares all of those features of caring more generally (amongst two other features peculiar to loving, discussed below). One of these general features is that what we care about *makes* that object of care important to us.[47] As we have already noted, a person does damage to herself when she fails to act as her practical identity mandates; so a person *needs* – as a matter of volitional necessity – to act as she has reason to. And, what we need is important to us,[48] so the object of a person's care is necessarily important to her. Now, when it comes to caring, generally, this need not be the only kind of worth something may have: it may have objective value apart from its subjective importance to me. Say, for instance, I care about recycling: I might do this because I think that recycling is valuable, and I might think this because I think recycling is a fruitful means to a valuable end: environmental preservation. And, in coming to stand in a special relationship to recycling by caring about it, its value is made important *to me*. However, this is not the case with loving.

---

[45] Frankfurt 2004, p.85
[46] Frankfurt 1999, pp.155 & 165
[47] Frankfurt 1988, p.92; 1999, p.172; and 2004, p.38
[48] Frankfurt 1999, p.163

As I read Frankfurt, loving is to be distinguished from caring on two counts. Firstly, loving is 'disinterested' – obviously not in the sense that it is indifferent to the beloved object, but in the sense that it is not motivated by any ulterior purposes; the good of the beloved is desired for its own sake.[49] This is not the case when it comes to my caring about recycling: I care about it *as a means of* environmental preservation; so I do not *love* recycling. The second distinctive feature of loving is that it is 'ineluctably particular': love is not a generic concern; the beloved object is not a token of some type, other tokens of which would be identically loved by the lover; the beloved object cannot be substituted.[50] This goes hand-in-hand with the first feature of loving: were I to discover that Φing, as opposed to recycling, were a better means of preserving the environment than recycling, then I would care about Φing rather than recycling; so I do not *love* recycling. Were I to *love* recycling, then, it could not be *because* I think it's a valuable activity. As Frankfurt writes:

> It is true that the beloved invariably *is*, indeed, valuable to the lover. However, perceiving that value is not at all an indispensable *formative* or *grounding* condition of the love. It need not be a perception of value in what he loves that moves the lover to love it. The *truly essential* relationship between love and the value of the beloved goes in the opposite direction [i.e. that loving *makes* the beloved object important to the lover].[51]

It is interesting that Frankfurt takes this non-responsiveness of loving to objective value to be what is 'truly essential' to it – especially when we bear in mind that, on Frankfurt's account, "(t)he origins of normativity […] lie in the contingent necessities of *love*",[52] that the forces that shape our lives as persons are the 'reasons of *love*'. That is to say, those forces that define our very identities are fundamentally *unrelated* to the external reality of objective value, on the Frankfurtian account. This stands in stark contrast to the Korsgaardian position where our contingent practical identities express our *valuing* our humanity – not just our making it important to ourselves, but our taking it to be objectively valuable. Indeed, in her earlier work, Korsgaard apparently uses 'value' and 'reason' as synonyms for each other.[53]

2.9.2 Objective value and autonomy

However, Frankfurt has good reason for reaching this conclusion, even though (as I will go on to argue) I think it is the wrong conclusion. I think Frankfurt is correct that judgements of objective value are not 'indispensable formative or grounding conditions' of our caring and our reasons. As Korsgaard notes: "Because there are many things of value, your [coming to care about *this*] is

---

[49] Frankfurt 2004, p.79
[50] Frankfurt 2004, p.44; see also 2004, p.79; and 1999, pp.169-170.
[51] Frankfurt 2004, p.38 (latter emphasis added)
[52] Frankfurt 2004, p.48 (emphasis added)
[53] Korsgaard 1996
She does draw a distinction between the two in later work, though (see 2006), which becomes clear in my later discussion.

clearly not caused merely by the fact that you see it as of value."[54] If our practical identities were a response to our judgements of value then (assuming we all judged correctly) we would all end up caring about exactly the same things, and our respective practical identities – that make us who we are – would be lost, and we would be indistinguishable from the next person.

There is a deeper reason for Frankfurt's emphasizing love over caring, though. I think Frankfurt is quite correct that something can only come to have volitional necessity over me and exert a shaping influence on my life if I count it important to me, by way of embracing the relevant practical identity, and not because it has value independently of me. If we concede objective value the power to reach into our lives and cause us to confrom to its dictates, then we relinquish our autonomy. And, if that it the case, then my acting as I then would have reason to – *because it's valuable* – is not an expression of my autonomously constituted self, and is no different in kind to falling towards the earth when I jump from a height, or coming down with a cold. Thus, I agree with Frankfurt that, "(m)ost profoundly, perhaps, it is love that accounts for the value to us of life itself. […] It is a powerful – indeed, a comprehensively foundational – generator of value",[55] insofar as this mean that it is *our embracing* – loving – something, and not that thing in itself, that gives us a reason to act one way rather than the other.

However, this is the extent of my agreement with Frankfurt, and, I take it, the extent of Korsgaard's too. As I will go on to argue in section 5.3, the above idea – that it is our *valuing*, and not the objectively valuable thing itself that imposes volitional necessity on us – is really not inconsistent with Korsgaard's view on objective value.

2.9.3   Implications of Frankfurt's position

Where Frankfurt and I (and Korsgaard) diverge is that, because Frankfurt denies the value-responsiveness of our practical identities, he allows that a person may love something that he judges to be objectively *valueless*, while I reject his apparent assumption that a denial of value-responsiveness entails a total disconnection of reason with objective value: I contend that a person cannot have a reason to Φ – as the product of her autonomously adopted practical identity – when she simultaneously takes Φing to be utterly valueless. My view on objective value will be the subject of section 5.3, but, for now, we can note the implications of Frankfurt's. He is quite explicit that his view "leaves open the possibility that someone may wholeheartedly love what is evaluationally nondescript, or what is bad, or what is evil."[56] That is to say, on Frankfurt's account, a person may adopt as a practical identity a way of being that he knows to be valueless and immoral – and, furthermore, that such a person thereby has a *reason* to behave immorally. I

---

[54] Korsgaard 2006, p.75
[55] Frankfurt 2004, pp.40-41
[56] Frankfurt 2004, p.98

reject this possibility, and concur with Korsgaard that there is a moral restriction on practical identities. First, though, let us further examine the implications of Frankfurt's view of the reasons of love as the shaping forces of a person's life.

3   Frankfurtian restrictions on practical identities

Up until this point, I have treated Frankfurt and Korsgaard as offering more or less the same account of what it is to have a practical identity and for one's life to thereby be shaped by reasons. There is, however, a significant difference between the two philosophers' positions, signs of which we have already seen towards the end of the previous chapter. The fundamental difference between Frankfurt and Korsgaard, as I will put it, is that Korsgaard places a restriction on the practical identities a person may plausibly adopt that Frankfurt does not: namely, a moral restriction. Whereas, for Frankfurt, "there is always *something* to be said in favour of caring about *any* object,"[57] because caring about something – anything – is better than caring about nothing at all, given that it is our practical identities that make us the persons we are. For Korsgaard – and myself – however, it is *not* the case that any practical identity is as good as any other.

Frankfurt does, though, place *some* restrictions on the practical identities a person may feasibly adopt, which will be the subject of this present chapter. I take it that Korsgaard acknowledges, as a start, these two Frankfurtian restrictions, although she goes on to add a third, the moral restriction on practical identities. I will argue, in line with Korsgaard's position, that just these two Frankfurtian restrictions – on inexpressible and deeply incompatible practical identities – are inadequate, my contention being that their 'person-particularity' does not establish a connection between the individual and the world in which she lives, including those other persons with whom she interacts, with the unacceptable consequence that she is rendered insular, and therefore – as I will go on to argue in Chapter 6 – unintelligible. Let us begin, then, with the two Frankfurtian restrictions on practical identities.

3.1   Frankfurt's two restrictions

We have established that, in order to express one's fundamental nature as a person, one must adopt contingent practical identities, the aggregate of which constitutes one's real self. Since our necessary practical identities furnish each of us with a reason to adopt a further contingent practical identity, the pertinent question seems to be: are some contingent practical identities

---

[57] Frankfurt 2002, p.246 (latter emphasis added)

better to adopt than others? And the answer, according to Frankfurt at least, is: it depends. As he writes,

> What makes it more suitable, then, for a person to make one object rather than another important to himself? It seems that it must be the fact that it is *possible* for him to care about the one and not about the other […] (T)he worthiness of the activity of caring commands that he choose an object which he will be able to care about.[58]

I can think of two broad ways in which a person might be incapable of caring for something, which will be the subject of sections 3.1.1 and 3.1.2, below, and which, as I will go on to explain in section 3.2.1, depend on the identity of the particular person. These are the Frankfurtian restrictions on practical identities.

3.1.1 Inexpressibility

Firstly, a person may be unable to care for something due to some 'mechanical' failure: if she lacks freedom or if she lacks motivation altogether. For instance,

a) A blind man who has his heart set on becoming the World Target Shooting Champion will find his practical identity inhibited. Given his condition, he is physically incapable of acting in the ways mandated by such an identity; no matter how much he may want to, he lacks freedom of action.

b) Similarly, Frankfurt's unwilling addict finds himself unable to express his care for sobriety. This, however, is due to a lack of freedom of the will: this character can't get his desire to refrain from taking drugs to be his will[59] in the first place, so he has no chance of expressing it, whether he enjoys (would enjoy) freedom of action or not.[60]

c) Contrast both of these cases with one in which the character lacks (strictly speaking) neither freedom of action nor freedom of the will, because the descriptive identity with which she identifies herself is only a *possible* one, as described in section 2.6.2. For example, I may take myself to have a reason to go for a run, yet find that I have absolutely *no* drive to do so, nor, frustratingly, does this motivation develop in time.

I call these cases of 'mechanical failure' because in all these examples, the problem lies in the absence or impediment of some volitionally functional element, which renders the character incapable of expressing his care. Because an unexpressed, repressed practical identity inflicts some suffering on the person, it is (*ceteris paribus*) better for her to not embrace a way of being that she is incapable of expressing.

These examples do not, however, strike me as clear-cut cases in which it is a bad idea for the character to have his/her particular practical identity. (b), I think, is the most revealing: it is near

---

[58] Frankfurt 1988, p.94
[59] That is, the first-order desire which is effective in moving the agent to action (when he enjoys freedom of action). (Frankfurt 1971, p.8)
[60] See Frankfurt 1971, pp.12-13.

impossible for the unwilling addict – being physiologically addicted – to act as his caring about sobriety mandates, to refrain from taking drugs, and this makes life very difficult for him; so Frankfurt's counsel would seem to suggest that he ought *not* to care about sobriety. But, intuitively, we think that it's in fact a very *good* idea for the unwilling addict to care as he does. I take it that this is due to certain assumptions involving the 'logic of caring', which I discuss below.

3.1.1a   The logic of caring

This is actually Korsgaard's term, which she uses to describe a Frankfurtian idea: that "caring essentially implies – or *entails*, as [Frankfurt] puts it – certain commitments that go beyond its immediate object."[61] To use an example from Frankfurt, if a person cares about his health, then this health-conscious practical identity entails that he cares about something else, something 'beyond the immediate object' of this practical identity (i.e. his health): in his example, the logic of caring entails that the health-conscious person cares also about (the avoidance of) background radiation, because background radiation is damaging to one's health.[62] Such a person's caring about background radiation is in fact an expression of some further practical identity of hers: her health-consciousness. Because "(n)early everyone cares about staying alive, […] and about avoiding severe injury, disease, hunger, and various modes of psychic distress and disorder",[63] it is safe to say that just about everyone should care about background radiation.

I think some similar considerations are at work behind the scenes in our intuitive response to the unwilling addict. Because, as we noted above, nearly everyone – unwilling addict included – cares about their health, and because drug-use is a one-way ticket to compromised health, it is quite likely entailed by the logic of caring that it is a *good* idea for the unwilling addict to care about sobriety, in spite of what it costs him in other ways.[64] This is because, even though the unwilling addict is unable to express in action his caring about *sobriety*, his caring about sobriety is in fact an expression of a further practical identity of his: as I have construed it, his caring about his *health*. So, we should note a qualifier on the Frankfurtian restriction on unachievable practical identities: a person ought not to care about something if he is unable to express his caring in action, *unless* his so caring is entailed by his other practical identities.

However, we must note that it only follows that the addict ought to care about sobriety if he does care about his health. If he does not care about his health, then (*ceteris paribus*) he has *no* reason to be sober. As Frankfurt writes with regards to caring about background radiation:

---

[61] Korsgaard 2006, p.56
[62] Frankfurt 2004, p.22
[63] Frankfurt 2004, p.27
[64] We might construe this case differently: perhaps as the addict's caring not about his health but about his reputation, or about staying out of prison, or about being employable, *etc*. Either way, I think the point stands that we take the unwilling addict's caring about sobriety to be a good idea (even though he is unable to actually express his so caring) because of further things he cares about.

> Suppose that someone genuinely does not care a bit about his health, or about any of the effects that radiation may produce. Suppose that he really is completely indifferent to whether the environment, or other people, or he himself, is or is not affected in those ways. In that case, the level of background radiation is not important to him[65]

This points to the *person-particularity* of this Frankfurtian restriction, which I will take up in section 3.2.1.

3.1.2 Deep incompatibility

The former restriction took into account, as I called it, the mechanics of caring: whether or not a person is actually able to express her practical identity in action. This second restriction is concerned, more fundamentally, with whether a person is even able to adopt a particular practical identity in the first place. A person may be volitionally incapable of caring about something, even if she is mechanically able to behave as that practical identity would mandate. As we saw above, the logic of caring may entail that a person ought to have a particular practical identity (even if he is incapable of expressing it); likewise, I think the logic of caring can also entail that a person ought *not* to adopt a certain further practical identity: when so caring would result in a conflict of reasons.

Because a person generally embraces more than one way of being, her assortment of practical identities will generate various reasons. And this may give rise to the regrettable situation in which some of a person's reasons conflict with others.[66] Unfortunate as it is, I take this to be true to everyday life: I should really attend the engagement party, but I actually just need some time to myself; I am 'proudly South African', but, as a vegetarian, I don't support 'National Braai Day'; I should go to my daughter's ballet recital, but I can't miss the board meeting; I want to buy environmentally friendly shampoo, but it's so expensive, and it doesn't even clean one's hair that well, and so on. These reasons-conflicts are not easily resolved, and any outcome always comes at a cost to the person: "(l)overs are characteristically vulnerable to profound distress if they must neglect what is required of them by one love in order to meet the requirements of another".[67] This is because, whichever way the person ends up acting, the action is simultaneously an expression of *and* a repression of herself. Say she chooses to skip the board meeting and attend her daughter's recital instead: then she expresses herself as a parent, yet represses her careerism; say she chooses the other way: then she expresses herself as a businesswoman, but represses her parenthood. (And, if she chooses to 'abstain', and not take any action, then she doesn't express herself at all, and she hurts herself twice as badly.)

