

Friendship with an Ideally Coherent Eccentric:
A Problem for Korsgaard's Attitude-Dependent
Account of Value

by

Senthuran Bhuvanendra

A thesis submitted to University College London in
partial fulfilment of the requirements for the degree of
MPhil Stud in Philosophical Studies

I, Senthuran Bhuvanendra, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

This thesis proposes an interpersonal version of Derek Parfit's case of future Tuesday indifference, as a challenge to attitude-dependent accounts of value. Parfit's character and other "ideally coherent eccentrics" (ICEs) are used as counterexamples to the claim that a coherent set of attitudes is sufficient for an agent to have legitimate values. Sharon Street argues that a detailed picture of ICEs can dispel our intuitive concern. Using Christine Korsgaard's attitude-dependent account, I argue that an examination of different aspects of coherence cannot dispel this concern. This is explained by a failure of her account to prevent my ICE from choosing values which threaten his friend's integrity as an individual.

In chapter 1, I explain my case of an ICE who is indifferent to his friend's sadness on Tuesdays, and my choice to use Korsgaard's account to investigate that case. Korsgaard's account offers the opportunity to consider separately three notions of coherence: coherence in action-guidance, affective coherence (the coherence of evaluative experiences) and social coherence (coherence between group participants' actions and experiences).

In chapter 2, I argue that the requirements of coherence in action-guidance and affective coherence cannot be used to deny that my ICE has legitimate values. He can choose universalizable and consistent rules, and can avoid affective incoherence if he has no emotions or desires prompting him to reduce his friend's sadness on Tuesdays.

In chapter 3, I argue that the context of interaction with his friend cannot prevent my ICE's actions either. A further restriction, that both friends must have the same evaluative experiences, can be circumvented if both friends share Tuesday indifference to the one friend's problems. This is arguably a troubling result for Korsgaard's account, as it suggests that the friend's integrity is now undermined in both friends' perspectives, and by both friends' actions.

Table of Contents

Acknowledgements.....	5
Summary	6
Chapter 1: What’s the Problem with Friendship with an Ideally Coherent Eccentric?.....	15
1.1 Opening the Tuesday Case for Investigation	16
1.1.1 Tuesday cases, attitude-dependence and attitude-independence.	17
1.1.2 An interpersonal Tuesday case	21
1.2 Attitude-Dependent Accounts and Coherence	29
1.3 Korsgaard’s Story of Agency	34
1.3.1 The core story of action-guidance	36
1.3.2 The affective dimension	37
1.3.3 The social dimension	42
Chapter 2: The Ideally Coherent Eccentric’s Individual Agency	46
2.1 Coherence in Action-Guidance.....	47
2.2 Affective Coherence	54
2.2.1 Defining a notion of affective coherence	55
2.2.2 Application to <i>Friendship</i>	62
Chapter 3: Interacting with the Ideally Coherent Eccentric.....	66
3.1 Social Coherence in Action-Guidance	67
3.1.1 The universality and interpersonal consistency conditions	67
3.1.2 Application to <i>Friendship</i>	73
3.2 The Argument for Affective Social Coherence.....	78
3.3 A Final Picture: Shared Tuesday Indifference.....	86
Conclusion.....	94
References	97

Acknowledgements

First I would like to thank Lucy O'Brien for excellent supervision, and guidance over the many permutations of my work. I thank Peter Hulme, Bruce Macdonald, Rory Madden and Sander Werkhoven for detailed comments and discussions, and also Alex Geddes, Alec Hinshelwood, Mark Kalderon, Mike Martin, Veronique Munoz-Dardé, Léa Salje and Maarten Steenhagen for many helpful suggestions. Finally, but most of all, I would like to thank my mother, for all her support, her patience and her encouragement.

Summary

“How recognizable, how familiar to us, is the man so beautifully portrayed in the Grundlegung, who confronted even with Christ turns away to consider the judgment of his own conscience and to hear the voice of his own reason...this man is with us still, free, independent, lonely, powerful, rational, responsible, brave, the hero of so many novels and books of moral philosophy.” (Murdoch 1970: 80)¹

Consider the following scenario, which I will label *Friendship*.

Theodore is friends with Susan. Mostly, he acts towards her in the characteristic way that friends do. In particular, when Susan is sad, he responds to that situation as a friend. When it comes to other members of society, he might just smile sympathetically. But he takes time to ask Susan what is wrong, listens carefully to her problems, and takes action if he can solve them. He also acts similarly if he anticipates that something will make her sad, for instance, if she breezily tells him about a problem that he knows will ultimately be serious.

However, if this sadness occurs, or will occur, on a Tuesday, then he responds to her sadness as if she is just another member of society. If it is a Tuesday and he notices that she is sad, and if he has already greeted her that day and they have no other plans to spend time together, then he will give the sympathetic smile and walk on. This is not because Tuesday is their prearranged day for taking time off from the friendship: on some Tuesdays, they do spend time together.

¹ Permission to reproduce this extract has been granted by Taylor and Francis Group UK.

Similarly, if he anticipates that something will make her sad on a Tuesday, he does nothing about it, unless it is a severe problem that he would help avert in the case of any other member of society.

Susan knows that he behaves like this, so he is not leading her to rely on his emotional support and then unexpectedly letting her down, and he is not deceiving her in any way. Nor does Theodore's behaviour on Tuesdays prevent him from responding as a friend to Susan's sadness on other days. For instance, if a problem has made Susan sad on a Tuesday and it persists to a Wednesday, then he will ask about it, and Susan knows that it is worth talking to him despite his seeming indifference on Tuesday. Similarly, if a problem occurs on a Tuesday and he anticipates that it will cause sadness on a Wednesday, then he will act. He does this even if it is also causing sadness on a Tuesday – but it is not Susan's sadness on a Tuesday that is prompting him to act.

Something seems wrong with Theodore's behaviour here: the way he acts towards Susan and her sadness on Tuesdays seems unjustified. Perhaps his actions are not immoral, since he does respond in at least the minimal way that he does to other members of society, and he is not misleading or deceiving her. But there is something troubling about a friendship where he, or they both, find his behaviour acceptable.

Suppose we conclude that something is wrong: Theodore's action is unjustified, or he is not recognising the value that Susan has to him. Here is one way of explaining this. In this scenario, it seemed at first that Theodore valued Susan in a particular way: as a friend, and not just another member of society. We might define a "good" as a thing valued in some particular

way, so that Theodore's good seemed to be "Susan *qua* friend" – that is, understood by him as his friend. But when it comes to how she is affected on Tuesdays, Theodore does not act in a way that counts as an appropriate response to that good. He acts in a way that counts instead as an appropriate response to a different good: "Susan *qua* member of society". A good is either one thing or another. Theodore cannot define it one way on Tuesdays and another the rest of the week. It is true that Susan is not the same good to everyone. Not everyone is Susan's friend. So Susan's relation to Theodore matters. But even that relation is something Theodore has to recognise and respect. Susan has a certain authority independent of Theodore, and that determines how he should respond to her.