---

[65] Frankfurt 2004, p.22
[66] Frankfurt 2004, p.62; Korsgaard 1996, p.126
[67] Frankfurt 2004, p.62

The examples I have listed above, though, are not cases of *deep* incompatibility. Most of the conflicts arise due to a contingent shortage of resources: if I had all the time in the world, I would attend both the board meeting and my daughter's recital; if I had more money, I would buy the environmentally friendly shampoo (and twice as much of it, to actually clean my hair), *etc*. These kinds of 'shallow' reasons-conflicts are endurable: though you will not survive wholly unscathed, you will not be torn apart. Your practical identity "can take a few knocks", as Korsgaard puts it.[68] So, a person's practical identity as a parent need not entail that her being a businesswoman is restricted. There is nothing inherently incompatible with these two ways of being, even though they may throw up a reasons-conflict every now and then; many people get along as both businesspeople and parents just fine.

The kind of incompatibility – deep incompatibility – that *does* amount to a restriction on practical identity, is not the result of having too little time or too little money. Rather, two (or more) practical identities are *deeply* incompatible when there is some inherent contradiction between them; for example, the person who cares about her health but is simultaneously a 'willing addict'[69]; the person who identifies himself as a Christian pastor, but who also embraces his homosexuality;[70] or the person who cares about her own physical destruction. In this last case, the clash is between, not contingent practical identities, but a contingent practical identity and the necessary practical identity of a person: if you embrace your personhood, then you can't simultaneously care about something that will bring about the end of your personhood by bringing about your bodily death; and you can't identify with the termination of your personhood, because that (as any caring does) would require your embracing your personhood. Adopting a deeply incompatible pair of practical identities will put you in an intractable reasons-conflict: the health-conscious willing addict, for instance, will have reason both to take drugs and to refrain from taking drugs; the Christian homosexual both to have a sexual relationship with another man and not to; the self-destructive person both to kill herself and to save her life. An individual torn between mutually exclusive reasons like this will disintegrate, be torn apart by her own reasons that, as the expression goes, 'pull her in two opposite directions'. So, if our practical identities are expressions of our self-love, then any practical identity that (along with another) undermines the integrity of the self is not feasible. Thus, I think it is correct that these characters' practical identities are restricted in that the health-conscious person ought not to care about taking drugs, the Christian ought not to embrace his homosexuality, and a person ought not to identify with her own demise.

---

[68] Korsgaard 1996, p.103
[69] See Frankfurt 1971, pp.19-20.
[70] Thanks to Dr Elisa Galgut for this example.

3.2   Implications of Frankfurt's two restrictions

3.2.1   Person-particularity

Though the above characters' practical identities are restricted in the ways outlined, it is crucial to an accurate depiction of Frankfurt's position to note that they are *not* (as I have presented it) restricted as they are because taking drugs/having homosexual sex/killing oneself is *morally wrong*. Likewise, as we saw in section 3.1.1a, the unwilling addict's caring about sobriety is not entailed by sobriety's being *morally right*, but by his caring about his health. Rather, all these reasons to care or not care are a product of *who the particular person is*: someone who cares about her health, someone who embraces Christianity, someone who has the practical identity of a person. Frankfurtian restrictions on the practical identities available to a person are, therefore, *person-particular*: the 'rightness' or 'wrongness' of a practical identity is not given by a mind-independent force that imposes itself on one's will, as I have explained in section 2.9.2; rather, it is a function of that particular person's aggregate practical identity.[71] In Frankfurt's words:

> Whether something is to be an object of our love cannot be decisively evaluated either by any *a priori* method or through examination of just its inherent properties. It can be measured only against requirements that are imposed upon us by other things that we love.[72]

3.2.2   Immorality

Although morality is concerned with persons' conduct, it is not concerned with *particular* person's conduct; morality is impartial. What a person is and isn't allowed to do is not a matter of who she is: a king is not permitted to cut off his wife's head just because he's blue-blooded; a member of one race isn't disqualified from voting because he happens to be browner; a rich person is no more entitled to healthcare just because he's got more money in the bank; and so on. And, this impartiality of morality is quite at odds with the person-particularity of Frankfurt's account.

This is not to say, though, that on Frankfurt's account no person ever has a reason to act morally. The Frankfurtian account is not an immoral one, but a thoroughly *amoral* one. A person may well have a reason to act morally – however, that is *only if* she cares about morality, if she embraces the practical identity of a morally good person (or if she cares about something else – for instance, other people's opinions of her – which entails her caring about behaving morally). As Frankfurt writes:

> Even when volitional necessity arises in connection with actions which are required or forbidden by duty, it does not derive from the person's moral convictions as such but from the way in which he cares about certain things.[73]

---

[71] Even having a reason to not kill oneself is person-particular. See Korsgaard's interesting discussion: 1996, §4.4.1, pp.160-161.
[72] Frankfurt 2004, pp.47-48 (emphasis added)
[73] Frankfurt 1988, p.90

A person's having a reason to act morally is particular to, dependent on, who she is. So, if a person does not identify herself as morally good, then, on Frankfurt's account, she does not have a reason to be moral (*ceteris paribus*)[74] – even if she judges accurately what is right and what is wrong. Here is Frankfurt's most lucid statement (in my opinion) of his position, quoted at length:

> self-love is in itself neutral with respect both to moral and to nonmoral values. It has no essential evaluational vector. A person loves himself insofar as he loves anything at all. The value of what he loves is irrelevant to the question of whether he is wholehearted in loving it.
> This leaves open the possibility that someone may wholeheartedly love what is evaluationally nondescript, or what is bad, or what is evil. Attempts are sometimes made to demonstrate that an unconflicted and unequivocal love of things like that is not possible. Many philosophers and religious thinkers have hoped, and have purported to demonstrate, that a will must be inescapably in conflict with itself unless it is effectively guided and constrained by the requirements of morality. […]
> In fact, however, their arguments are not convincing. It seems to me, indeed, that the project in support of which they argue is hopelessly unpromising. Being wholehearted is quite compatible not only with being morally somewhat imperfect, but even with being dreadfully and irredeemably wicked.[75]

So, as Susan Wolf has correctly observed, it is a consequence of Frankfurtian person-particularity that

> (i)f our make-up and circumstances are such that we will be more rewarded by caring about helping people rather than hurting them, then we should cultivate our sympathies. If, however, we would be more fulfilled by taking up the call of sadism, nothing in Frankfurt's remarks seems to discourage it.[76]

However, not only does 'nothing in Frankfurt's remarks discourage' a person's taking up the 'call of sadism', the Frankfurtian account even leaves room for the possibility that such an immoral practical identity may in fact be *entailed* by the logic of caring. Say a certain person thrives on the negative attention she receives from upsetting others, and she comes to identify herself in this way. She might elicit this negative attention by telling inappropriate jokes or by using explicit language – but to *really* get people's backs up, she should behave *sadistically*. That is, she would have a *reason* to be despicably immoral. I find this prospect utterly unacceptable, as many philosophers have, for reasons briefly described below.

3.2.3   Insularity

Frankfurt concedes that "(i)t is possible, I am sorry to reveal, that immoral lives may be good to live."[77] Good for whom? Good for the particular person, of course. Frankfurt clearly espouses a kind of relativism. However, despite the literary catalogue of problems with relativism, and despite my repulsion by the idea that is may be 'good' for a person to be a sadist, I can see the sense in Frankfurt's position, for reasons we have already canvassed the reason in section 2.9.2:

---

[74] This clause is important: there may be other ways by which a person, who does not identify herself as morally good, nevertheless has a reason to act morally; for instance, if she cares about what other people think about her, and does not want to make them think badly of her by acting badly.
[75] Frankfurt 2004, p.98
[76] Wolf 2002, p.229
[77] Frankfurt 2002, p.248

the preservation of our autonomy. If a person's reasons are determined by value – moral value – then her free will is compromised, and her acting morally will not be an expression of her real self. So, what a person cares about cannot be restricted by impartial, impersonal morality. Thus, the allowance for immoral practical identities is the price paid for securing our freedom in expression.

However, as I will argue in the final chapter, I think this price buys us the wrong kind of freedom: an unintelligible freedom. I prefer to describe Frankfurt's position as insular, rather than relativistic: island cultures are generally described as 'insular', being cut off from the rest of the world (the word clearly shares the etymology of 'insulate'), or a person may be so described if he is 'caught up in his own world' or fails to take notice of other people; relativism, on the other had, conveys the idea of relating things to other things – and this is exactly where I think Frankfurt goes wrong. As I will go on to argue, the Frankfurtian assertion that a person can love what she judges to be valueless, and therefore that she can care about something she knows to be immoral, serves to render the person unintelligible by severing her connection to the world, to others, and even to herself. As such, there cannot be only person-particular restrictions on practical identities, and there must also be a moral constraint, that serves to relate persons to other persons.[78] This is the force of the Korsgaardian restriction on practical identities, to which we turn next.

---

[78] Let me nip one possible objection in the bud, before we proceed: one may be tempted to think that Frankfurt's account is *prima facie* correct, because there are plenty of immoral people out there – yet they are persons nonetheless; so, any account that *disallows* for the possibility of immoral people is clearly mistaken. However, let us take care to distinguish between immoral *descriptive* identities and immoral *practical* identities. My contention is that, although many of us may suffer immoral descriptive identities, no person can embrace an immoral practical identity.

4   Korsgaardian restrictions
on practical identities

It is time to examine more closely the philosopher I have pitted against Frankfurt. To my mind, Korsgaard presents a much more thorough, intricate – at times, quite difficult – argument than Frankfurt does. And it is this greater attention to detail that I think leads Korsgaard to the correct conclusion, the lack of which leads Frankfurt astray. In other words, I regard the Korsgaardian account as the full, correct working out of the Frankfurtian account to its proper conclusion. In this respect I apparently share Korsgaard's opinion on Frankfurt; in her words:

> Frankfurt, if I understand him correctly, thinks that it follows from these views that the normativity of morality for any given agent is contingent on whether that agent cares about morality […]. And I don't agree with that. That is to say, I don't think that it *follows*, and I also don't think that it is true.[79]

In the previous chapter, we examined the two Frankfurtian restrictions on practical identities, and how their person-particularity may give an individual reason to act immorally. Korsgaard, I assume, would accept these two restrictions too,[80] yet she takes a third to follow as well: the moral restriction on practical identities. And, this is not a person-particular restriction: irrespective of any particular person's other practical identities, she cannot feasibly adopt an immoral practical identity. Korsgaard's main argument for this claim is of Kantian origin, developed on the basis of Immanuel Kant's argument for his 'Formula of Humanity',[81] described in section 4.1. The bulk of this chapter will be concerned with defending two particular premises of that basic Kantian argument, which will be the subjects of section 4.2 and 4.3. As we will see, though, we are *still* left with an insular account, and further argument is required to establish a moral restriction on practical identities, which will take us into the next chapter, and the next.

4.1   The basic Kantian argument

As Frankfurt's seminal paper inquired into free will,[82] the Kantian argument also begins with a consideration of free will. According to Kant, the free will is "a rational causality which is effective without being determined by any alien cause",[83] where – like Frankfurt[84] – Kant allows for the

---

[79] Korsgaard 2006, pp.55-56
[80] And, I assume, Korsgaard's justifications of the former two restrictions would be the same as those I have presented in the previous chapter, so they will not be rehearsed here.
[81] See Kant 1785/1959, pp.45-47, as referenced by Korsgaard 1996, p.122.
[82] See Frankfurt 1971
[83] Korsgaard 1996, p.97, referencing Kant 1785/1959, p.446

desires and inclinations to which a person is subject to count as 'alien causes'.[85]   Since this present project has not been directly concerned with the matter of free will, let me put the point, rather, as follows: if a person's actions are expressive of her real self, then she must regard herself – her real self – as the cause of those actions.   Recall that a person's real self is constituted by her identifying with some (possible) way of being and dissociating herself from others.   And, "(s)ince the concept of a causality entails that of laws",[86] there must be some law or principle according to which the person so identifies herself, thereby constituting the self that is expressed in action.   So, in constituting herself, a person is a law unto herself.

Now, since the will is not yet subscribed to any principle of choice, there are no constraints as to which principle(s) it may adopt – except that the principle must have the form of a law.   But this is just Kant's categorical imperative: act only on a maxim that can be willed as a law.   This is why, according to Kant, the categorical imperative is the law of the free will.[87]   Now, since laws are universal, this means that all free wills are subject to the categorical imperative.   And, if every person's actions are governed by the same law, this means that all persons "act only on maxims that all rational beings could agree to act on together in a workable cooperative system" – and this is just Kant's 'moral law'.[88]   So, the law according to which a person may identify herself one way rather than the other turns out to be the moral law – that is to say, the practical identities available to a person are restricted by morality, *contra* Frankfurt.

4.1.1   Dubious moves in the basic Kantian argument

If you do not find this argument convincing, you are not alone.   Although I will ultimately defend (a version of) it, as Korsgaard does, there are two conspicuously dubious moves in this basic Kantian argument.   The first is the presumption that a person can only identify herself according to some law.   G.A. Cohen, for instance, denies this, contending that:

> If, as Korsgaard says, 'the necessity of acting in the light of reflection makes us authorities over ourselves', then we exercise that authority not only in making laws but also in issuing singular edicts that mean as much to us as general principles do.[89]

However, in section 4.2, I will defend Kant in this regard: a person can only act in accordance with a law.   Here, I will follow Korsgaard, although my discussion will take us further into the Kantian theory of causality than is strictly necessary in order to make this point.   However, it will prove indispensable for forthcoming arguments.

---

[84] See Frankfurt 1971
[85] Korsgaard 1996, p.97
[86] Kant 1785/1959, p.65, as cited by Korsgaard 1996, p.97
[87] Korsgaard 1996, p.98, referencing Kant 1785/1959, pp.64-67; and 1788/1956, pp.28-29
[88] Korsgaard 1996, p.99
[89] Cohen 1996, p.176, citing Korsgaard 1996, p.165

The second questionable element of the basic Kantian argument is the move from the categorical imperative to the moral law. We should be suspicious of this move because the categorical imperative – act only on a maxim that can be willed as a law – is, as I will demonstrate, really just an empty formalism: it is just the law that one act only according to a law. The moral law, on the other hand, is richly contentful – it tells us exactly what we can and can't do. Therefore, we should be suspicious of Kant's apparently treating the two as the same thing. Korsgaard's response to this will depart from strictly Kantian theory, and as we will see, bring her somewhat closer to Frankfurt.

Since, as I have described it, Korsgaard's views are the proper extension of what are otherwise Frankfurt's – and, therefore, of all the material covered so far – much of the argument (besides my digression into the Kantian theory of causality) will consist in the tying together of loose ends, and there will be much back-referencing to previous chapters. I hope my various references to both Korsgaard and Frankfurt make it clear that I take Frankfurt to be as much committed to the following arguments as Korsgaard.

4.2   Causality and laws

If a person acts – that is, if her behaviour is expressive of her real self – then she – her real self – caused that behaviour; she is not merely the "location of a causally effective desire",[90] as the unwilling addict is, for instance, when it comes to his drug-taking. As Korsgaard notes,

> To will is not just to be a cause, or even to allow an impulse in me to operate as a cause, but, so to speak, to consciously pick up the reins, and make myself the cause of what I do. And if I am to constitute myself as the cause of an action, then I must be able to distinguish between my causing the action and some desire or impulse that is 'in me' causing my body to act.[91]

So, the notion of causation is as central as the notion of the real self when it comes to a person's action (as opposed to mere behaviour) or self-expression. And, as we will see, they both come together: as I will explain, the synchronic and the diachronic unity of the self are bound up with each other.

My discussion in this section will build on that of volitional necessity in section 2.7. As we noted there, volitional necessity has a certain unifying, integrating function. It is the force that "connects and binds us to ourselves",[92] imparting "coherence to our volitional lives."[93] In short, it establishes our *synchronic unity*, our identity in a particular action. For instance, imagine two persons engage in a descriptively identical behaviour: they both eat a meat-free meal, say, a cheese and

---

[90] Korsgaard 1996, p.228
[91] Korsgaard 1996, pp.227-228
[92] Frankfurt 2004, p.17
[93] Frankfurt 1999, p.162

tomato sandwich from the cafeteria. Now, one of these persons is a vegetarian, and the other not. The vegetarian is to be distinguished from the non-vegetarian in the performance of this act – even though they *both* eat a meat-free meal – because the vegetarian stands in a special relationship to the action, which her counterpart does not: the vegetarian's eating a cheese and tomato sandwich is an *expression* of herself, and the omnivore's not. The vegetarian has acted as she is volitionally constrained to and thereby exercises her autonomy,[94] and this is not the case for the non-vegetarian. This is because the vegetarian has constituted herself, by way of her practical identity, in a way that the non-vegetarian has not.

How is it that the vegetarian's action is an expression of herself, and the non-vegetarian's not? This means that the vegetarian construes her real self as the *cause* of her action, and the non-vegetarian does not: she is merely the location of the causally effective desire to eat (anything). So, to get to grips with this, we should inquire further into causation. To this end, I will present Kant's theory of causality, by way of contrast with Hume's. The subsequent discussion will be more detailed than is strictly necessary for present purposes, but the material will prove useful to subsequent arguments I wish to present, so I permit myself the digression. After this slight excursion, I will return to the present topic in section 4.2.3, and show why the synchronic unity of the self is inseparable from its diachronic unity.

4.2.1 Hume on causality

Hume and Kant held very different views of causality – perhaps surprisingly so, given that they also concurred to a large degree. (This strikes me as interestingly analogous to the Frankfurt/ Korsgaard debate.) Both Hume and Kant agreed that, since causality involves the necessary connection of cause and effect,[95] and this necessity cannot be given by experience alone,[96] causality does not exist 'out there', in the world, but is/is the result of, rather, "a fundamental feature of human nature […] which is well known by its effects."[97] On both philosophers' accounts, therefore, causality is something contributed by the human mind. However, the manner in which the human mind furnishes itself with causality is significantly different on Hume's and Kant's respective accounts.

Hume began by establishing the logical independence of the antecedent of a causal conditional from its consequent. Take as an example of such a causal conditional the familiar: If one billiard-ball hits a second billiard-ball, the second will move. As Hume writes:

---

[94] Note that the vegetarian would still be distinct from the non-vegetarian if they both *ate meat* (therefore also behaving descriptively identically): here the vegetarian would be not expressing herself, but repressing herself, hurting herself, and the omnivore would not.
[95] See, for instance, Kant 1787, B5, where he explicitly references Hume.
[96] Hume 1748, §4; Kant 1787, B3
[97] Hume 1748/2010, §5, p.20

> The mind can't possibly find the effect in the supposed cause, however carefully we examine it, for the effect is totally different from the cause and therefore can never be discovered in it. Motion in the second billiard ball is a distinct event from motion in the first, and nothing in the first ball's motion even hints at motion in the second.[98]

Hume takes this to imply that causality is not known *a priori*, and therefore can only be known on the basis of experience.[99] However, as is well known, Hume also noted that causality is never directly given in experience; we are never in touch with any 'causal nexus', we only ever observe one kind of event regularly following, or regularly being preceded by, another.[100] This means that the way in which the human mind furnishes itself with causality must be by way of *inference* on the basis of experience: events of type A were regularly followed by events of type B in the past, therefore events of type A will continue to regularly be followed by events of type B in the future.

Of course, this inference depends on the further premise that the future will resemble the past[101] – and this is a premise for which we have little-to-no justification, since we have not yet experienced the future. Besides which, no matter how well we may be able to substantiate this latter premise, it does not bring with it the necessity of causation – nor could any premise ultimately grounded in experience. Our human propensity to infer causation, the necessary connection of cause and effect, is therefore not a matter of rationality, according to Hume, but is rather merely a result of 'custom' or 'habit'.[102] This, Hume takes to "(announce) a discovery about the weakness and narrow limits of human reason and capacity."[103]

4.2.2   Kant on causality

Kant, in a sense, turns the Humean theory of causation on its head. Where Hume reasoned that, because causality is not to be discovered 'out there' in the world, it must be our human minds that contribute causality as an inferential response to past experience, Kant reasoned that, because causality is not to be discovered 'out there' in the world, it must be our human minds that contribute causality as a *precondition* of experience, which "is possible only by means of the representation of a necessary connection of perceptions".[104] Kant thus

> rescues the *a priori* origin of the pure concepts of the understanding […] in such a way that their use is limited only to experience, because their possibility has its ground merely in the relation of the understanding to experience, however, not in such a way that they are derived from experience, but that experience is derived from them, a completely reversed kind of connection which never occurred to Hume[105]

---

[98] Hume 1748/2010, §4, p.13
[99] Hume 1748, §4
[100] Hume 1748, §4
[101] Hume 1748, §4
[102] Hume 1748, §5
De Pierris and Friedman (2008) use the term 'feeling of determination' (see §2).
[103] Hume 1748/2010, §7, p.37
[104] Kant 1787, B218
[105] Kant 1783/2003, § 30, as cited by De Pierris & Friedman 2008, §1 (emphasis added for consitency's sake)

Causality for Kant is a pure concept of the understanding because it is (transcendentally) necessary to unify the manifold of experience, and thus to produce life as we know it: each of us with an integrated point of view on an integrated world.

In his 'Second Analogy' of the Critique of Pure Reason[106] Kant establishes that "(a)ll alterations take place in conformity with the law of the connection of cause and effect."[107] He asks us to consider the visual perception of a house. This could

> begin with the apprehension of the roof and end with the basement, or could begin from below and end above; and [we] could similarly apprehend the manifold of the empirical intuition either from right to left or from left to right.[108]

Contrast this with the visual experience of a ship moving downstream: in this case

> (m)y perception of its lower position follows upon the perception of its position higher up in the stream, and it is impossible that in the apprehension of this appearance the ship should first be perceived lower down in the stream and afterwards higher up.[109]

Kant's point is that there is no determinate way in which the intuition of a house's basement and that of its roof must be ordered because the relation between these two objects is not causal – a basement doesn't 'bring about' a roof, nor *vice versa*; on the other hand, there is only one way in which my intuition of a ship upstream and that of the ship downstream can be ordered when this is a causal relationship: the boat's being upstream *must* precede its being downstream. When a perception is one of causation, there is some necessity as to the order of my intuition of the antecedent and of the consequent, and this is the necessity supplied by my intuitions' being brought under a concept of the pure understanding: the law of cause and effect. This law of cause and effect brings with it the expectation that I should see similar antecedents followed by similar consequences in the future: when next I see a ship upstream, I should expect to see it move downstream. (Whereas my future perception of a basement would in no way necessitate the subsequent perception of a roof.) This Kantian extension to future cases is no Humean inference upon the unsubstantiated premise that the future will resemble the past; rather, it is the idea that *what it is* to perceive an event as causal is to see it as falling under a law which extends to future cases (as well as covering all past ones). My projection that future As will also be followed by Bs is constitutive of my seeing today's A as the cause of B at all.

Because Hume regards causality, not as a precondition of experience (as Kant does), but as an inference based on past experience – a bad inference – he draws his skeptical conclusion that our human mind's contributing causality indicates "the weakness and narrow limits of human

---

[106] Kant 1781/1787, A189-210/B234-257
[107] Kant 1787, B234
Or, as he first worded it: "Everything that happens, that is, begins to be, presupposes something upon which it follows according to a rule." (1781, A189)
[108] Kant 1781/1787, A192-193/B238
[109] Kant 1781/1787, A192/B237

reason and capacity".[110]   For Kant, on the other hand, our pure concept of causality is not a weakness at all – nor, though, is it a strength. The categories of bad/good or weak/strong aren't applicable to Kant's constitutive account, because we are bound by our human condition, the concepts of our pure understanding being what they are, to take it that if today billiard ball A caused billiard ball B to roll, then tomorrow, if circumstances are relevantly similar, another billiard ball A would do the same, and so on for every other day in the future; this is just part of what it is to have a human point of view on the world.

4.2.2a Law and objectification

Before I return to the central discussion of synchronic and diachronic unity, I should make two quick points related to the Kantian theory of causality. The first is about the objectifying effect of laws. As noted, for both Hume and Kant, causality is not to be discovered 'out there', in the world, but is, rather, a contribution of the human mind. According to Hume, it is a matter of my inferring on the basis of past experience; as such, causality retains some kind of connection to the outside world: the inference is a contribution of the human mind, but the perception of regularity that provides the basis for this inference comes from 'out there'. On Kant's account, causality is a concept of the pure understanding – not a response to some experiential 'evidence' but the precondition for that kind of experience in the first place. So, we might wonder how, on the Kantian account, our perception of, say, a ship moving downstream is a perception of a ship at all – that is, an object 'out there', with independent existence – and not just an observation about the ordering of my subjective intuitions. But this, according to Kant, is exactly what this concept of the pure understanding – or, the law of cause and effect – ensures: because it is a law, it has the effect of *objectification*. Kant's phrasing of this point in the *Prolegomena* is useful:

> Empirical judgements, in so far as they have objective validity, are judgements of experience; they, however, in so far as they are only subjectively valid, I call mere judgements of perception. […] All of our judgements are at first mere judgements of perception: they are valid merely for us, i.e., for our subject, and only afterwards do we give them a new relation, namely to an object, and we intend that [the judgement] is supposed to be also valid for us at all times and precisely so for everyone else; for, if a judgement agrees with an object, then all judgements about the same object must also agree among one another, and thus the objective validity of the judgement of experience signifies nothing else but its necessary universal validity.[111]

'Objectivity', here, is a rich term: the law of cause and effect 'objectifies' in the sense that it makes (say) the ship we perceive an *object*, a real thing, and not just an intuition; and it also 'objectifies' in the sense that this perception is not just a subjective intuition, but is an experience of something *objective* – experiences of which *other* (human) subjects could have too. So, though I am bound by my human condition to perceive the world by way of the pure concepts of

---

[110] Hume 1748/2010, §7, p.37
[111] Kant 1783/2003, §18 (emphasis removed), as cited by De Pierris & Friedman 2008, §1

understanding (*viz.* the law of cause and effect), my experience is nevertheless of a world which other subjects might perceive in the same way I do.

We will return to this point in the following chapter, in section 5.3.

4.2.2b The law of cause and effect versus *particular causal laws*

Above, I appealed to Kant's argument in the 'Second Analogy'. The conclusion of that section of the *Critique* is that: "Everything that happens, that is, begins to be, presupposes something upon which it follows according to a rule."[112] That is, the Second Analogy is the *rule* that 'everything that happens presupposes something upon which it follows according to a *rule*'. Of course, this latter rule can't be the former rule – the Second Analogy itself – otherwise its definition would be circular. The Second Analogy, as I have frequently referred to it in previous sections, is also known as 'the law of cause and effect'. This latter term conjures up ideas of the velocity at which a billiard ball will travel when it is hit by another billiard ball, about how a ship will move downstream or even about how a lit stove will heat a room.[113] These, however, are what Kant calls "particular laws",[114] which are not identical with the law of cause and effect, but are, rather, instantiations of it.[115] The law of cause and effect is just the *law* that every effect follows upon a cause according to some *law*. This second use of 'law' therefore refers to some particular causal law. The law of cause and effect itself, then, is somewhat empty: all it says is that there are such things as particular causal laws. Particular causal laws, on the other hand, are contentful: they tell us not only *that* some event is caused by another event, but *which* events are caused by which events.

Now, though the law of cause and effect is *a priori*, a pure concept of the understanding, particular causal laws are given by experience:

> Special laws, as concerning those appearances which are empirically determined [such as particular causal laws], cannot in their specific character be derived from the categories [such as the law of cause and effect], although they are one and all subject to them. To obtain any knowledge whatsoever of these special laws, we must resort to experience; but it is the *a priori* laws [such as the law of cause and effect] that alone can instruct us in regard to experience in general, and as to what it is that can be known as an object of experience.[116]

The law of cause and effect, then, which is not much more than a formalism, is instantiated as particular causal laws when combined with experience. That is to say, experience provides the

---

[112] Kant 1781, A189
[113] This is Kant's (1781/1787, A202/B248) own example.
[114] Kant 1787, B165
To be precise, I will call them 'particular *causal* laws'.
[115] De Pierris & Friedman 2008, §2
[116] Kant 1787, B165

content, and the law of cause and effect determines its form, setting the boundaries of the content a particular causal law can take. I return to this point in section 4.3.1.

4.2.3 Self-expression: being a law unto oneself

Earlier, we left off with the question of what it is for the vegetarian to *cause* her action of eating the cheese and tomato sandwich, and thereby to express herself, where this is not the case with the non-vegetarian. The driving idea here is that "normativity holds our inner world together in something like the way causality holds the outer world together."[117] As Frankfurt puts it,

> The moments in the life of a person who cares about something […] are not merely linked inherently by formal relations of sequentiality.[118] The person necessarily binds them together, and in the nature of the case also construes them as being bound together, in richer ways.[119]

So, we need a way for a person to distinguish between her causing her action – binding it to herself – and merely temporally linked states. And, the above discussion of causality provides just that. As both Hume and Kant note, causality is connected to regularity, the idea (somewhat simplistically) that the same thing would happen in future, relevantly similar situations, because "we could never identify the element of necessitation and therefore distinguish cases of causal connection from cases of mere temporal sequence without regularity."[120] So, this is how a person's synchronic unity goes hand-in-hand with her *diachronic unity*[121] – that is, her identity 'across time', on various occasions: the vegetarian, for instance, is who she is, and is distinguished from the non-vegetarian, on the occasion of her eating the cheese and tomato sandwich (i.e. her synchronic unity) because the vegetarian construes herself as the cause of her actions, and not merely the location of a causally effective desire; and, according to Kant, what it is for her to regard herself as causing her action is for her to take it as a matter of law: that similar antecedents (her self) will have similar consequents (opting for the vegetarian meal) in the future (i.e. her diachronic unity). This is why "(a) person who cared about something just for a single moment would be indistinguishable from someone who was being moved by impulse."[122] A person's expressing herself and acting as she is volitionally constrained to on any one occasion, therefore, cannot be separated from her taking herself to be committed to so acting in other relevantly similar situations – that is, as acting according to a law, and not a 'singular edict'.