What I have outlined here is an attitude-independent account of value. There are facts about Susan's value which Theodore has to work out, which are independent of his attitudes towards her, and which he could be mistaken about. This account can extend to other normative terms: terms like "right", "justified" and "reason", which point out that something has authority over our actions, and maybe over our feelings.

Sharon Street (2009) and Christine Korsgaard (1996c, 2009) both believe that there are no such facts independent of our attitudes. Our attitudes give us a perspective, and what we value is explained by that perspective. Our perspective might be incoherent, or we might be mistaken about non-normative facts, but we cannot otherwise be wrong about value.

Derek Parfit (1984, 2011) believes that there are such facts, and that he has a good example to prove it. This is the example of someone who usually cares about pain and pleasure, except if they happen to him on a future Tuesday. Then he is indifferent. In fact, he would be willing to undergo an agonising

dental appointment on a future Tuesday without anaesthetic. This character's motivations are not incoherent, and he is not mistaken about the non-normative facts. So according to an attitude-dependent account, when he claims that he does not put a negative value on agony on Tuesdays, his claim must be legitimate. But Parfit says he is clearly wrong. We all have a reason to avoid agony, regardless of our motivations. Attitude-dependent accounts of value are wrong.

Street thinks Parfit's example does not prove that attitude-dependent accounts are wrong. We simply do not know enough detail about what this character is like to make a judgement. She gives us two versions of the character, Indy and Hortense. Indy does not care about agony on Tuesdays at all. When he experiences it, he has the ability to distance himself from his physical self's suffering, like a highly trained Buddhist monk. Once we have seen what his motivations are like, says Street, we know enough to know that Indy's values are legitimate. He is just different to us. But if we are struggling to accept this, an account of his origins, say of his species' evolution on another planet where such a Tuesday-focused ability is an advantage, may be enough to sway us.

Hortense, meanwhile, does care about the agony on the Tuesday itself. This means that she must plot against herself to make sure she does not escape the dental appointment on the actual day. Hortense is different to us too, but in a different way. She almost seems like two separate people, and the way she treats her Tuesday self is cruel. Street argues that we should see the problem here as a moral failure. There is nothing specific about her indifference to Tuesdays that we need to explain.

I think Street is right to argue that we need more detail to understand Parfit's case. But I think she is wrong to conclude that she has given enough detail. Indy and Hortense are actually quite similar. They both seem divided against themselves: Indy from his physical self, and Hortense from her Tuesday self. But this way of describing their cases makes it harder to see what is happening. To what extent, and in what ways, are they coherent? Are there internal conflicts in the way they direct their actions? In their experiences? Are we really convinced that Indy's case is not troubling? That Hortense's case has nothing specific to do with indifference to Tuesdays?

Because the idea of one person having two selves is confusing the issue, I turn to my case of Susan and Theodore, *Friendship*. Theodore has a type of indifference to Tuesdays, like Indy and Hortense. But it is someone else's experiences on Tuesdays, Susan's, to which he is indifferent. So we can focus how things are for Theodore without that layer of internal conflict. We can work out what it would be for him to be coherent in the way he guides his actions. I call this *coherence in action-guidance*. We can work out whether there is a requirement for him to have coherent experiences of the world, and of Susan's sadness in particular, and what he would be like if he did. I call this *affective coherence*. At the same time, although the internal conflict may have gone, this case introduces an interpersonal context. Theodore is interacting with Susan, since they are friends. So they will also be deciding together on actions, and sharing experience. We need to consider whether they need to have what I call *social coherence*, and again, what they would be like if they did. In the end, we will have a detailed picture, and we will be able to judge whether Theodore's values are legitimate or not.

To investigate these questions in detail, I use one particular attitude-dependent account, Christine Korsgaard's. Korsgaard's account allows us to consider the contribution of each of the three forms of coherence I have just mentioned, one at a time. This is because the way she explains action naturally falls into three components. She tells a story of how humans act, or successfully achieve agency. The core story is about guiding one's actions, by choosing rules to govern them. Another dimension involves the experiences involved, our emotions and desires, which make things seem good or bad to us. Those feelings prompt us to act, and we choose rules as we decide whether to listen to those promptings. Another dimension is to extend the story of how individuals act to give an account of how groups act. We can think of each dimension of the story as explaining how to achieve one of the three forms of coherence. So proceeding through the story, and applying it to our case *Friendship*, will allow us to gradually build up our picture of Theodore and Susan.

This is my strategy for investigating our case of Tuesday indifference. But before this, in chapter 1, I explain the context for my investigation in more detail than I have here. In §1.1, I give Street's set-up of the metaethical debate between attitude-independent and attitude-dependent accounts. I discuss her treatment of her cases of Tuesday indifference, and I explain my own case, *Friendship*. In §1.2, I explain the importance of the idea of coherence to attitude-dependent accounts. I then explain the three notions of coherence that structure my investigation. In §1.3, I outline Korsgaard's account, and the three dimensions of her story of agency.

In chapter 2, I discuss the first two dimensions of the story, which apply to the case of an individual agent. This allows us to consider Theodore's coherence, outside the context of his interaction with Susan.

In §2.1, I explain that Korsgaard's account of how individual agents guide their actions gives us two conditions for achieving coherence in action-guidance. In Korsgaard's account, agents must choose rules which are applicable across their lives and in all possible situations. This implies a condition of *universality across a life*. Korsgaard also tells us that agents must overcome possible conflicts between the different roles and relationships they have. This implies a condition of *consistency*. I argue that Theodore could have a rule that satisfies both these conditions, and allows him to act as described in *Friendship*.

In §2.2, I set up a notion of affective coherence. I use Korsgaard's claim that evaluative experiences of objects are our judgements of their goodness or badness. The point of these judgements is not to describe the world, but to feel a certain way about it. But since they are judgements, they can be related to judgements about the rightness or wrongness of actions towards those objects. I use the idea of judgements about rightness confirming judgements about goodness to define the notion of affective coherence. This notion allows us to differentiate everyday inner conflicts from the kind of conflict Theodore would have if he had emotions and desires to reduce Susan's sadness on Tuesdays. Theodore would be affectively incoherent. I conclude that, if we want to keep analysing Theodore as a coherent agent, then he cannot have emotions and desires to reduce Susan's sadness on Tuesdays.

In chapter 3, I add the social dimension of the story. This involves new versions of coherence in action-guidance and affective coherence. These help us consider how Theodore must be in order to share actions with Susan as a "group agent". We also start building a picture of what Susan must be like.