Several points follow from this. 'Construing oneself as a cause' obviously involves taking a view of oneself in reflection – and how one reflectively conceives of oneself is just what we called in Chapter 2 one's practical identity. So, the law according to which a person acts is her *practical*

---

[117] Korsgaard 1996, fn.12, p.139
[118] *Cf.* Kant on the visual experience of the parts of a house.
[119] Frankfurt 1988, p.83
[120] Korsgaard 1996, p.227
[121] The terms 'synchronic' and 'diachronic unity' come from Korsgaard (2006, pp.60-61, referencing Frankfurt 2006).
[122] Frankfurt 1988, p.84

identity, and, as we said, a person constitutes herself by way of her practical identity. (And this brings us back to synchronic unity – so we have come full circle).[123] So, the law according to which a person acts is a law that she is unto herself, and she is therefore autonomous. When she acts as her practical identity gives her reason to, she freely expresses herself, and when she fails to so act, she does violence to herself – she *violates* the law that she is unto herself. Volitional necessity, therefore, is the force of law.

4.3   The categorical imperative and the moral law

A further point that follows from the above – that a person can act only according to a law, because this is how she construes herself as the cause of her action – is that this is basically Kant's categorical imperative: act only on a maxim that can be willed as a law. So, Kant's claim, that the categorical imperative is the law of the free (freely expressed) will, follows: any person who acts, if she acts at all, acts according to the categorical imperative. However, in light of the previous section, we can see that this is something of a truism or an empty formalism: it doesn't tell us *which* maxims we're allowed to will as laws, only that they have to have the form of laws[124] – which of course they do, if they can be willed at all. The moral law, on the other hand, is richly contentful: it tells us exactly what we can and can't do. How, then, do we move from the categorical imperative to the moral law, as Kant does in his argument, described above in section 4.1?

4.3.1   Bridging the gap: practical identities

We encountered a similar situation above, in section 4.2.2b, with regards to the relation between the law of cause and effect and particular causal laws, where the former is merely a law about laws. There, the gap was mediated by experience: experience provides the content of the particular causal laws, and the law of cause and effect determines their form. And, when it comes to the categorical imperative and the moral law, the gap is bridged by none other than our practical identities.[125] The following risks putting too much strain on the analogy, but where we move from the law of cause and effect to particular causal laws by looking towards the world, we move from the categorical imperative to the moral law by looking towards our*selves*.

Kant, though, had nothing to say about 'practical identities'; this is a modern notion. In fact, Kant doesn't even draw the distinction between the categorical imperative and the moral law, as I (following Korsgaard) have done. Why not? We have, at the end of section 4.2, established the

---

[123] It is hardly surprising then, given how Hume's theory of causality differs from Kant's, that Hume (1739) argued that there is no self but only a "bundle" of perceptions to be met with in experience.
[124] Korsgaard 1996, p.99
[125] Korsgaard 1996, pp.99-101

categorical imperative, which amounts to the claim that a person construes herself as acting only when she does so according to a law – that is, only when she is volitionally constrained to act in the same way in other relevantly similar situations. So, another way to articulate the categorical imperative would be that a maxim must be 'universalizable' – that's just what it means to be a *law* and not a singular edict: a law holds universally. From this point, the basic Kantian argument proceeds as follows: since the laws according to which persons identify themselves are *universal*, this means that *all* persons are subject to the same law(s), which means that all persons "act only on maxims that all [persons] could agree to act on together in a workable cooperative system" – and this is just Kant's moral law.[126] So, according to Kant, the moral law is entailed by the categorical imperative; there is no need to invoke practical identities.

4.3.2 Universality

However, while I (like Korsgaard) wish to defend the conclusion that there is a moral restriction on the ways a person can act, the Kantian route to it is problematic: it depends on a fallacious equivocation of 'universal'. Perhaps the most natural understanding of the word is: pertaining to everyone, ranging over everybody. This is the use of the term upon which Kant relies when he deduces that all persons are therefore subject to the same law(s) – however, this is *not* the kind of universalizability indicated by the categorical imperative. In the latter case, the law according to which a person identifies herself is 'universal' in the sense that it applies to her all the time, in any relevantly similar possible future situation – that is, to *her*, as an individual person. And this is a very different kind of universality to that involved in the moral law, which applies to *all possible persons*, all the time. As Korsgaard explains the problem,[127] the categorical-imperative sense of 'universal' has a much smaller *domain* – namely, the individual – than the sense of 'universal' needed to get us from the categorical imperative to the moral law, the domain of which extends to all possible persons. So, we cannot simply equate the categorical imperative with the moral law, as Kant does. In order to achieve the broad-domain universality of the moral law, which ranges across all possible persons, we must demonstrate that a person who acts according to the categorical imperative necessarily conceives of herself as a possible person.

It is Korsgaard's introduction of practical identities that allows us to do just that: establish that any person who acts in any way necessarily conceives of herself as a person. As we have already established, in section 2.8, any person with any contingent practical identity has the necessary practical identity of a person. So, any person who acts according to the categorical imperative (that is, who acts at all) necessarily falls within the domain of the moral law, which ranges over all

---

[126] Korsgaard 1996, p.99
[127] Korsgaard 1996, p.99

35

possible persons. However, we are not yet home and dry, and this doesn't get us all the way to a moral restriction on practical identities.

4.3.3 Residual insularity

We have established that every person falls within the domain of the moral law, because she necessarily conceives of herself as a person – but how does this help to establish a moral restriction on practical identities? How does another's having the practical identity of a person tell me what I can and what I can't do – particularly what I can and what I can't do *to her*? All we have established so far is that *my* having practical identity implies *my* embracement of *my* personhood, and that *your* having practical identity implies *your* embracement of *your* personhood. Our respective practical identities, rooted in our respective personhoods, engender our respective reasons. But, as of yet, we have no cause to think that your reasons, expressive of your personhood, are in any way curtailed or otherwise influence by my reasons, expressive of my personhood. As Korsgaard correctly observes, "there is still a deep element of relativism in the system",[128] or, as I have called it, insularity. What remains to be shown, then, if Frankfurtian person-particularity is to be refuted, is that *a person's regarding herself as a person consists in her regarding others as persons too*. This will be our task in the following two chapters.

---

[128] Korsgaard 1996, p.113

5 Reasons Publicity

At the end of Chapter 3, I tentatively criticized Frankfurt on the grounds that his solely person-particular restrictions on practical identities render the person insular, disconnected from the world and from others. In the last chapter, Chapter 3, we examined the Kantian basis of Korsgaard's argument for a moral restriction on practical identities, which would serve to build "the ideal of human relations that morality embodies"[129] into the very formation of persons' practical identities, so that they would not be insular. However, the basic Kantian argument did not quite achieve its objective, and we were left with a residually insular account, where I identify with my personhood and you with yours, yet a large gulf may nevertheless exist between us, with your reasons imposing no restrictions on mine, nor *vice versa*.

The final suggestion of the previous chapter, was that we establish that a person's necessary practical identity entails her conceiving of others as persons too. As we will see, this is exactly where the neo-Kantians' Analogical argument picks up, as described in section 5.1. However, the neo-Kantian amendment fails in just the same way as the basic Kantian argument does. So, in section 5.2, we turn to Korsgaard's own proposed solution, which relies on Wittgenstein's private language argument. However, Korsgaard's consequent view of the public nature of reasons entails, I contend, too demanding a view of morality, where I am obligated not only by my own reasons, but by yours too. So, in section 5.3, I lay the groundwork for my own proposed solution (to be detailed in the following chapter), in which it is not reasons *per se* – normative force or volitional necessity – that is shared amongst persons, as it is under Korsgaard's account, but rather a world of objective values.

5.1 The neo-Kantian solution

5.1.1 The Analogical argument

The way in which neo-Kantians[130] have traditionally tackled residual insularity is to argue that because *my* personhood entails that I am subject to certain reasons – my reasons – I must, on pain of inconsistency, admit that *your* personhood generates its own reasons – your reasons – and that, if I respect my reasons, I ought also to respect yours. This is just an argument from analogy: I am a person and I have reasons; you are like me in that you too are a person; therefore you also have reasons, which I must take into account when formulating my reasons.

---

[129] Korsgaard 2006, p.55
[130] Korsgaard (1996, p.133) cites Gewirth (1978) as one such philosopher.

That is to say, your reasons limit the sort of reasons I might possibly have, and (by the same analogy, in the other direction) my reasons limit the sort of reasons you might possibly have; which is to impose a moral restriction on practical identity.

5.1.2   The problem with the Analogical argument

I think there are several problems with the Analogical argument. For one thing, I think it takes the wrong view of personhood, in much the same way that Hume takes the wrong view of causation, as compared to Kant – but this must wait till the next chapter. Secondly, I think it runs the risk of suggesting, by extension of the analogy, that the reasons you – as a person, like me – have are the very same reasons that I have. And this is unacceptable, because then there is no way to distinguish between us as persons. As covered in sections 2.7 and 2.8, though our personhood is common between us, our various reasons are what distinguish us as individuals.

But, the biggest problem with the Analogical argument – and the one that Korsgaard picks up on – is that it doesn't actually overcome insularity. The Analogical argument is something of a *non sequitur*: it doesn't actually follow from my recognizing that you have reasons that I ought to *respect* those reasons or that your reasons present any sort of restriction on the reasons I myself may have (and the same for you with regards to me). As Korsgaard puts it:

> Consistency can force me to grant that your humanity is normative for you just as mine is normative for me. But it does not force me to share in your reasons, or make *your* humanity normative for me. It could still be true that I have my reasons and you have yours, and indeed that they leave us eternally at odds.[131]

My recognizing that you are subject to reasons may have no deeper implications than my recognizing that we both drive the same kind of car or that we share a birthday. Just as our ID numbers' sharing some digits brings us no closer together, my recognizing that you have your own reasons (whatever they might be) may still 'leave us eternally at odds', and each insular.

5.2   Korsgaard's Wittgensteinian solution[132]

5.2.1   Private reasons

Korsgaard's diagnosis of the failure of the Analogical argument is that it implicitly relies on a conception of reasons as inherently private. The Analogical argument then tries to move from a person's private reasons to 'public' ones – reasons that take other people's (private) reasons into

---

[131] Korsgaard 1996, p.134
[132] To be fair, section 5.2 covers only one of two solutions offered by Korsgaard: that presented in *The Sources of Normativity* (1996). Korsgaard has gone on to offer a different response more recently, which appears in 'Morality and the logic of caring' (2006). However, Korsgaard takes her two arguments to "come to the same thing in the end" (2006, fn.6, p.108).

account, amounting to a moral restriction – so that "the public character of reasons is as it were *created* by the reciprocal exchange of inherently private reasons".[133] So, on such an account a 'public' reason is really just a private reason to take others' (similarly private, to them) reasons into account.[134] And, the reason (private to the individual) that a person is given to have these 'public' reasons is that she neglects to respect others' reasons on pain of inconsistency with her own case. But there is something inconsistent about the argument itself: it trades in the currency of private reasons – which, being private, are person-particular – yet goes on to appeal to a reason all persons (apparently) must share: the reason to avoid inconsistency. To see this, just consider how Frankfurt would respond to the Analogical argument: unless a person *cares* about not being inconsistent, then this *isn't* a reason for her.

So, Korsgaard proposes that in order to really mount a defence of the moral restriction on practical identities, we must begin with a different conception of reasons altogether: as essentially *public*. In her words:

> the kind of argument we need here is not one that shows us that our private reasons somehow commit us to public ones, but one that acknowledges that our reasons were never more than incidentally private in the first place. To act on a reason is already, essentially, to act on a consideration whose normative force may be shared with others.[135]

"So", she continues, "what we need here is some help from Wittgenstein."[136]

5.2.2   The Private Language argument

To establish that reasons are not – as the Analogical argument mistakenly assumes – private, but rather, public, Korsgaard[137] employs Wittgenstein's 'Private Language argument'[138]. The Private Language argument is supposed to establish that there could be no such thing as a 'private language', where this means a language in which the meaning of its terms is given by reference to some private mental entity. The original argument targets sensations in particular, but Korsgaard extends it to reasons. Very briefly, the argument is this: meaning is normative – there is a right way and a wrong way to use a word; and if this is the case, then there must be some non-private way of fixing its meaning. Suppose I give a particular sensation of mine the name 'S'; on a private language account, then, the meaning of 'S' is given by my attending inwardly to the sensation. But how do I know if the sensation I attend to today when I use 'S' is the same sensation when tomorrow I use 'S'? That is, how do I know if I use the word 'S' correctly if I have

---

[133] Korsgaard 1996, pp.133-134
[134] To this end, Korsgaard (1996, pp.132-134) compares it to Hobbesian arguments, of the form: you should do X (submit to the will of the Sovereign, behave morally), because X is good for you.
[135] Korsgaard 1996, p.136
[136] Korsgaard 1996, p.136
[137] Korsgaard 1996, pp.136-138
[138] See Wittgenstein 1953, §§243ff.

no other way of identifying the sensation other than by means of its name, 'S'? As Wittgenstein puts it:

> A definition surely serves to establish the meaning of a [word].—Well that is done precisely by the concentration of my attention; for in this way I impress on myself the [connection] between the sign and the sensation.—But "I impress it on myself" can only mean: this process brings it about that I remember the [connection] *right* in the future. But in the present case I have no criterion of correctness. One would like to say: whatever is going to seem right to me is right. And that only means that here we can't talk about 'right'.[139]

So, if a word is to have any meaning at all, it must have some connection to something outside of the subject, some point of comparison by which it can be determined whether a word is used correctly or not.[140] Therefore, there can be no such thing as a private language. Korsgaard applies this same argument to the possibility of private reasons: reasons, like meaning, are normative; and if this is the case then reasons, like meaning, cannot be private. And this is apparently just what we need, because, if reasons are not private but *public*, then we can avoid the insular situation in which 'I have my reasons and you have yours, and we are left eternally at odds'.

5.2.3 Public reasons

If reasons are not private entities, what are they? Again, the parallel with meaning helps: according to Wittgenstein, the meaning of a word is not given privately, by some inner state; rather, "the meaning of a word is its use in the language",[141] where its 'use' is something like: what it gets people to do in response to it (like 'bringing me a slab'[142]). So, the meaning of a word is a function of how it relates people to each other. According to Korsgaard, a reason is, likewise, the function of a relation between people, and what it gets them to do.[143] And anything that gets its existence *between people* cannot be possessed by one individual alone, that is, privately. On Korsgaard's account, then, reasons are public in the sense that "the normative force of the reason I legislate should be public and shared."[144] And this is Korsgaard's 'publicity as shareability'[145] thesis: "If reasons are, as I have suggested, public and universal for all rational beings, then anyone's reasons are reasons for me";[146] "public reasons are reasons that have normative force for everybody."[147] And – more importantly, for our broader purposes – if it *can't*

---

[139] Wittgenstein 1953, §258
[140] I take this to be the force of Wittgenstein's (1953) remarks with regards to blood pressure in §270. As McGinn comments: "The point is not that the meaning of 'S' is exhausted by its connection with a rise in blood-pressure, but that its meaning derives from its use in the language-game, and not from its connection with anything that is identified by 'turning the attention inwards'." (1997, pp.138-139).
[141] Wittgenstein 1953, §43
[142] See Wittgenstein 1953, §19
[143] Korsgaard 1996, pp.137ff
[144] Korsgaard 2006, p.64
[145] Korsgaard 1996, p.135
[146] Korsgaard 2006, p.69
[147] Korsgaard 1996, p.133

be shared as a reason for me,[148] then it can't be a reason for you, and your practical identity is thereby restricted by consideration of another's reasons – that is, a moral restriction.

5.2.4   The problem with Korsgaard's public reasons

However, I find this to entail too demanding a view of morality. I certainly do not think that your reason, say, to climb Kilimanjaro requires me to help you do so[149] because, if it's a reason for you, it's shared as a reason for me. At most, I think your wanting to climb Kilimanjaro requires me not to get in your way. In fact, we sometimes find it quite offensive when other people try to take positive action on *our* reasons. For instance, imagine that I have a reason to lose weight (because I embrace a health-conscious practical identity, or something to the same effect), and you, acting on *my* reason, feel the need to point it out to me that I really shouldn't have had that second doughnut. As Korsgaard correctly notes: "we may resent it when strangers point out to us that we are not doing what we have reason to do, […], and we […] tell them that it is none of their business."[150] Not only might we tell such a person to 'mind her own business', we might well add that she should 'get her own life' – that is, her own practical identity, with her own reasons that shape her own life. In a similar vein, I do not think it is true, as Korsgaard does, that "(i)f I call out your name, I make you stop in your tracks […] because *I* am a law to *you*."[151] If my calling out your name gives you a reason to stop, then this is because of a law you are to *yourself*, having embraced the practical identity of the attentive and responsive person, or something like that.

What these complaints appeal against, I think, is the demolition of the distinction between *your reasons* and *my reasons* that Korsgaard's reasons publicity seems to imply. You have your reasons and I have my reasons, and although they might be 'shared' in the sense that you must respect my reasons (and *vice versa*) they are nevertheless *my* reasons that you should respect, not 'ours'. And, as we have noted now several times, if the sense is lost in which these are *my* reasons – if my reasons are everybody's reasons – then my sense of self altogether, my synchronic unity, is threatened. So, I think Korsgaard is mistaken to take the publicity of reasons to imply that reasons "have normative force for *everybody*"[152] or that "the reasons of others have something like the same standing with us as our own desires and impulses do"[153].[154] Rather, I think that *your reasons* are those that have normative force for *you* and *my reasons* are those that have normative force for *me*, and this is how the distinction between your and my reasons –

---

[148] I take it, because of those restrictions discussed in Chapter 3.
[149] This is the thrust of Korsgaard's (1993/1996) critique of Nagel (1970 & 1986). See p.298 in particular.
[150] Korsgaard 1993/1996, p.298
[151] Korsgaard 1996, p.140 (emphasis added)
[152] Korsgaard 1996, p.133 (emphasis added)
[153] Korsgaard 1996, p.140, referencing Ulrike Heuer.
[154] There is a sense, therefore, in which what Korsgaard is doing is 'extending the analogy' of the Analogical argument, in the manner discussed in section 5.1.2.

indeed, between *you* and *me* – is retained, as we have noted several times, in sections 2.7, 2.8, 4.2, and 5.1.2.

However, let me not overstate the extent of my disagreement or dissatisfaction with Korsgaard's account. I *do* agree that reasons are not private entities, but are public in the sense that they are *shareable*; however, I do not think this means that reasons are necessarily *shared*,[155] that my reasons are actually reasons for you too. I think, rather, that reasons are public and share*able* in the milder sense that my reasons could conceivably be reasons for you too, that another person must be able to find my reasons *intelligible* even if she does not herself experience their normative force.[156]

As I will go on to detail in the following section, I take reasons to be public in the sense that they must be 'co-cognizable'. This I hope to establish not by directly tackling the nature of reasons themselves – as Korsgaard has done – but by an examination of what it is to conceive of oneself as a person, to have the necessary practical identity of a person. That is to say, I propose we return to the original tack suggested at the end of the previous chapter: showing that one's very identity as a person entails a restriction on the practical identities available to her by other persons' reasons. First, though, some groundwork must be laid.

5.3   Objective value

What I do think is shared amongst persons is, not the normative force of their reasons *per se*, but the *world*, with which a person's caring serves to connect her. This lays the groundwork for my own argument for a moral restriction on practical identities, which I present in the following chapter.

5.3.1   Practical identity and objectification

Recall that, as noted in section 4.2.2a, according to Kant laws have an objectifying function. I perceive an event as causal when there is a lawfully necessary ordering of my intuitions: the boat's moving downstream is causal, because my intuition of it upstream *must* precede my intuition of it downstream. However, this is not a fact about the ordering of my intuitions; what I perceive is objective, something out there in the world. The laws that we are to ourselves – our practical identities – as I see it, have this same objectifying function.

---

[155] To be fair, Korsgaard does sometimes describe the requirement as one of being *shareable* rather than *shared*. See, for instance, 1996, p.135.
[156] Thus, I think Korsgaard gets it right when she says: "To act on a reason is already, essentially, to act on a consideration whose normative force *may be* shared with others." (1996, p.136; emphasis added).

If I regard my seeing billiard ball A as causing billiard ball B to roll as the result of solely who or what I am, as a fact about my intuitions, then the sense is lost in which I am in touch with the world, and have learnt something about the billiards balls – not my own intuitions. Likewise, if I do not regard my caring about an object, and thereby my making it important to me, as coming to stand in a special kind of relationship to something which has *objective value*, then, when I express that practical identity of mine, the sense is lost in which I am in touch with the world, acting in it and interacting with it. To this end, I think Frankfurt is spot-on (although he apparently fails to see the problem with it, or the solution) when he remarks that, on his account:

> once we begin asking how people *should* live, we are bound to find ourselves helplessly in a spin. The trouble is not that the question is too difficult. Asking the question tends to be disorienting, rather, because it is inescapably self-referential and leads us into an endless circle.[157]

And, as any sports coach will tell you, the way to keep your balance when you find yourself in a spin is to fix your eyes on the world. Thus, I think the correct lesson to draw from Wittgenstein's Private Language argument is, not that reasons are public in Korsgaard's sense, but that reasons, like meanings, must stand in some connection to the outer world – to objective value.

Korsgaard, however, is quite clear that she thinks "(v)alues are not discovered by intuition to be 'out there' in the world."[158] But values, on my understanding, are not 'out there', if this means, as I take it to: mind-independent. Causality, on both the Humean and Kantian accounts is not discovered 'out there' in the world; it is not mind-independent, but a contribution of the human mind. Nevertheless, when I perceive a ship moving downstream, this is not something I have noticed about myself, but about the *world*; because my perception is objectified by the relevant causal law. I suggest that just the same objectification of value takes place when a person makes something important to herself by taking up the relevant practical identity and becoming a law unto herself. That is to say, when a person cares about something, she necessarily takes that thing to be objectively valuable. (And, when a person positively rejects a way of being, she necessarily takes it to be objectively valueless.)

5.3.2 Value as objective yet mind-dependent

However, Frankfurt vehemently denies this necessary connection between reasons and objective value.[159] As I explained in section 2.9.2, I take it that this denial is a reaction to the threat of a loss of autonomy: if my practical identity is determined by objective value, then *I* am not the one who constitutes my real self by way of my practical identity. While I do endorse as a genuine worry any threat to autonomy, it is not a worry that is provoked by considerations of objective

---

[157] Frankfurt 2004, p.24
[158] Korsgaard 1996, p.112
[159] See, for instance, Frankfurt 2004, p.38.

value. To see it in this way, as Frankfurt apparently does, is, I think, a result of having succumbed to the almost irresistible temptation to affirm the consequent.

The lesson to be learned from the example Kant made of Hume is that Hume got things the wrong way round: he thought we extrapolate to future case because we infer causality; but, according to Kant, we see an event as causal because we take it to be the instantiation of a law which extends to future cases. My Kantian claim, with regards to value, is this: if you count it as a reason, then you must take it to have objective value. Now, this is not that same as: if you deem it to have objective value, then it must be a reason for you. And this latter, misrepresented claim would pose a threat to our autonomy – but that is not the claim I am making. Of course, a parent doesn't love her child *because* she deems him valuable. (There is something quite perverse in the idea of it.) However, I am in agreement with Frankfurt insofar that if she loves her child then she must take him to be valuable. I diverge from Frankfurt, though, in that I think this implies, not purely person-particular importance, but objective value.

To take a causal interaction to be objective is not to say that it occurs mind-independently, because causality is contributed by the human mind; it is, though, to say that other people, with minds like mine, might also perceive it, and that it is not purely a fact about *my* intuitions' necessary ordering. Similarly, to say that value is objective, in the sense I am advocating, is not to say that it is mind-independently 'out there', but that other persons might come to stand in the same caring relationship to that value as I autonomously choose to. A superficial look at the term 'valuable' reveals just that: *able to be valued* – by me, of course, since I do indeed value it, and by others, if they so constitute themselves. And whether we do or don't respond with care, embracing it and calling it our own, is where our autonomy resides. So, value is mind-dependent, but this need not imply person-particularity, as Frankfurt apparently takes it to. On my understanding, then, value is mind-dependent but nevertheless objective, not in the sense that it is objectively real or mind-independent, but in the sense that it is not dependent on *me*, or *you*, or any other particular person or perspective; value is, in a word, *intersubjective*. Precisely this view emerges in Korsgaard's later work:

> I agree with Frankfurt that love is not, or not necessarily, a response to value [...]. Yet I think that in loving something you do accord universal or public value to its object. And I think that someone who loves something with a certain kind of value is committed to that kind of value in general.[160]

As it stands, this does not constitute an argument for a moral restriction on practical identities. It does, though, lay the groundwork for my argument to follow.

---

[160] Korsgaard 2006, pp.75-76

6   Shareability as intelligibility

As part of her argument that reasons are public, Korsgaard presents the following scenario:

> A student comes to your office door and says: 'I need to talk to you. Are you free now?' and you say 'No, I've got to finish this letter right now, and then I've got to go home. Could you possibly come around tomorrow, say about three?' And your student says 'Yes, that will be fine. I'll see you tomorrow at three then.'
> What is happening here? On my view, the two of you are *reasoning together*, to arrive at a decision, a single shared decision, about what to do. And I take that to be the natural view. But if […] reasons cannot be shared, then that is not what is happening. Instead, each of you backs into the privacy of his practical consciousness, reviews his own reasons, comes up with a decision, and then re-emerges to announce the result to each other. And the process stops when the results happen to coincide[161]

Although I have my doubts, expressed in the previous chapter, about Korsgaard's view on the publicity of reasons, I think her description of what is happening here is exactly right: the student and the lecturer are *reasoning together*. And this strikes me as the same notion, now well established, though in a discrete area of philosophy, by (predominantly) Jane Heal: namely, *co-cognition*.[162]

Co-cognition, discussed in section 6.1.2, is a unique version of Simulation theory, which is a theory of interpersonal understanding. I contrast Co-cognition with what I call the empirical version of Simulation theory, discussed in section 6.1.1. Interestingly, empirical Simulationism has basically the same form as the failed Analogical argument. So, as Co-cognition improves upon empirical Simulationism, my hope is that a Co-cognitive account of the shareability of reasons, as put forward in section 6.2, will likewise improve on the Analogical argument, and successfully refute Frankfurtian person-particularity. Unfortunately, it will not amount to a fully worked out theory, but I hope it comprises the beginnings of a good one.

6.1   Two versions of Simulation theory

Heal's theory of C-cognition is broadly Simulationist, though she took up the term 'Co-cognition' in order to avoid confusion of her claims and arguments with others going under the heading of Simulation theory.[163] Heal's Co-cognition, and the other Simulation theories, are related in that they stand opposed to what is know as Theory theory. Theory theory and Simulation theory are rival hypotheses as to the underpinnings of 'folk psychology'; that is, that faculty possessed by all normal adult humans which enables us to navigate the social world, by way of describing,

---

[161] Korsgaard 1996, p.141 (emphasis added)
[162] 'Co-': together; 'cognition': thought; hence 'co-cognition': thinking, or reasoning, together.
[163] Heal 1998b/2003, p.92

predicting and explaining the behaviour of others (and of ourselves) by reference to mental states such as belief, desire, *etc*. We should note just how remarkably good we are at folk psychologizing; only very, very seldom is it the case that we are totally taken off guard by, or cannot make any sense of, another person's behaviour. Indeed, in such cases when we find ourselves at a folk psychological loss, it is usually an indication of the other's pathology rather than of our folk psychological incompetence. Folk psychologizing, therefore, is central to our lives.[164]

Theory theory gets its somewhat inelegant name from the fact that it is the *theory* that our folk psychologizing is subserved by appeal to some *theory* as to which mental states cause which behaviours. Since this is precisely what our understanding of everything else in the world (the weather, disease, gravity, the stock market, *etc*) consists in – namely, coming up with a theory that subsumes our observations – this is some *prima facie* support for the Theory theory. However, as the Simulationist points out, the sort of thing we're concerned with when we're engaged in folk psychology – namely, other *persons* – is the same sort of thing *we* are: that is, of course, persons. When it comes to folk psychology, the subject and the object of inquiry are alike in a special way, in a way not shared by, say, persons and the weather. I will call this the 'subject-object similarity' insight. And *this* seems to support the Simulationist hypothesis (to be explained in detail below) that in fact our understanding of other persons is quite different in kind to that of the weather, or anything else for that matter: it is an *interpersonal* understanding.[165]

There are, though, two ways of cashing out the Simulationist subject-object similarity insight. The first takes a distinctly empirical approach, and will be the subject of section 6.1.1.

6.1.1   Empirical Simulationism

Martin Davies and Tony Stone introduce Simulation theory by way of the following analogy. They ask: "How could someone predict the change in pressure of the gas in a cylinder when its temperature is raised?"[166] And they describe two possible answers:

> One possibility would be to use an empirical [generalization] about the way in which the pressure of a volume of gas increases as its temperature increases. In this case, the predictor would be drawing on a body of information about gases, in line with a (T)heory theory account. […]

---

[164] Korsgaard seems *prima facie* to hint towards the same kind of approach with her comment: "I take ['publicity as shareability'] to be equivalent to another thesis, namely, that what both enables us and forces us to share our reasons is, in a deep sense, our *social nature*." (1996, p.135) However, Korsgaard's actual working out of this approach is very different to mine: where she stresses our 'animal nature' (p.145ff), I take our 'sociability' to be a specially personal feature.

[165] Note that my intention in this section is *not* to provide an argument for the truth of Simulation theory, as opposed to Theory theory. (Although, if the anti-Frankfurtian arguments of this project work, Simulation theory may gain some credit by association.) I introduce Theory theory only because, as Heal notes, "Simulationism is best understood by contrast with […] the '(T)heory theory'." (1998a/2003, p.28).

[166] Davies & Stone 1998, p.55

> There is, of course, an alternative to [this] theory-based [strategy] for arriving at a prediction about the pressure of the gas in a cylinder, A, after its temperature is raised. We could take another cylinder of similar gas, B, heat it to the temperature in question, and measure the pressure. Provided that the cylinder B really is relevantly similar to cylinder A, this method is liable to yield an accurate prediction.[167]

The former method is an analogue of a Theory theory prediction, and the latter of Simulation theory. In the latter case, a prediction is made of cylinder A without appeal to any theory of gases, by way of a *simulation* of cylinder A on cylinder B. In the case of interpersonal understanding, the Simulationists' distinctive idea, therefore, is that, because of subject-object similarity, we are able to predict, explain and/or describe each other's behaviour not by appeal to some folk psychological theory (as the Theorists theorists claim), but by way of running a simulation of the other's states on our own minds. (Hence the name 'Simulation theory'.)