In §3.1, I argue that Korsgaard's account of group agency implies two conditions, which perform analogous functions to the two conditions set out in the individual case (§2.1). Group agents must have rules that apply across their members, implying a condition of *universality across group membership*. They must overcome possible conflicts between their members, implying a condition of *interpersonal consistency*. These conditions mean that we have to adjust Theodore's rule for action slightly, but he does not need to adjust his behaviour towards Susan. The conditions imply that Susan would have to act towards herself like Theodore does, if she were in his situation. This means, for instance, that if she lacked information about her problems (as Theodore often does), she would have to use some method to find out. But she does not have to ignore her own sadness on Tuesdays, because she can carry out her own sadness-reducing activity independently of cooperating with Theodore. Susan does have to accept his behaviour, but we can stipulate that she does in our case.

In §3.2, I present a suggestion for what the social version of affective coherence would look like. Evaluative experiences in the individual case are representations of objects which automatically prompt responses. These are produced by our valuing capacities, psychological mechanisms which link representations with the call to respond. I use David Velleman's (2000) account of shared intention as a model, to set out an account of shared evaluative experiences. These would be shared representations of objects which automatically prompt responses by group members. They would be produced through our participation in social practices of representing objects, which again link the representations with the call to respond. These could include bodily expression of emotions as well as language. But these social practices presuppose that participants have the same individual evaluative experiences of the relevant objects. This implies a notion of *social*

affective coherence for the individuals. To have shared evaluative experiences, their individual evaluative experiences must be the same.

In §3.3, I apply this final part of the story to Susan and Theodore. The requirement of social affective coherence means that Susan must have the same feelings as Theodore, with regard to the objects their friendship focuses on. We have to redefine our case: since the problems making Susan sad on Tuesdays do not worry Theodore, Susan cannot be sad about them either. But we can still refer to these problems, or Susan's underlying projects. In our final picture, Susan and Theodore do not respond to changes in Susan's projects which bring about certain results on Tuesdays, even though they do on other days. They also do not care about these results on Tuesdays. I argue that this is problematic. But it is not because of cruelty, as in Street's case of Hortense. Rather, it threatens Susan's integrity, and the integrity of the individual is a crucial ethical commitment.

Chapter 1: What's the Problem with Friendship with an Ideally Coherent Eccentric?

Unlikely scenarios can be useful to consider. This certainly seems to be the opinion of many metaethicists, who have devised many examples of extremely implausible characters. These are people who supposedly deem important things that, surely, no one in the real world does: counting blades of grass, not scratching their fingers, or whether something happens on a Tuesday. Sharon Street calls these characters “ideally coherent eccentrics” (2009). They have been proposed, for instance, to attempt to prove that our values are independent of our attitudes. As Street points out, this is not just a technical issue. The debate ultimately is about whether we should stay faithful to ourselves or to something independent of us, and this can shape how we live our lives.

In this chapter, I set up an investigation of a version of perhaps the most implausible, and thus most troubling, of these characters. Theodore thinks it is legitimate to ignore his friend Susan's sadness if it occurs on a Tuesday. He has a pedigree running through Street's own discussion back to an example given by Derek Parfit (1984, 2011). I explain the role of ideally coherent eccentrics in the debate between attitude-independent and attitude-dependent accounts of value, and my motivation for investigating Theodore's behaviour, in §1.1.

Part of that investigation will involve ensuring that our character Theodore is genuinely coherent, and setting out a detailed picture of what he looks like if he is. I explain the role of coherence in attitude-dependent accounts, and

suggest three notions of coherence that Theodore might need to satisfy in §1.2.

I will use Christine Korsgaard's attitude-dependent account to investigate my case. That account can be understood as a story of how agency is achieved, with the story having three dimensions. I outline them in turn in §1.3.

1.1 Opening the Tuesday Case for Investigation

In this section, I will explain my choice to discuss *Friendship*, the case of Tuesday indifference (for short, "Tuesday case") that I set out at the beginning of the summary. I will put the case in the context of a debate in metaethics. This is between those who see ethical claims and judgements as attitude-independent, and those who see them as attitude-dependent. The main character in our Tuesday case, Theodore, can be seen as a version of a certain kind of example used in these debates. Sharon Street labels such characters "ideally coherent eccentrics" (2009). My own Tuesday case is derived from a case from Derek Parfit: a character who is indifferent to pain on future Tuesdays (1984, 2011). Street defends attitude-dependent accounts against Parfit's original example, and against an adaptation of it.

I carry out this exposition in §1.1.1. In §1.1.2, I will argue that Street too quickly dismisses the challenge of Tuesday indifference to attitude-dependent accounts. My own Tuesday case aims to capture this challenge. I conclude the section by putting my case, *Friendship*, in the context of my discussion of Street's cases.

1.1.1 Tuesday cases, attitude-dependence and attitude-independence

Suppose I make the claim that some entity, *X*, matters, or is good, or is valuable. Perhaps this entity is an object, or a person, which I claim should be protected, or encouraged to flourish. Or perhaps it is a state of affairs that I claim should be achieved, or an action that I claim is worthwhile to perform. According to an attitude-independent account of value, there is a fact about whether *X* is valuable (etc.) or not (Street 2009: 274). In making my claim, I can be mistaken about that fact. That fact obtains independently of my attitudes towards *X*, or my mental activity about it. If my attitudes or activity were different, the fact would still obtain.

By contrast, an attitude-dependent account holds that there are no such independent facts. I have some perspective on, or stance towards, *X* (and the world generally), that is constituted by my attitudes or mental activity. My valuing *X*, or seeing it as good, or it mattering to me, follows from that perspective or stance. I can be criticised as incoherent, or as uninformed (for instance, if I am mistaken about the nature of *X*). But there is no normative fact about *X* for me to be mistaken about.

Sharon Street suggests the term “ideally coherent eccentric” to label a type of character often used to attack the attitude-dependent account (2009: 273-4). Such a character has a coherent perspective: her attitudes are consistent with each other. She is also fully informed, in that she knows all the non-normative facts. According to an attitude-dependent account, since these conditions are fulfilled, whatever such a character values is indeed valuable for her. But the ideally coherent eccentric values something so strange that it is difficult to accept this conclusion. Street gives a number of examples, including some proposed by adherents of the attitude-dependent account: David Hume’s example of someone preferring the destruction of the world

to the scratching of his finger, and John Rawls' example of a man who only enjoys counting blades of grass. Perhaps the strangest example (as she points out) is Derek Parfit's hedonist who is indifferent to pains and pleasures on future Tuesdays (1984: 124; 2011: 56, 79). This person's preferences are not only objectionable, but arguably make no sense. It therefore seems the most troubling for the attitude-dependent account. I will focus on Street's treatment of this case.