There are, though, two main differences between the kind of simulation described by Davies and Stone, above, and that at work (according to Simulationists) in interpersonal understanding. The first is that most of us don't often carry around spare gas cylinders, but we do always 'carry around' our own minds; we are never without them. (Indeed, it would seem that the very reason we went on to develop a theory of gas in the first place is because we *don't* always have the right kind of model ready-to-hand.) The second difference is that, in the gas cylinder case, the simulation is actually performed in real life, whereas in the case of interpersonal understanding it is performed in imagination. So, for instance, when I see you come inside after mowing the lawn on a hot day, red in the face and wet with perspiration, I don't myself have to go out and actually mow the lawn in order to work out that the next thing you will do is walk to the sink and pour yourself a glass of water; all I have to do is think to myself, "What would I do if I were him?"[168] – to 'imaginatively entertain' the desire for water – and the answer comes. As this is often put, what I do is I take my relevant cognitive mechanism 'off-line', temporarily severing its connection with my motor control centre, and sending its output instead to a belief-forming module, so that instead of actually pouring myself a glass of water (which I would, were this a real-life simulation) I only formulate a belief about what *you* will do.[169] Given subject-object similarity, I can infer that whatever answer comes to me in imagination, 'off-line', will be true of you in real life.

6.1.1a Subject-object similarity on the empirical account

Despite these two disanalogies between gas cylinder and interpersonal simulation, the fact that the two are comparable (to any extent) brings to light the driving idea behind empirical

---

[167] Davies & Stone 1998, pp.55-56
So as not to misrepresent these two philosophers, I should note that they do not endorse this description of Simulation theory, as the rest of the cited article makes clear.
[168] This description may be misleading. Most of the time, our folk psychologizing is quite automatic, and we hardly ever (though we might sometimes) have to explicitly *ask* ourselves: What would I do if I were him? Indeed, for those individuals who *do* have to work through it step-by-step like this, it is generally an indication of pathology.
[169] This is Stich and Nichols' (1995) presentation of Simulation theory, which has been very influential in the literature.

Simulationism: because you and I are sufficiently alike, I am able to use my own mind as a *model* of yours in order predict or explain your behaviour. The subject-object similarity insight is therefore cashed out by empirical Simulationism as implying that we can extrapolate from one case (the model) to another, when the two cases are of the same kind. Thus, I hope it is clear why I take the Analogical argument to be a particular application of empirical Simulationism. They both have the form: I am of kind X, and Y is the case with me; you are of kind X too, so what is the case with me is the case with you too (i.e. I am a suitable model of you.); therefore, Y is the case with you too.

I call this the 'empirical' version of Simulation theory for a number of related reasons. Firstly, because this is generally the understanding of Simulation theory had by those philosopher-psychologists who argue in its favour by way of appeal to empirical studies (particularly developmental ones).[170] Secondly, in order to contrast it with the Co-cognitive version of Simulation theory which is argued to be *a priori* true, and which will be the subject of section 6.1.2. And, thirdly, because this way of cashing out the subject-object similarity insight begs a question – an apparently empirical one: *am I* sufficiently and relevantly similar to you, so as to serve as a model of you? Davies and Stone note this condition with their remark, quoted above: "*Provided that* the cylinder B really is relevantly similar to cylinder A, this method [of simulation] is liable to yield an accurate prediction."[171] So, empirical Simulationism regards my folk psychologizing about you as an inference from me to you, from the outputs of an imaginative exercise in my case to your actual case – the accuracy of this inference depending on a further premise or condition: that I am relevantly and sufficiently similar to you, and thus a suitable model of you; that we share the same mental apparatus, and the cognitive mechanism on which I, off-line, simulate you is indeed the cognitive mechanism at work in producing your behaviour.

6.1.1b The Collapse argument

Now, this condition for reliable simulation – where this is conceived as an inference from me to you – poses a large problem to empirical Simulation theory (as advocated as an alternative to Theory theory). In the case of the gas cylinders, for instance, for A and B to be relevantly similar, and thus for B to be able to serve as a model of A, A and B must contain the same gases (helium, oxygen, methane, whatever). Ideally, they should also contain the same amount of gas in the same volume (i.e. under the same pressure) and should have the same initial temperatures. But what if they don't – what if, say, cylinder B contains the same gas, at the same temperature, as A, but contains less of it? Well, given that a change in the volume of a gas is inversely proportional

---

[170] One such notable Simulationist is Alvin Goldman, who states quite explicitly: "But whether [simulation] can in fact unlock any doors depends heavily on the outcome of empirical research in cognitive science" (1993/1995, p.185), and: "My framing of the problem deliberately construes it as an empirical question about the psychology of attributors" (p.186). See also Harris 1992.

[171] Davies & Stone 1998, p.56 (emphasis added)

to a change in its pressure, it shouldn't be too hard to make the necessary adjustments. But wait – how do I know that, that a change in the volume of a gas is inversely proportional to a change in its pressure? This is, of course, a derivative of Boyle's law – that is, a piece of the *theory* of gases, and that's the problem. It seems that in order to run a simulation at all, the simulator must rely on theory. Even if no adjustments need to be made – say the type, volumes and initial temperatures of the gases are identical – how do I know that these are the relevant factors? Why, for instance, is the colour of the cylinder not relevant?

Consider the folk psychological case: in order to work out what you will do after you come in from mowing the lawn, since I am not in exactly the same state as you (*I* didn't mow the lawn) I have to make an adjustment: I have to imaginatively entertain the desire for water (a desire I don't current experience). But how do I know that *this* is the adjustment I need to make? Why not some other desire, or some other mental state altogether? Well, the Simulationist may wish to respond, we know that the correct state to adjust for in imagination is the desire for water because we know that, generally, after people have engaged in strenuous activity, especially in the heat, they desire water – but, again, that sounds just like a piece of *theory* about when people desire what.

This is commonly posed as the 'collapse' argument against Simulation theory: the objection that, when all is said and done, Simulation theory collapses into Theory theory,[172] because any reliable simulation depends on theoretical assumptions.[173] So, the very assumption upon which empirical Simulationism rests – that I am relevantly and sufficiently similar to, and so a suitable model of, you – turns out to be its weakness. Although the empirical Simulationists do have certain ways of responding to these criticisms,[174] I will not entertain them here, given that there is a readily available, alternative version of Simulation theory, that avoids these issues altogether: Co-cognition.

6.1.2   Co-cognitive Simulationism

The Simulationist approach I have in mind here is something of a combination of the views of Jane Heal and, to a lesser extent, R.M. Gordon. On this account, the deepest difference between empirical Simulationism and Co-cognition can be put in terms of direction of gaze. When my simulating you, as the empirical approach claims, consists in an inference from me to you, my gaze is necessarily directed at you and at me: to discover what situation you are in; to discover what would happen with me, were in I in that situation; and to determine whether I am a suitable model of you (which, as we have seen, may well involve some impure appeal to theory). When it comes to Co-cognition, though, I, the simulator, gaze neither at me, nor at you. My gaze as the

---

[172] Ravenscroft 2003
[173] See Heal 1998b/2003, p.34ff, for further problems with empirical Simulationism.
[174] For instance, see Heal 1996a/2003.

simulator is directed, rather, at the *world*, *from your point of view*.[175] Heal captures this distinctive feature of Co-cognition as follows: when one person simulates another, they both have the same *content* to mind, because the simulator attends not to another person, but to the same outward phenomena, the same world, as the other is attending to. As Gordon puts the point, the simulator attends, not *to* another, but *as* another.[176] Say, for instance, we are trying to arrive at a folk psychological prediction about what deduction another person will make, given that she has the beliefs *p* and *If p then q*. According to Heal, what we do in order to arrive at the prediction is we imaginatively entertain *p* and *If p then q* – that is, we consider the *subject matter* of the other's thoughts, not the thoughts themselves, and work out what follows on that basis.[177] On this account, when one person simulates another, she doesn't think *about* the other's thoughts – she *thinks his thoughts*. That is to say, the subject and the object of simulation *co-cognize*, think together, by attending to the same subject matter;[178] hence the name of this particular version of Simulationism.[179]

6.1.2a   Interpersonal understanding on the Co-cognitive account

The Co-cognitive emphasis on point of view contrasts it with both Theory theory and empirical Simulationism. Recall, as we noted above, that the Theory theorist reasons that, because the way in which we go about achieving understanding of most of the rest of the world is by theorizing about it, our folk psychologizing must also be subserved by a theory. That is, Theory theory implies that the sort of understanding of other persons available to us is the same in kind (namely, understanding by subsumption) as that of other non-personal phenomena, like the weather, for instance. Perhaps surprisingly, the empirical version of Simulation theory – even though it stands opposed to Theory theory – has just the same consequence: the sort of understanding of other persons available to us by empirical simulation is the same kind of understanding that might be had of non-personal systems, such as Davies and Stone's gas cylinders, or of volcanoes, fields of wheat or kitchen mixers,[180] when we extrapolate from one case, the model, to another.[181] (I know why your kitchen mixer is behaving like that: my kitchen mixer behaves like that when I set it to 'supermix', so someone must have set yours to 'supermix'.)

---

[175] Blackburn 1992/1995, pp.277-278
[176] Gordon 1992/1995, p.55
[177] Heal 1998a/2003, p.38
[178] Heal 1998b/2003, p.98
[179] To be precise, Gordon does not describe himself as a Co-cognitivist. I hope I have not misrepresented him by taking those few elements of his view, which I present here, to be in line with and congenial to Heal's Co-cognition.
[180] These are Heal's (1998a/2003, p.42) eclectic examples.
[181] I think it is for this reason that Blackburn (1992/1995, p.277) refers to what I have called the empirical version of Simulation theory as the 'Model model' – as a play on 'Theory theory', so as to draw out their similar nature.

Co-cognition, on the other hand, offers us a unique kind of understanding, one that cannot be had of the weather, gas cylinders or kitchen mixers because Co-cognition exploits something these inanimate systems – unlike other persons – don't have: a point of view on the world. So, the special kind of understanding facilitated by Co-cognition is understanding *from the inside*, from the other person's point of view. As Heal explains:

> We come to something more distinctive of mind in the fact that, on the view sketched, when I consider the nature of what is 'inside' another person, in the [sense] suggested by [Co-cognition], what I find myself postulating is a set of thoughts which represent the world from a point of view. So the 'inside' which I find is not mere mechanical or biological complexity. If the inside were of that kind, there would not be any question of anything being 'from' it. But things can be 'from the inside' with a person because what is 'inside' is itself outwardly directed. […] Their thoughts and behaviour therefore have sense and can be justified in ways which have no analogue in the explanations provided for the behaviour of inanimate items.[182]

Thus, the interpersonal understanding afforded us by Co-cognition is not an understanding *of* the inside, such as that paramechanical insight facilitated by empirical Simulationism, relying on the assumption that you use the same kind of cognitive machinery as I do. Rather, it is an understanding that renders intelligible, an understanding *from* the inside, where the 'inside' is itself outwardly directed. To have an inside, then, on this Co-cognitive account, is just to have a point of view on the world, which is something gas cylinders and kitchen mixers do not have.[183]

6.1.2b   Subject-object similarity on the Co-cognitive account, and its a priority

So, the subject-object similarity insight is understood by this Co-cognitive version of Simulation theory to mean something like this: you are like me, not in the empirical, 'suitable model' sense, but in the sense that you, like me, occupy a point of view on the world. And, in true Kantian style, Heal argues that my taking you to be similar to me in this sense – as occupying a point of view on the world – is not a consequence of regarding you as, like me, a person; it is, rather, the *precondition* of my regarding you as a person at all.[184] Another person just is another intelligible point of view. Consider all these characteristically personal interactions: arranging a time to meet that suits both of us;[185] asking your tutor to explain to you why the 'gambler's fallacy' is indeed fallacious;[186] speculating as to whether a mutual friend will accept a job offer;[187] even something

---

[182] Heal 1998a/2003, pp.40-41

[183] This raises the interesting question of whether non-human but animate creatures – that is, the other (higher) animals – are simulable, given that they ostensibly also have a point of view on the world. (*Cf*. Nagel 1974.) This *is* an interesting question, but unfortunately not one I will be able to take up in this project. I will note, though, that the higher animals' points of view are limited in a way that ours are not: we are able to take a reflective point of view on ourselves, whereas they apparently are not. The interesting question that all this poses, to my mind, is whether the moral restriction on our practical identities encompasses not only other persons with points of view but animals too. Korsgaard (1996, pp.145ff) takes this up, but, of course, not from the Simulationist approach I have advanced here.

[184] Heal 1998b/2003, especially pp.100-103

This is apparently also the force of Gordon's remark (with regard to spatial simulation specifically): "When we are aware of other – that is, aware of them *as* others – we are constantly, automatically projecting onto them our own beliefs about the environment." (1992/1995, p.105)

[185] Korsgaard 1996, p.141

[186] Heal 2003, pp.242ff

[187] Heal 1996b, p.82

as simple as stepping aside so that somebody else can pass you;[188] and – most prominently, for this project's purposes – holding each other morally accountable. All these actions – mundane as they are, but nevertheless central to our lives as we know them – proceed on the assumption that the other, like you, has a point of view on the world.[189] (Try entering into any of these activities with a gas cylinder or a kitchen mixer.) And, if this is the case, then it means that for all individuals with whom I enter into such activities, there is a special kind of interpersonal understanding available to me: understanding from the inside, finding your actions intelligible. And, when I co-cognize with you, I view the world from your point of view, from the inside – my gaze is directed at the world from your point of view, not at you and then at me – so there is no inferring to do.[190] In Gordon's words, this kind of simulation is 'not a transfer but a transformation',[191] not requiring any inference from me to you.

By this stage, I hope that the significance of Co-cognition's non-reliance on inference is becoming clear – but my discussion of it must wait till the following section. First, let me clarify one point: in previous chapters (see section 2.4 in particular), we have spoken of 'persons' with the understanding that a person is the sort of creature who has practical identity. However, I have spoken of 'persons', above, as being the sorts of creatures who are intelligible and who have points of view on the world. I am not, though, working with two different conceptions of 'person'. Consider the following passage by Heal:

> A person becomes aware of her world and builds up a picture of it, through perception, memory and reasoning. And that view must be unified in the way sketched. But let us note also that her view will necessarily include, woven in among the rest, many indexical thoughts, defining her beliefs about herself, her placement, role, capacities and so forth. For example, they will include beliefs of the form 'I am in such an such a location', 'I am capable of these or those actions', 'I occupy such and such a role', 'These and those achievements, dangers, disappoints or pleasures are possible for me' and so forth. These elements may be said to define a 'point of view' on the world[192]

Heal, therefore, clearly takes the 'world' on which a person has a point of view to include the person's *self*. As we noted in section 2.8 (with reference to section 2.4), not only is a creature subject to reasons a person, she must also necessarily have a conception of herself as a person – likewise, not only does a person have a point of view on the world, her point of view includes

---

[188] Korsgaard 1996, p.141
[189] Interestingly, 'mundane' comes form the Latin: belonging to this world.
[190] To be fair, Heal's early work (see 1995) apparently *does* endorse the view that my co-cognizing with you involves an inference from me to you. However, I take this to be out of step with the rest of Heal's views, as I go on to describe them – and, indeed, this claim is not at all prominent in her later work on Co-cognition, although, to the best of my knowledge, she does not explicitly reject it. Furthermore, I take Gordon's (1995) argument against Heal in this regard to be correct. This is why I have described my presentation of Co-cognition as a combination of Heal's and Gordon's views.
[191] Gordon 1995, p.54
[192] Heal 1998a/2003, p.40

herself as in the world. [193] I take it, then, that a person's having a point of view on the world in this special sense includes her having a practical identity – that is, a point of view on herself. And it is persons' having points of view that facilitates their intelligibility to each other: the kind of point of view a person occupies is one another might also have. I build on these remarks, below.