Street selects a few examples of ideally coherent eccentrics and gives a case-by-case defence of the attitude-dependent account. She aims to present each case in enough detail to make clear that it satisfies the conditions of consistency and full information mentioned above. This detail exceeds that usually given for such cases in the literature. Street argues that this fuller picture either dissolves our original intuitions about the unacceptability of the character's values, or even reverses them (2009: 281).

Street works with two versions of Parfit's case of future Tuesday indifference, the original version and an adaptation. She claims that the original case prompts a mix of intuitions from different sources (2009: 284-5). To separate intuitions about pain and the arbitrariness of Tuesdays from those about personal identity and cruelty, she considers the adapted version of the case first. Street labels her adaptation *Consistent Tuesday Indifference (CTI)*, and the original version *Future Tuesday Indifference with Present Tuesday Horror (FTI-PTH)*. In both versions of the case, her characters (Indy and Hortense, respectively) are indifferent to pain on future Tuesdays (2009: 282-4). They prefer scheduling an agonising dental appointment on the following Tuesday, when anaesthetic will be unavailable, to scheduling a mildly painful one on the Wednesday.

In *CTI*, when Tuesday arrives, Indy persists in his indifference. Street uses Parfit's terminology to explain this. Indy has a strong hedonic dislike of the sensations he will experience during his Tuesday appointment, compared to a mild hedonic dislike of the sensations he would experience on the Wednesday. Thus we can understand Indy's agony as the complex made up from two components: the sensation on Tuesday and his strong hedonic dislike of it. The mild pain he avoids is the complex made up from the sensation on Wednesday and the mild hedonic dislike of it that he would have had.

He also has desires directed at these experiences, called meta-hedonic desires. In particular, in the days preceding the appointment, he has a meta-hedonic desire to avoid the mild pain on Wednesday, and no meta-hedonic desire about the agony on Tuesday. Equivalently, he has a meta-hedonic preference for Wednesday's mild pain over Tuesday's agony. When Tuesday's appointment arrives, he has the same set of meta-hedonic desires.

Street fills in the detail of what the appointment is like for Indy. As the sensations begin, he feels his hedonic dislike of them. He feels "hatred and resistance" of them, "as though his physical self is in all out revolt, begging the dentist to stop" (2009: 285-6). At the same time, he is meta-hedonically indifferent. His "core emotional being" is calm, and "at the level of his innermost self", he contemplates his sensations and hedonic dislike with curiosity and detachment. Street makes a comparison with the experience of meditation, and the self-immolation of the Buddhist monk Thich Quang Duc in South Vietnam in 1963².

² Permission to reproduce the extracts quoted in this paragraph has been granted by John Wiley and Sons.

Street argues that this motivational structure is enough to justify Indy's choice to schedule his appointment on Tuesday. But to strengthen the reader's intuitions, she presents a genealogical story of how this motivational structure arose (2009: 286-7). Indy is a member of a species from another planet, where it is an evolutionary advantage to be meta-hedonically indifferent to pain at a certain point in the planetary cycle. This translates into Tuesday indifference when such a being is on Earth. With this full picture of Indy in sight, Street argues that it is no longer intuitive to say that his Tuesday indifference is irrational.

I now turn to Street's treatment of the original case (2009: 289-92). In *FTI-PTH*, Hortense has the same hedonic dislikes as Indy, but her meta-hedonic desires are different. As soon as Tuesday arrives, she has a strong meta-hedonic desire not to experience the agony of the appointment. However, like all ideally coherent eccentrics, Hortense is fully informed, so she knows when she makes the appointment that her meta-hedonic desires will change. Therefore, she must take drastic steps to ensure that her future self does not escape the appointment. Street mentions her hiring thugs to take custody of her on Monday evening, and maybe every Monday evening (in order to prevent her Tuesday self coordinating counter-measures across Tuesdays).

Street grants Parfit's stipulation that this character has no false beliefs about personal identity, and therefore does consider herself as one person. But she argues that Hortense must have a lack of compassion (and an inimical stance) towards herself on Tuesdays which makes her very different to ordinary people. Street insists that we should see Hortense as failing to be moral. She suggests turning to how attitude-dependent accounts treat the failure of an ideally coherent person to be moral. Street herself discusses

Allan Gibbard's case of an ideally coherent Caligula who only values maximising others' suffering (I will not consider that discussion here).

1.1.2 An interpersonal Tuesday case

Street's analysis, I claim, differentiates her two versions of the Tuesday indifference case too sharply, and amalgamates the original version with ideally coherent immorality too quickly. Indy in *CTI* is more similar to Hortense in *FTI-PTH*, and therefore more troubling, than Street suggests. And the way in which they are troubling is more distinctive than she allows. Street dismisses the troubling nature of Tuesday indifference too easily.

She rightly points out that it is natural to think of Present Hortense and Tuesday Hortense as different people. But we could also interpret Indy this way, if we looked at his description more closely. Notice that, on the one hand, Street refers to Indy's "physical self" as the personification of his hedonic dislike of the Tuesday appointment sensations; whilst on the other hand, she refers to his "core emotional being" and "his innermost self" as the personification of his meta-hedonic indifference. Presumably, it is his "innermost self" that is preventing Indy's writhing around and resistance from spilling over into an actual attempt to escape the dentist's chair. This is somewhat similar to Present Hortense's thug-assisted machinations against Tuesday Hortense.

Yet despite using the language of intrapersonal splitting to describe Indy, Street never suggests that he might falsely believe that he is two people. By contrast, she suggests that this could be true of Hortense: that Parfit's stipulation might not hold, and that Present Hortense might falsely think of Tuesday Hortense as another person. Similarly, Street does not accuse Indy of the kind of interpersonal failings that she does for Hortense, the lack of

compassion and mercy. Yet we might say that there is something cruel in Indy's detachment from his physical suffering, especially when he imposes it on himself for so little reason.

I suspect that Street's way of describing Indy confuses our intuitions and makes him seem less problematic than he might actually be. Elements that might seem conflicting can be attributed to different personas. At the same time, we feel familiar with applying the locutions for these personas, phrases like "physical self" and "core self", to a single individual. We therefore do not go so far as to see him as a person divided against himself like Hortense.

For instance, by writhing around in response to his pain, Indy *is* trying to escape, in a sense. Yet he is also keeping still in response to his meta-hedonic desire. In terms of the actions they prompt, Indy's motivational states are in severe conflict. We might think we can attribute one motivational state to Indy as really his own, because it is his "core self" doing the keeping still. That is who he really is, supposedly. But perhaps we are wrong to think that he is coherent in this way³. We might characterise Hortense as similarly lacking coherent desires, which forces her to perform actions that (deliberately) contradict her actions at other times. Indy might just be a less striking version of Hortense.