6. 2   The Co-cognitive solution

We are now in a position to explore how Co-cognition improves upon empirical Simulation theory, and, more importantly, how Co-cognition is able to offer an alternative solution to the neo-Kantians' Analogical argument, which I take to be a particular application of empirical Simulationism.

The way in which Co-cognition overcomes the 'weakness' involved with empirical Simulationism's inference from me to you is the same way in which Kant overcomes Hume's sceptical conclusion about causality and the weakness he took it to imply: to deny that there is any inference involved at all, and to replace it with a constitutive claim. Recall from section 4.2 that, under Hume's theory, causality is conceived of as an inference from past cases to future ones (amounting to a causal law), with the inference relying on the problematic assumption that the future will resemble the past; and this Hume took to "(announce) a discovery about the weakness and narrow limits of human reason and capacity."[194] Kant's response to this sceptical conclusion was to deny that there is any inferring going on in the first place, and to contend that, rather, *what it is* to perceive an event as causal just *is* to regard it as being subsumed by a law which extends to future possible cases. Likewise, the Co-cognitive response to the threat of collapse, as faced by empirical Simulationism, is to deny that there is any inference from me to you (relying on problematically theoretical assumptions); and to contend, rather, that *what it is* to see you as a person in the first place is to regard you as being relevantly like me, in that you have a point of view on the world – which, as noted above, includes a point of view on yourself.

In section 5.1.2, I complained that the Analogical argument takes the wrong view of persons – and not, as we then went on to see Korsgaard complain in section 5.2, that it takes the wrong view of reasons. I can now give sense to this claim: the problem with the Analogical argument (which is a problem for empirical Simulationism in general, I think) is that it erroneously supposes that I can establish your being like me (in that you are a person) *independently* of my taking you to be subject to reasons, as separate premises of an analogical inference. Rather, there is no

---

[193] Furthermore, the roles and capacities to which the Heal passage makes reference strike me as being just the sort of thing we have called practical identities and the Frankfurtian restrictions (i.e. what a person is and isn't capable of caring about).
[194] Hume 1748/2010, §7, p.37

'you' – no other person – who is relevantly like me, unless I *already* take you to have a point of view on the world and a practical identity, like me. However, the main problem noted in that section, 5.1.2, was that the Analogical argument didn't actually achieve its objective of getting us from the categorical imperative to the moral law. Can Co-cognition do any better?

6.2.1   Co-cognizing practical identities

The idea as to how Co-cognition supplies us with the content of the moral law is really quite basic, but it does the job. The idea is this: a person knows what another's reasons are by simulating the other's practical identity, by putting herself in the shoes of the parent, or the businesswoman, or the Muslim, or whomever.

Insofar as we actually share one of the same practical identities – say, we are both parents – there is quite a vivid sense in which I know 'what it's like' to, for instance, experience your child's hurts as you own. However, as I argued in section 2.6, I take the ability to think about other *possible* ways of being – even to embrace non-actual descriptive identities – to be quite uncontroversial, and really just a product of our capacity for hypothetical thought. So, we are able to simulate others' reason, not only to the extent that we do actually share their practical identities, and therefore their reasons, but to the extent that we are able to co-cognize with them, to think the same thoughts as them, where, in our case, this may consist in hypothetical thought, (or 'imagination', as we put it earlier). I take this ability to co-cognize other practical identities to be quite plain, even on the basis of our discussion in this project so far (especially Chapter 2): I am not a parent myself (perhaps you aren't either) but it's quite clear to me (and hopefully as clear to you) that a parent has a reason to act in her child's best interest. Note that this is not just to say that most parents happen, as a descriptive matter of fact, to act in their children's best interests (because, unfortunately, some parents I have met do not), but that they are volitionally constrained to do so, even if their descriptive identities diverge. And this ability to know the reasons of others, even when we ourselves do not share their practical identities, extends well beyond parenthood: I am neither a businesswoman nor a Muslim nor an athlete, but it is quite clear to me what reasons, in what circumstances, these persons are subject to, when I take up their perspective on the world and co-cognize with them. So, the content of the moral law, whatever it might be, is given by co-cognition: considering the world from other points of view, which may be actual or possible points of view.

Furthermore, my knowing another's reasons in this way does not entail that I myself am therefore subject to those reasons too, as Korsgaard advocates and as I criticized in section 5.2.4. This depends on the idea, proposed in section 5.3, that when a person cares about something, she necessarily takes it to not only be subjectively important to her, as Frankfurt insists, but to have

objective value. This is because a person's practical identity is a law unto herself, and laws, according to Kant, have the effect of objectification. So, I may know another's reasons – say, the businesswoman's – by co-cognizing with her, gazing upon the world of objective value from her point of view, in which she chooses to stand in a special relationship to that objectively valuable thing – say, making it to the board meeting on time – by way of caring about it. Now, the objective value of making it to the board meeting on time means that this action is valuable from other points of view too – including my own. But, as we established in section 5.3, a person's recognizing a thing's objective value does not entail its subjective importance to her. So, I am able to co-cognize with the businesswoman, and thereby to find intelligible her action of trying to make it to the board meeting on time, without myself being subject to the volitional necessity of that reason for action. I can share her reasons without being subject to those reasons. So, Co-cognition can supply us with the content of the moral law, and can do so in a way that retains the distinction between my reasons and your reasons, and therefore the distinction between you and me.

6.2.2 Co-cognition and restriction

Consider the sentiment captured by, for instance, this not-uncommon expression in reaction to some perceived wrongdoing: "How would you like it if someone did that to you?"[195] or even the 'Golden Rule': 'Do unto others as you would have them do unto you'. Although these moral expressions strike me *prima facie* as calls to simulation, they do not tell us *why* another's not wanting one to act in a particular way places a restriction on one's reasons – unless, as Frankfurt might add, one cares about acting only in ways that others would want one to act, but this is again just a person-particular restriction. What remains to be shown, then, is that the content made available to a person by co-cognition (that is, others' reasons) places a moral restriction on feasible practical identities. My argument is that there must be a moral restriction on practical identity, because the resultant insularity of Frankfurt's solely person-particular restrictions poses an obstacle to interpersonal understanding. Although an individual may be surrounded by others, insularity renders her unintelligible and therefore isolated, in a deep sense, from other persons, from the world, and even from herself.

Consider Wolf's sadist,[196] as discussed in section 3.2.2. On Frankfurt's account, such an individual is indeed a person: she has constituted her self by way of the practical identity of a sadist; she cares about sadism, and her life is thereby shaped by the attendant reasons, or 'commands of love'. On my view, however, this is only half the story as to what it is to be a person. I contend that the sadist is *not* a person, insofar as she identifies herself with her sadism,

---

[195] Korsgaard 1996, p.142, referencing Nagel 1970
[196] See Wolf 2002, p.229

because her behaviour, would, in this regard, not be intelligible. But, why not? We are able to offer some explanation of her behaviour: she does what she does because she cares about hurting people, and therefore has a reason to hurt them. On the face of it, this does have the form of an intelligible explanation; compare, for instance: the businesswoman does what she does because she cares about making it to the board meeting on time, and therefore has a reason to skip breakfast. However, this form of explanation is also compatible with distinctly non-personal explanations; compare: the pressure of the gas cylinder increased as it did because its temperature increased – that is, the reason for the increase in pressure is that the temperature increased. So, we cannot glean too much from the surface form of explanation. The sadist's caring about hurting people may *look* like a reason for action, but is it one – in the distinctly personal sense in which we have used the term, and not in what I have called the mechanical sense? To achieve that specially personal understanding, understanding from the inside, we need to ask if we can share the sadist's reasons: can we co-cognize with her?

I certainly cannot co-cognize with the sadist. This is because, if she adopts the practical identity of a sadist – if she is a law unto herself in this way – then she must take hurting people to be not only important to her, but *objectively valuable*. However, hurting people is *not* objectively valuable, because it's not valuable from other points of view – namely, my point of view. I, for one, reflectively *reject* that possible way of being; that is to say, not only do I not-embrace sadism, I positively dissociate myself from that way of being. And this law that I am to myself serves to objectify my subjective stance, with the effect that I take sadism to have objective *disvalue*. So, from my point of view, sadism is valueless. But, when I attempt to co-cognize with the sadist, when I attempt to view the world from her perspective, I discover a world in which sadism is *valuable*. But this can't be. If sadism is valueless from my perspective – that is, *objectively* valueless, valueless from other possible perspectives – then somebody who takes it to be valuable is not attending to the same thing I'm attending to; that is, we're not co-cognizing. And if the sadist and I can't co-cognize, I can't understand her from the inside, and – though I may be able to cite some mechanistic explanation of her behaviour – she (insofar as she identifies herself as a sadist) is unintelligible to me.

But, one may object, that's just from *my* perspective. There might be other perspectives from which sadism is valuable, and the sadist's actions are intelligible. What about, for instance, the willing victim? There are some people out there, just as disturbed as the sadists themselves, who *want* to be hurt, who want to have the sadists do what they do *to them*. However, I contend that such an individual, who cares about her own harm (even if she isn't the one doing the harming) would find herself in a deep reasons-conflict (such as that described in section 3.1.2), so she isn't really a unified person herself. Her caring about anything necessitates her fundamental practical

identity as a person; yet her caring about her own harm serves to undermine her own life, and is thus deeply incompatible with her personhood. So this is not a perspective from which the sadist is intelligible, because it conflicts with a rational regard for the continuation of one's self. What, then, about other sadists? Sadism must also be valuable from the perspective of other persons with the practical identities of sadists, so they would all be intelligible to each other. A kind of insularity would still result: sadists wouldn't be intelligible to the rest of us, who regard sadism as valueless – but sadists would, presumably, nevertheless be intelligible to other sadists. However, I'm not sure this would work either: what sadists, I presume, count valuable is hurting *others*. So, when sadist 1 takes up sadist 2's point of view, he discovers a world in which he *himself*, sadist 1, being hurt is valuable – because sadist 1 is an 'other' from sadist 2's point of view. And this cannot be, for the same reason that it cannot be for the willing victim. So, even sadists, insofar as they each identify themselves with sadism, aren't (insularly) intelligible amongst themselves. And this means that, not only is the sadist unintelligible to *us*, but *we* are unintelligible to *her* (insofar as she is a sadist, and we are not). And, given the centrality of interpersonal understanding to our lives as we know them, the sadist is thereby drastically impaired, and her agency is eroded to the extent that she identifies herself with sadism.

6.2.3 Intelligibility and immorality

Does all this mean that we can never understand an immoral action? No, of course not; we can often offer quite comprehensive explanations of immoral behaviour. But notice how our explanations of immoral actions take either of four forms.

Firstly, we might appeal to some mechanical failure, as I called it earlier. For example: he says those rude things because he has Tourette's syndrome. In these cases, the behaviour is much like the unwilling addict's drug use, where the behaviour is not regarded as an action or expression of the person's real self, but rather as the result of some kind of physical malfunction. Whether or not such action is *intelligible*, therefore, is not an appropriate question. As Paul writes: "I do not understand what I do. For what I want to do I do not do, but what I hate I do."[197]

Secondly, we might explain a person's wrongdoing by reference to their *mis*judgement of value, generally the product of a poor upbringing. For example: having grown up in Nazi Germany, he *thought* that exposing Jewish refugees was the right thing to do. Here we make sense of the person's behaviour in the same way we make sense of the person who thinks that the moon is made of cheese, because the moon is a planet, and planets are made of cheese: though these beliefs are false, their *content* is connected in the right kind of way. Clearly, a person who thinks it true that the moon is made of cheese, or valuable that Jews be exposed and killed, is not

---

[197] Romans 7:15

attending to the same world we are, so we cannot co-cognize. There is some sense to be made in that they are connected to the world – *their* world, so to speak – in the right kind of way, but this does not amount to an understanding from the inside.

Thirdly, we might well view a person as having acted for reasons, yet we mitigate the person's blameworthiness by substantiating the motive for his action with further reasons. For example: he robbed the bank because his daughter is very ill and the medical aid refuses to pay for the treatment she needs to save her life. Such a person's actions, though they are immoral, are nevertheless intelligible: he didn't rob the bank because he identifies with robbing banks; he acted as he did because he cares about his daughter, and robbing the bank was a means to that end – and, *this* is an end, a practical identity, with which we *can* co-cognize: even though we do not stand in the special relationship of love to his daughter, she has objective value from our points of view too.

Lastly, we might explain an immoral action by appeal to the agent's nature. For example: he did it because he's criminally insane. In these cases, we can't co-cognize with the agent – and thus do not regard him as a person – because, even though his thoughts may be about the same content as ours, his thoughts do not map the relations of the subject matter as they should. When people are immoral in *this* way, we don't put them in prison, to teach them a lesson – because the precondition for the kind of personal interaction involved in 'teaching someone a lesson' is taking them to have a co-cognizable point of view on the world – rather, we put them in padded cells: we further insulate the unintelligibly insular.

Though we might be able to make some sense of a person's immoral behaviour in most cases, the important point is that, whatever understanding there is to be had, it does not consist in our co-cognizing with the immoral practical identity of a fellow person. An individual who takes an action to be objectively valuable that we take to be objectively valueless is not intelligible to us, insofar as that individual embraces that practical identity.

6.2.4   Intelligibility and other persons

It seems, however, that Frankfurt is *still* able to retort that this argument doesn't touch on a person unless she *cares* about being intelligible. The sadist, on Frankfurt's account, might be perfectly content, with a well-shaped life, replete with her own reasons for doing this and that – albeit reasons that are wholly unintelligible to others. The sadist would be insulated from others, but if she doesn't care about her intelligibility, she need be no less content, just her and her reasons. However, this *can't* be the case. As I hope to show, if we shun intelligibility, we undercut our personhood. If the sadist is not intelligible to other persons, she is not intelligible to

herself, insofar as she identifies herself with sadism. And if *she* can't make any sense of her own actions, she loses her grip on herself, her personhood, and her reasons. Given that, under Co-cognition, our understanding of other persons is the same as our understanding of ourselves, it follows that, when others are unintelligible to a person (as, for instance, the rest of the world is unintelligible to Wolf's sadist) that person is to that extent unintelligible to *herself*.