We can bring out the questionable nature of Street's description by applying a different set of personifications to Indy's situation. In general, we do not

³ Christine Korsgaard offers a different perspective of what it is like to endure an unpleasant sensation calmly (1996b: 147). Pain is an impulse to escape the sensation, and it is key to her picture that the painfulness decreases as the agent stops fighting it. Thus there is not the same conflict in tendencies to act that we see with Indy. Perhaps this is also more what Thich Quang Duc's experience was like.

say that whenever one has a pleasurable or painful sensation, one must desire to act in accordance with it, on pain of inconsistency. Suppose we speak here of one's overall, enduring self resisting one's momentary self. This repels the temptation to imagine that there are two distinct people here, of equal reality. But can we think of the Indy who writhes in agony as a momentary self, resisted by his calm enduring self? Imagine the dental operation lasts two hours. That is a lot of writhing. On the other hand, perhaps such reactions occur moment to moment, whilst his endurance can be understood as an overarching action. So this set of personifications does not settle the question of the coherence of his actions. But it gives a different perspective to Street's.

Similarly, consider Indy's experience of what is happening to him, the way he sees it. On the one hand, he feels a sense of revolt against it, but on the other, he is detached and deeply calm. Again, because Street describes his "core emotional being" as the one who is deeply calm, we have the sense that these are his *real* feelings, and there is a lessened sense of conflict. But maybe this description is misleading. On the other hand, we often do find ourselves in situations where we say that we really feel one emotion and that we only seem to feel another. So this question, the question of the coherence of his experience, is one that needs to be unpacked.

I have suggested that whether Indy is coherent is more of an open question than Street's description of him might indicate. Without settling that question, we cannot adjudicate whether his values are legitimate. I have also suggested that he is more similar to Hortense than he at first appears. But that comparison works both ways. Hortense is more similar to Indy than she at first appears. This might mean that Street should not move so quickly towards seeing the issues surrounding Hortense's case as moral. For

suppose, on further investigation, that it turned out that Indy is coherent, and that his is just a rather unusual case of managing one's desires, emotions and pains, along the lines that Street suggests. Then it might be possible to characterise Hortense in the same way. Hortense's agony on Tuesdays might be like Indy's physical agony, something she thinks it worth enduring, from her overall perspective. It is just that she needs external help, from her gang of thugs, whilst Indy can rely on his genetically given powers of self-restraint. Street compares her to Ulysses lashing himself to the mast to resist the call of the sirens. But we do not think of Ulysses as lacking compassion for himself. A clear difference between Ulysses and Hortense is that Ulysses is clearly responding to something legitimately valuable, his survival. Whether Hortense and Indy are failing to respond to something legitimately valuable when they put themselves through agony on Tuesdays is precisely what is at issue.

I have argued that we have yet to delineate a clear picture of a coherent agent with Tuesday indifference. So we cannot yet tell if such a picture would satisfy our intuitions. My comparison between Hortense, Indy and Ulysses suggests that Tuesday indifference may indeed involve a failure to respond to value.

I have also argued that the way that Street uses different personifications to describe Indy's case may make it more difficult for us to make an assessment. The same may apply to locutions like "Present Hortense" and "Tuesday Hortense" in Hortense's case.

One way forward would be to shift to a genuinely interpersonal case, translating these intrapersonal issues so that we can see them clearly, in the open. We could do this in a way that avoids bringing in moral issues *per se*.

In this way, we could keep the focus on the challenge presented by Tuesday indifference cases, namely, the apparent failure to respond to a value. However, the value would now be located within another person who the ideally coherent eccentric is interacting with.

This shift has another advantage. It makes the case, as a problem for attitude-dependent accounts, more forceful. It is not clear that an Indy-type explanation, even when spelt out at length, will be as convincing for a genuinely interpersonal case. Let us put this in terms of pain rather than agony, since agony might bring in moral issues. If someone ignores another's pain because it occurs on Tuesday, it may not be sufficient justification to explain that his motivational structure just happens to be that way. Iris Murdoch writes compellingly about the necessity, when interacting with others, of moving beyond our existing motivational structures to focus properly on the other person (1970). If the attitude-dependent account is to defeat the challenge of Tuesday indifference, it should defeat it in its most threatening form.

I have argued that we should shift from the cases of Tuesday indifference that Street examines to one involving interpersonal interaction. I will now restate the Tuesday case I set out in the summary, *Friendship*, and explain how it relates to the discussion above.

Theodore is friends with Susan. For the most part, he acts towards her in the characteristic way that friends do. In particular, when Susan is sad, he responds to that situation as a friend. When it comes to other members of society, he might just smile sympathetically. But for Susan, he takes time to ask her what is wrong, listens carefully to her problems, and takes action if he can solve them. He also acts

similarly if he anticipates that something will make her sad, for instance, if she breezily tells him about a problem that he knows will ultimately be serious.

However, if this sadness occurs, or will occur, on a Tuesday, then he responds to her sadness as if she is just another member of society. If it is a Tuesday and he notices that she is sad, and if he has already greeted her that day and they have no other plans to spend time together, then he will give the sympathetic smile and walk on. This is not because Tuesday is their prearranged day for taking time off from the friendship: on some Tuesdays, they do spend time together. Similarly, if he anticipates that something will make her sad on a Tuesday, he does nothing about it, unless it is a severe problem that he would help avert in the case of any other member of society.

Susan knows that he behaves like this, so he is not leading her to rely on his emotional support and then unexpectedly letting her down, and he is not deceiving her in any way. Nor does Theodore's behaviour on Tuesdays prevent him from responding as a friend to Susan's sadness on other days. For instance, if a problem has made Susan sad on a Tuesday and it persists to a Wednesday, then he will ask about it, and Susan knows that it is worth talking to him despite his seeming indifference on Tuesday. Similarly, if a problem occurs on a Tuesday and he anticipates that it will cause sadness on a Wednesday, then he will act. He does this even if it is also causing sadness on a Tuesday – but it is not Susan's sadness on a Tuesday that is prompting him to act.

I have chosen to look at friendship as a paradigm case of two people valuing each other directly. This contrasts with a working relationship, or other cases of fellow members of an organisation, where a large part of valuing each other has to do with the shared goal which they are pursuing together. This is also a case where what is centrally at issue is how Theodore values Susan as a friend, rather than his moral obligations towards her. Here I am understanding morality as a universal code, applying to our interactions with each other as humans (or at least as fellow members of some society)⁴. Theodore does have moral obligations towards Susan, as his fellow human being. But I have set up the case so that he seems to be fulfilling those. It might turn out, on further investigation, that his Tuesday indifference towards her means that he is actually failing her morally. But at first sight, what seems troubling is the way he is valuing her within their particular personal relationship.

In my description of the case, I have said nothing about Theodore's motivational structure – about his emotions, pleasures, pains or desires – but only described his actions. I leave this to be filled in as we discuss the case further, and in particular, to depend on the particular attitude-dependent account we will be considering (an issue I will discuss in the next section). I have referred to Susan's emotion, sadness, but this is because it is what (supposedly) activates Theodore's concern for Susan. I have thus moved

⁴ Another motivation for sidelining the issue of morality is that I find promising Bennett Helm's current project of understanding personal relationships first and then building on that account to understand morality (cf. Helm 2010, 2014, forthcoming), an approach that can also be suggested by a reading of P.F. Strawson (1974). In addition, I share concerns expressed by Elizabeth Anscombe (1958) and Bernard Williams (1985) that an excessive focus on morality, understood as a universal code governing behaviour, might distort our understanding of ethics.

what was an internal object of concern for Indy and Hortense in Street's cases, *CTI* and *FTI-PTH*, to be located in another person in *Friendship*.

I have shifted from considering physical agony to a painful emotion (sadness). I make the shift from agony to minimise the role of moral considerations in the case. Barring extenuating circumstances, most societies will impose a moral obligation on their members to respond to each other's agony in some way. Parfit presumably needs to use ignoring agony in his prudential case as, supposedly, a particularly egregious example of prudential failure. But given the context of friendship, ignoring someone's sadness is enough to also be egregious. I make the shift from physical to emotional pain to provide an example of a more common situation (someone being sad) that might arise in a friendship.

Sharon Street's treatment of Derek Parfit's future Tuesday indifference case dismisses the challenge it poses to attitude-dependent accounts of value. I have argued that it does this too quickly. Street separates the case into two versions and seemingly isolates the aspect of intrapersonal conflict. But this makes the apparent arbitrariness of Tuesdays too easy to explain away. Her detailed spelling out of the case raises helpful questions about the actions her characters are motivated to perform and about their felt experience. It is still worth investigating a case of Tuesday indifference as a challenge to attitude-dependent accounts, but adapting the case to involve personal interaction. Having explained my own Tuesday case, *Friendship*, in this section, I will discuss the attitude-dependent account I will use to investigate it in the next.

1.2 Attitude-Dependent Accounts and Coherence

In the last section, I argued for investigating my Tuesday case, *Friendship*, as an adaptation of Parfit's original case. I will be investigating *Friendship* using a particular attitude-dependent account of value, Christine Korsgaard's (esp. 1996c, 2009, forthcoming). This is because of the way it allows us to consider different aspects of coherence. In this section, I explain the importance of the notion of coherence to attitude-dependent accounts. I explain why I will consider different aspects of coherence in turn. I then define the three aspects of coherence that I will be working with.

Street gives three examples of attitude-dependent accounts of value (2009: 274, fn.10): her own Humean constructivism (2008, 2010), Korsgaard's Kantian constructivism, and Bernard Williams' account in his essay on internal versus external reasons (1981). Recall that, according to an attitude-dependent account, an agent has a standpoint constituted by her evaluative attitudes (2009: 274). From this, in combination with the non-normative facts, we can derive her normative reasons, that is, her reasons to perform certain actions. Street, Korsgaard and Williams each fill in the details of what an agent's standpoint is, what her evaluative attitudes are, and how the reasons are generated.

In particular, each account gives a different characterisation of what it is for an agent's evaluative attitudes to be coherent. The idea of coherence is implicit in Street's summary. A standpoint will be unified only if the evaluative attitudes that constitute it form a coherent set. Reasons will be successfully generated insofar as that standpoint is unified⁵. This is why it is

⁵ In fact, a disunified standpoint, where some evaluative attitudes contradict some others, can generate some reasons for action, as long as the evaluative attitudes

important that the “ideally coherent eccentric” is coherent: such a character does have a unified standpoint, and thus has genuine reasons for action, according to the attitude-dependent account.

An account’s characterisation of coherence will depend on its characterisation of the evaluative attitudes. In Street’s account, the relevant evaluative attitude is a “normative judgement”, or taking something to be a reason to perform an action (2008: 19-24). Evaluative attitudes form a coherent set if each instance of what one takes to be a reason to perform an action is consistent with all the others.

In Korsgaard’s account, the main evaluative attitude is towards actions, and is the judgement that they are right or justified (1996c: 97-8; 2009: 68-73; forthcoming: 37, fn.28). The agent makes this judgement by acting as described by a maxim that one wills as a universal law. These evaluative attitudes form a coherent set if the agent can indeed live by the complete set of laws.

Williams does not aim to offer a full-fledged metaethical theory in his essay. However, he does tell us that the relevant evaluative attitudes can be, as well as desires, “dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects” (1981: 105)⁶. These generate a deliberative question over what to do, where that question is answered by some action (1981: 103). Williams talks of a subjective motivational set made

relating to those actions form a coherent subset. For instance, I may have a reason to drink water when I am thirsty, as long as my attitudes towards alleviating my thirst, promoting my long-term health etc., form a coherent subset. This can be true even if I my evaluative attitudes towards my career do not form a coherent subset (and thus make my overall set of attitudes is incoherent).

⁶ Permission to reproduce this extract has been granted by Cambridge University Press.

up of these evaluative attitudes. It is possible for an element of this set to give rise to a reason for some action without the set as a whole giving overall reason for that action. This occurs when deliberative reasoning using the whole set does not yield that action as the result (1981: 104). By implication, the evaluative attitudes form a coherent set if deliberative reasoning gives determinate answers about which actions to perform.

I will not be able to consider all these accounts in detail. Indeed, there are many other attitude-dependent accounts⁷. Nor will I be able to consider “the attitude-dependent account” in the abstract. I will therefore choose one account, with the aim that my analysis will yield results that may be indicative of how attitude-dependent accounts in general might handle my Tuesday case. The basis for my choice is the way the account explains coherence.

In particular, I am interested here in how an account allows the identification and examination of different aspects of coherence. In §1.1.2, I discussed Street’s treatment of her Tuesday cases. I suggested that a closer look at the case of Indy, who was consistently indifferent to his pain on Tuesdays, called into question whether he really counted as an ideally coherent eccentric. We want a picture of someone with Tuesday indifference but who is clearly coherent, so that our intuitive verdict about his values is secure. But also, considering different aspects of coherence in detail will show us what contribution they each make towards our character’s values being legitimate, and where, if anywhere, things go wrong.

⁷ Parfit’s discussion of future Tuesday indifference (as a challenge to “subjectivism about reasons”) suggests other central authors in ethics, such as Michael Smith, Harry Frankfurt and John Rawls (2011: 78-80, 96-107). These would need to be included in any comprehensive review.