Now, this may seem to be a *non sequitur* – to suppose that others' unintelligibility entails one's own unintelligibility – but I contend that to think this way is to cling to the empirical version of Simulation theory, which, as we have seen in section 6.1.1b, is problematic. Let me explain. On the empirical approach, my simulating you is conceived of as an inference from me to you: mental state X precedes Φing in my own case, therefore – because I am a suitable model of you – your Φing is due to your X. Say, though, that I discover that your Φing is not actually due to your X (and that I have no alternative explanation of your Φing available to me), so this behaviour of yours is unintelligible to me. But your unintelligibility, of course, need not be inherited, by inference, from the unintelligibility of my own case – that is, it need not be that you are unintelligible to me because *I* am unintelligible to me. Rather, you are unintelligible to me, my simulation of you has not been successful, because I am *not* (contrary to our initial assumption) a suitable model of you; as it turns out, you do not have and use the same kind of mental apparatus I do. And this has no consequences as to my intelligibility (to me). Indeed, empirical Simulationism takes for granted that, no matter what, I am intelligible to myself. So, if we take the empirical approach, it is true that your unintelligibility does not threaten my intelligibility.

Now, the Co-cognitive approach – the approach I endorse – also takes for granted that a person must be intelligible to herself. An individual who could make no sense, have no understanding 'from the inside', of her *own* actions would to that extent fall short of personhood. The Co-cognitive approach differs from the empirical one, though, in that, according to empirical Simulationism, this kind of self-understanding is true independently of anything else, 'no matter what', as I put it above. On the Co-cognitive story, this is not the case: a person's own intelligibility is not independent of anything else; it is dependent on others' intelligibility to her; they are two sides of the same coin. This is why, if the rest of the world is unintelligible to the sadist, it follows that she is unintelligible to herself, which I take to mean that there is no meaningful sense in which the sadist (insofar as she identifies herself as such) is a person.

The crucial element here – that a person's own intelligibility depends on the intelligibility of others – is more prominent in the writings of the founding fathers of what is now known as Co-cognition, as I have described it. Together, these Co-cognitive forerunners comprise the *Verstehen* tradition: Max Weber, Wilhelm Dilthey, R.G. Collingwood, and – unsurprisingly – Kant, amongst

others.[198] Collingwood, for one, makes the point that my intelligibility goes hand-in-hand with others' quite explicitly; as he is often quoted: "the child's discovery of itself as a person is also its discovery of itself as a member of a world of persons".[199] The same notion is apparently key to attachment theorist Peter Fonagy's take on the psychoanalytical concept of 'mentalizing': that the development of the child's own mind is only made possible by her situation in an intersubjective community of other minds; "that the psychological, emotional, reflective self is discovered (or perhaps created) primarily as it is recognized and understood by others."[200] Perhaps more accessibly, I take this to be the same idea as that captured by the African ideal of *Ubuntu*: 'A person is a person through other persons'. On the face of it, this is not a moral injunction, but *Ubuntu* is commonly taken to have ethical implications – and, if my views are correct, this should hardly be surprising.

The rest of the passage from which the Collingwood quote, above, is taken, is given in terms of 'speaking' and 'listening'. It is not at all a purely linguistic point, though. In fact, the passage appears as part of Collingwood's aesthetic theory, where he is concerned with what it is to see a work of art as an expression of the artist. Likewise, our speech acts are expressions, and when a listener understands what is expressed by a speaker's utterance, he finds the speaker intelligible. As I will argue, precisely the same point applies to action generally. The rest of Collingwood's passage is this:

> The discovery of myself as a person is the discovery that I can speak, and am thus a persona or speaker; in speaking I am both speaker and hearer; and since the discovery of myself as a person is also the discovery of other persons around me, it is the discovery of speakers and hearers other than myself.[201]

The idea is this: for an action to genuinely be expressive of a person's real self – to be real speech, in Collingwood's example – she must be able to find intelligible her doing so. This is just the notion, encountered several times already in this project, that the precondition of a person's expressing herself in action is her reflectively taking a particular view of herself. But *this*, according to Collingwood, requires that she also be, not only the actor, but the *audience* of her action. In speaking, a person is both speaker and listener, and in acting, a person is both actor and audience.

Now, being the audience of your own actions means that you are able to shift points of view. (Hence, Collingwood's being regarded as a forerunner of Co-cognition.) Taking up a perspective *on* yourself, cannot be done *as* yourself, because you just *are* that point of view on the world from

[198] Heal 1998a/2003, p.29
[199] Collingwood 1938, p.248, as cited by Blackburn 1992/1995, p.277
[200] Wallin 2007, p.51, referencing Fonagy *et al*. 2000
My thanks to Dr Elisa Galgut for bringing this article to my attention.
[201] Collingwood 1938, p.248, as cited by Blackburn 1992/1995, p.277
Donald Davidson (1991, p.165) apparently cites this passage from Collingwood in motivating his idea that "consciousness of self necessarily involves that of others" (Kemp 2009, fn.1)

which certain things are valuable (as a result of your practical identities). Thus, I take Frankfurt's and Korsgaard's reliance on reflection, as discussed in section 2.5 – the notion that not only are we able to think *as* ourselves, from our perspectives, but also *about* ourselves – to commit both philosophers, minimally, to the claim that we are not bound by our own contingent points of view. But, if thinking 'as ourselves' is to think from our points of view, then thinking 'about ourselves' – that is, reflecting – is to attend to ourselves from *another's* point of view. So, in order for a person to act, she must find her so acting intelligible from another's point of view. Whose point of view? It need not be from the point of view of another *actual* person (although we may sometimes want to know how we are perceived by our friends and colleagues), but only from the point of view of a person generally conceived, as a creature with an intelligible point of view on the world, as some possible person. Person A finds her action X to be intelligible and expressive of her practical identity, when she finds, from person B's point of view, that X may be valuable. Were A to find that, from B's point of view, X cannot be valuable, then A cannot embrace that practical identity that counts X valuable, because she cannot make sense of herself under that description. I take this to be the force of Kant's remark that "even the most malicious villain (provided he is otherwise accustomed to using his reasons)" – that is, provided he has a reflective point of view on himself – "imagines himself to be this better person when he transfers himself to the standpoint of a member of the intelligible world".[202] That is to say, if the villain himself is to find his own actions intelligible – which involves 'transferring himself to the standpoint of a member of the intelligible world' – he must conceive of himself as a 'better person' unfreely compelled to behave immorally, where his villainous behaviour does not express his real self. He cannot intelligibly embrace his immoral descriptive identity.

So, a person's acting for reasons, expressing her own practical identity, is constitutively bound up with her taking into account other possible points of view. Practical identities that are not connected in the right kind of way to objective values, or to values that turn out not to be objective (which we could only know from other points of view), are not feasible practical identities, because they render the individual (insofar as she embraces such a practical identity) unintelligible to others, and thereby unintelligible to herself – which is to say, not a 'self' at all, to that extent. There is a restriction on those practical identities that are available to a person, and it is a restriction that comes from taking into account other possible persons – that is to say, a *moral* restriction.

Loving something valueless or immoral, as Frankfurt allows, is therefore not an exercise of one's autonomy, which Frankfurt is so zealous to defend. It is, rather, like the composer who in a

---

[202] Kant 1785/1959, pp.73-75, as cited by Korsgaard 1996, pp.237-238

moment of artistic rebellion spurns time signatures, musical keys and melodies: what he's left with is not a freer piece of music, but unintelligible noise.

References

Please note that, in citation, when I have not used the original publication of the cited work, I have given both the original date of publication and the date of publication of the edition or reprint that I have used. Page numbers refer to the later publication.

Blackburn, S. (1992) 'Theory, Observation and Drama' (Chap.13) in M. Davies & T. Stone (Eds.) (1995) *Folk Psychology: The Theory of Mind Debate*. Oxford, UK: Blackwell Publishers Ltd. pp.274-290.

Cohen, G.A. (1996) 'Reason, humility and the moral law' (Lec.5) in O. O'Neill (Ed.) *The Sources of Normativity*. Cambridge: Cambridge University Press. pp.167-188.

Davies, M., & Stone, T. (1998) 'Folk Psychology and Mental Simulation' in A. O'Hear (Ed.) *Current Issues in Philosophy of Mind*. Cambridge: Cambridge University Press. pp.53-82.

De Pierris, G., & Friedman, M. (2008) 'Kant and Hume on causality' in *The Stanford Encyclopedia of Philosophy*. Retrieved January 10, 2011, from http://plato.stanford.edu/entries/kant-hume-causality/

Frankfurt, H.G. (1971) 'Freedom of the will and the concept of a person' in *The Journal of Philosophy, Vol.68(1)*. Journal of Philosophy, Inc. pp.5-20. Retrieved February 12, 2009, from http://www.jstor.org/stable/2024717

Frankfurt, H.G. (1988). 'The importance of what we care about' (Chap.7) in *The Importance of What We Care About: Philosophical Essays.* Cambridge: Cambridge University Press. pp.80-94.

Frankfurt, H.G. (1999). 'On caring' (Chap.14) in *Necessity, Volition, and Love*. Cambridge: Cambridge University Press. pp.155-180.

Frankfurt, H.G. (2002) 'Reply to Susan Wolf' in S. Buss & L. Overton (Eds.) *Contours of Agency: Essays on Themes from Harry Frankfurt*. USA: A Bradford Book. pp.245-252.

Frankfurt, H.G. (2004) *The Reasons of Love.* Princeton: Princeton University Press.

Goldman, A.I (1993) 'Empathy, mind, and morals' (Chap.10) in M. Davies & T. Stone (Eds.) (1995) *Mental Simulation: Evaluations and Applications*. Oxford: Blackwell Publishers Ltd. pp.185-208.

Gordon, R.M. (1992) 'The Simulation theory: objections and misconceptions' (Chap.4) in M. Davies & T. Stone (Eds.) (1995) *Folk Psychology: The Theory of Mind Debate.* Oxford, UK: Blackwell Publishers Ltd. pp.100-122.

Gordon, R.M. (1995) 'Simulation without introspection or inference from me to you' (Chap.3) in M. Davies & T. Stone (Eds.) *Mental Simulation: Evaluations and Applications*. Oxford: Blackwell Publishers Ltd. pp.53-67.

Harris, P.L. (1992) 'From Simulation to folk psychology: the case for development' (Chap.10) in M. Davies & T. Stone (Eds.) (1995) *Folk Psychology: The Theory of Mind Debate*. Oxford, UK: Blackwell Publishers Ltd. pp.207-231.

Heal, J. (1995) 'How to think about thinking' (Chap.2) in M. Davies & T. Stone (Eds.) *Mental Simulation: Evaluations and Applications*. Oxford: Blackwell Publishers Ltd. pp.33-52.

Heal, J. (1996a) 'Simulation, theory and content' (Chap.4) in E. Sosa (Ed.) (2003) *Mind, Reason and Imagination*. Cambridge, UK: Cambridge University Press. pp.45-62.

Heal, J. (1996b) 'Simulation and cognitive penetrability' (Chap.5) in E. Sosa (Ed.) (2003) *Mind, Reason and Imagination*. Cambridge, UK: Cambridge University Press. pp.63-88.

Heal, J. (1998a) 'Understanding other minds from the inside' (Chap.3) in E. Sosa (Ed.) (2003) *Mind, Reason and Imagination*. Cambridge, UK: Cambridge University Press. pp.28-44.

Heal, J. (1998b) 'Co-cognition and off-line simulation: two ways of understanding the simulation approach' (Chap.6) in E. Sosa (Ed.) (2003) *Mind, Reason and Imagination*. Cambridge, UK: Cambridge University Press. pp.91-114.

Heal, J. (2003) 'What are psychological concepts for?' (Chap.12) in in E. Sosa (Ed.) *Mind, Reason and Imagination*. Cambridge, UK: Cambridge University Press. pp.225-249.

Hume, D. (1739) 'Personal identity' (Bk.1, Pt.iv, Sect.6) of *A Treatise of Human Nature*, paraphrased by J.F. Bennett (2010). Retrieved January 13, 2011, from: http://www.earlymoderntexts.com/pdfbits/humtr14.pdf . pp.132-140.

Hume, D. (1748) 'Sceptical doubts about the operations of the understanding' (Sect.4) and 'Sceptical solutions of these doubts' (Sect.5) in an *Enquiry Concerning Human Understanding*, paraphrased by J.F. Bennett (2010). Retrieved January 13, 2011, from: http://www.earlymoderntexts.com/pdfbits/he1.pdf . pp.11-27.

Kant, I. (1781/1787) *Immanuel Kant's Critique of Pure Reason* (Abridged ed.), translated by N. Kemp Smith (1934). London: MacMillan & Co., Ltd.

Kemp, G. (2009) 'Language' (Pt.5) of 'Collingwood's aesthetics' in *The Stanford Encyclopedia of Philosophy*. Retrieved January 6, 2011, from: http://plato.stanford.edu/entries/collingwood-aesthetics/

Korsgaard, C.M. (1993) 'The reasons we can share: an attack on the distinction between agent-relative and agent-neutral values' (Chap.10) in (1996) *Creating the Kingdom of Ends*. USA: Cambridge University Press. pp.275-311.

Korsgaard, C.M. (1996) 'Prologue', 'The normative question', 'Reflective endorsement', 'The authority of reflection', 'The origin of value and the scope of obligation' and 'Reply' (Lecs.1-4 and 9) in O. O'Neill (Ed.) *The Sources of Normativity*. Cambridge: Cambridge University Press. pp.1-166 and 219-258.

Korsgaard, C.M. (2006) 'Morality and the logic of caring' in D. Satz (Ed.) *Taking Ourselves Seriously and Getting It Right*. USA: Stanford University Press. pp.55-76.

McGinn, M. (1997) *Routledge Philosophy Guidebook to Wittgenstein and the Philosophical Investigations*. UK: Routledge.

Nagel, T. (1974) 'What is it like to be a bat?' in *The Philosophical Review, Vol.83(4)*. Duke University Press. pp.435-450. Retrieved August 25, 2008, from http://www.jstor.org/stable/2183914

Ravenscroft, I. (2003) 'Simulation, collapse, and Humean motivation' in *Mind and Language, Vol.18(2)*. Oxford: Blackwell Publishing Ltd. pp.162-174.

Stich, S. & Nichols, S. (1995) 'Second thoughts on Simulation' (Chap.5) in M. Davies & T. Stone (Eds.) *Mental Simulation: Evaluations and Applications*. Oxford: Blackwell Publishers Ltd. pp.87-108.

Wallin, D.J. (2007) 'Fonagy and forward' (Chap.4) in *Attachment in Psychotherapy*. USA: The Guilford Press. pp.43-58.

Wittgenstein, L. (1953) *Philosophical Investigations*, transl. by G.E.M. Anscombe. Oxford: Basil Blackwell Publishers.

Wolf, S. (1990) 'The real self view' (Chap.5) in J. Fischer and M. Ravizza (Eds.) (1993) *Perspectives on Moral Responsibility*. Ithaca: Cornell University Press. pp.151-170.

Wolf, S. (2002) 'The true, the good, and the lovable' (Chap.8) in S. Buss & L. Overton (Eds.) *Contours of Agency: Essays on Themes from Harry Frankfurt*. USA: A Bradford Book. pp.227-244.