We can delineate different aspects of coherence based on the questions I raised about Indy and Hortense in §1.1.2. One question was over the conflict between the actions that they are motivated to perform. Indy's hedonic dislike of the sensations involved in his unanaesthetised dental operation seems to prompt him towards action of one sort (escaping the dentist's chair). His meta-hedonic indifference motivates him towards another (ignoring the pain and continuing with the operation). Indy manages to hold himself together, give or take some writhing in the chair. Hortense is horrified on Tuesday at the prospect of her operation. This motivates her, on Tuesday, to take actions that she actively tries to circumvent on previous days. The actions Hortense performs are incompatible with each other. Indy might not be in the same situation, but his motivational states seem to guide him towards incompatible actions⁸. His set of evaluative attitudes arguably fails to give him an overall guide to action. Call this a failure of *coherence in action-guidance*.

A second question was over Indy's felt experiences. His hedonic dislike involves feelings of hatred and resistance, whilst his meta-hedonic indifference involves calmness and detached curiosity. The contrast between these experiences is partly explained by the actions they prompt (such as "resisting"), already mentioned above. But we can distinguish the causal role of an evaluative attitude (the fact that it prompts someone to action) from its phenomenological character (the fact of feeling prompted to act). In addition, the phenomenological character of these experiences is not

⁸ Cf. Korsgaard's discussion of the oligarchic character in Plato's *Republic*, who rules despotically over his repressed desires and just about manages to hang together (2009: 166). If there is something problematic about that character, perhaps there is something problematic about Indy.

exhausted by pointing to the actions that Indy is prompted to perform. There is also a difference in the valence of the feeling. On the one hand, the dental operation feels hateful. On the other, it is of no immediate consequence. If we think this contrast in Indy's feelings is problematic, we might call this a failure of *affective coherence*.

The final issue is an attempt to capture what I think is right about Street's claim that Hortense's case is ultimately about moral failure. I have raised questions over whether Indy fails to achieve coherence in action-guidance and affective coherence. These questions perhaps press more insistently for Hortense. But it might be countered that, even in Hortense's case, it is not that her action and experience break down. She does not sit around not knowing what to do, nor does she not know how to feel. It is rather that her actions and experiences over time do not fit together. She should agree with herself, as it were.

In support of this point, notice that Hortense's meta-hedonic desire on most days towards her Tuesday pain is exactly the same as Indy's meta-hedonic desire on all days towards his Tuesday pain. This weighs against the idea that Hortense is cruel when Indy is not. What does make Hortense different is that she has an anomalous meta-hedonic desire.

I might be wrong to suggest that this notion of agreeing with oneself is distinct from the aspects of coherence already outlined. Hortense's anomalous meta-hedonic desire does frustrate her usual plan of action to some extent, indeed to the extent that she needs to hire external agents to fulfil her own actions⁹. In that sense, we can describe her action as breaking

⁹ Cf. Korsgaard's discussion (2009: 185-5) of Parfit's Russian nobleman (1984: 327-8), who needs to enlist his wife to force him to fulfil his youthful ideals.

down. But the analogue of this notion of agreement in the interpersonal case is a clearer example. For instance, we might think that a pair of friends shares a standpoint, constituted by their shared evaluative attitudes, and that this would do some explanatory work with regard to the values that each can hold. We might call this a question of *social coherence*.

I have set out three aspects of coherence for us to consider, coherence in action-guidance, affective coherence and social coherence. I argued that it would be helpful to consider each aspect of coherence carefully. In the next section, I will explain how Korsgaard allows us to add aspects of coherence to our investigation one at a time.

1.3 Korsgaard's Story of Agency

In this section, I outline Korsgaard's attitude-dependent account of value, telling it as a story of the role played by judgements about rightness and goodness in human agency. Both in this exposition, and in my application of Korsgaard's account to our Tuesday case *Friendship* through the thesis, I will add one dimension of the story at a time. First I will give the core story of how agents construct rules to guide their actions. Then I will add the affective dimension, which concerns evaluative experiences and judgements about goodness. Finally I will add the social dimension, extending the account to group agents. These dimensions correspond to the three notions of coherence suggested in §1.2.

For Korsgaard, evaluative attitudes are judgements about the rightness of actions and the goodness of objects (1996c: 97-8; 2008: 33; 2009: 68-73; forthcoming: 37). Judgements about rightness are made through the activity of constructing rules to govern our actions, and judgements about goodness

through the activity of having evaluative (affectively-laden) experiences of the objects in our environment.

Korsgaard tells a story of the role of those activities in human agency¹⁰. As we will see, they are in fact aspects of the same process. To aid exposition here, and analysis through the thesis, I will start by telling the story of how we construct rules to guide our actions (exposition in §1.3.1; analysis in §2.1). This introduces the notion of coherence in action-guidance. Then I will add the dimension of the story which involves evaluative experiences of objects, which I call the affective dimension (exposition in §1.3.2; analysis in §2.2). This introduces the notion of affective coherence. Agency can be understood from yet another angle by adding the social dimension of the story (exposition in §1.3.3; analysis in chapter 3). For Korsgaard, the conditions for being an agent can be satisfied by a group just as for an individual (1996b: 372-3; 2009: 197). So the individual story extends to the group. But by symmetry, the story of how a group achieves cooperation among its members can help us understand how an individual resolves the competition between her impulses towards different actions (2009: 134-5; 153-8). The social dimension of the story introduces the notion of social coherence. Thus, we will be able to apply one notion of coherence at a time to our case *Friendship*.

In the rest of this section, I will outline Korsgaard's story of agency. My analysis in subsequent chapters will involve some extrapolation from her account to fill in details, which I will leave until then.

¹⁰ Roughly, the three dimensions of the story I outline in this paragraph correspond respectively to chs. 2-5, chs. 6-8 and chs. 9-10 of (2009).

1.3.1 The core story of action-guidance

Let us begin with the core story about making judgements about rightness. For Korsgaard, every agent inescapably finds herself in the first-person, agential perspective (2009: 1). From this perspective, potential actions are presented to her (1996c: 93-4; 2009: 105, 115-121). Unlike the other animals, a human agent is conscious that she is inclined to act in the ways her mind proposes to her.

This sets her a question: should she act on these inclinations or not? The inclination presents the action as worth performing (2009: 111). This is equivalent to her being presented with an idea of an agent: in particular, the idea that the agent who performs such an action is one it is worth being (1996c: 101; 2009: 126). The equivalence can be made explicit by reference to rules. A potential rule would prescribe the action presented to her (2009: 105-6). But this rule – or rather, the maxim that can be willed as rule – also describes the sort of agent she could be by performing that action¹¹. So she in fact faces three equivalent questions: whether or not to act on an inclination, whether to be a certain kind of agent, and whether to endorse the corresponding rules. The task of deciding about each of these three questions is one and the same task (2009: 125-6).

Korsgaard calls this activity the construction of “practical identities” (1996c: 101). A practical identity is a conception the agent has of herself, a description of what she does according to which what she does is worthwhile. Her practical identity must describe, at least provisionally, her

¹¹ Here I use the term “rule” rather than “principle” as in the original text.

Korsgaard uses the latter term in various ways, and a full discussion here would not be relevant, as we will not need the term in our analysis.

actions across the course of her life and in all possible situations in which she might find herself (2009: 73-4, 198, 202). Her practical identity is likely to be composed of various particular practical identities, corresponding to particular roles and relationships. So she must also ensure that her practical identity is coherent, such that the particular roles are consistent with each other (2009: 126). Since a role can be equated with the rules which govern its performance, the achievement of a coherent practical identity is equivalent to constructing a consistent set of rules to guide her actions over her life.

I have presented the core of Korsgaard's story of human agency: the construction of a practical identity which applies to the agent's life as a whole. We can restate this using the terminology of evaluative attitudes. By constructing a rule which prescribes a certain action A, the agent comes to see herself as having reason to A (1996c: 93, 97-8). Thus, she comes to have a certain evaluative attitude towards action A, namely, seeing A as right, or justified. If she has constructed a consistent set of rules, then her judgements about rightness are consistent with each other. Her seeing one of her actions as justified is consistent with her seeing all of her other actions as justified. This means that she has a coherent guide to action for her life. Thus, having a consistent set of this type of evaluative attitude – seeing an action as justified – is the same as having achieved coherence in action-guidance.

I will give some more detail about coherence in action-guidance, and how Korsgaard's version of it applies to our case *Friendship*, in §2.1.

1.3.2 The affective dimension

I will now turn to the affective dimension of the story of agency, and how judgements about goodness relate to judgements about rightness. The core story I told above involves conscious experience, because the agent's

experience is involved in the decision-making that guides her action. However, Korsgaard's account of how that experience comes about is more detailed than that yet given. It can be considered in its own right, as Korsgaard has increasingly done in recent work (e.g. forthcoming).

Humans need to make decisions to act, but they share a more general form of life with the other animals, characterised by conscious experience (2009: 127-9; forthcoming: 19). Animals, like all living beings, maintain themselves, but they are different in needing conscious experience in order to do it. Korsgaard, like Aristotle, holds that all living beings maintain and reproduce their forms (2009: 35; forthcoming: 18). An object's form is the way it is arranged which allows it to carry out its function, its characteristic activity (2009: 27; forthcoming: 17). A living being's characteristic activity is to maintain itself as that kind of being and create new instances of that kind. Animals maintain and reproduce their forms partly through action (2009: 93-4; forthcoming: 19). Animals' self-maintaining and self-reproducing processes count as action because they involve consciousness. This contrasts with plants. In particular, animals use their power of perception to carry out these processes.

The core story of human agency in §1.3.1 began with the agent's awareness of potential actions. In common with other animals, the human agent perceives objects in her environment¹² which would contribute to her self-maintaining¹³ processes if she acts towards them in the right way (2009: 94, 104; forthcoming: 19-20). For instance, she sees food she could eat; she also sees threatening creatures or situations she should avoid. Unlike the other

¹² We can also conceive of objects which would contribute to our maintenance and seek them out (2009: 94), but I will restrict myself to the case of perception here.

¹³ Henceforth I will refer only to self-maintenance and not reproduction.

animals, the human agent is aware of these actions as potential actions for her.

But it is not just that she has an awareness of the object on the one hand, and an awareness of the potential action on the other. The two are linked. Her experience of the object includes the sense that the action is the correct one: she feels called to act, and feels that the object is to-be-acted-towards in that way (to-be-eaten, or to-be-avoided) (2009: 110-1). There is an established connection between the perception of the object and the prompting to act, which Korsgaard sometimes refers to as an instinct (2009: 111, 113-4). For Korsgaard, our “instincts”, our stock of established connections, can be extended through learning. The term “instinct” captures the idea of automatic responsiveness, rather than genetic origin. Korsgaard also uses the term “valuing capacity” to mean the capacity to find a perceived object engaging in some way: pleasant, interesting, enchanting, satisfying, or stimulating (forthcoming: 10). Since she describes such experiential responses as prompting action (forthcoming: 19-20), I take “valuing capacities” and “instincts” to refer to the same thing, a psychological mechanism linking perception of useful objects to the prompting to act towards them in ways that would maintain the agent.

But having such experiential responses towards the objects is also what it is to value those objects (forthcoming: 19). These experiential responses are our evaluative experiences of those objects. Some of our evaluative experiences make the object attractive and some repulsive – the experiences are welcome or averse (forthcoming: 20), where this is explained by the operation of the valuing capacities mentioned above. For an object to be good is for it to be experienced as attractive in this way, but also for it actually to contribute to the agent’s maintenance (forthcoming: 7-8, 33, 35).

Similarly, for an object to be bad is for it to be experienced as repulsive, and for it to damage the agent's maintenance.

Korsgaard explains judgements about goodness in terms of such experiences (forthcoming: 37, inc. fn.28). When I make a judgement that an object X is good for an agent Y, I am sharing Y's evaluative experience of the bearing of X on her own functioning. Thus, judgements about goodness are essentially sympathetic. I feel the same way about X as Y does. The function of making such a judgement is to share such feelings, rather than to describe the way the world is. Similarly, I can judge that some object X is good for *me*. This involves the limiting case of sympathy: here, I am sharing the evaluative experience with myself. Korsgaard describes this function, the sharing of evaluative experience, as "underneath" and partly prompting the lawmaking function of judgements about rightness of action. I take this to mean that judgements about goodness can hold independently of judgements about rightness, and are in some sense "prior" to them. Korsgaard labels her theory of the good "constructivist". I take this to mean that agents construct their own goods in virtue of having such evaluative experiences about objects, again independently and prior to constructing rightness (via rule-construction).

I have set out the affective dimension of Korsgaard's story of agency. Perception plays a central role in action, because agents' experiences of the objects that maintain them are part of the psychological explanation of how action comes about. For something to be good for the agent is partly a matter of the agent having such experiences, that is, of the agent seeing objects as good. Judgements about goodness are a matter of such experiences of objects as good, and it is through such experiences that the

agent constructs what is good for her. But how might this dimension of the story bring in a notion of affective coherence?

Korsgaard's account suggests the following possibility: someone would fail to show affective coherence if she fails to share judgements about goodness – that is, experiences of objects as good (and bad) – with herself. Remember Hortense, who is indifferent to the agony of an unanaesthetised dental operation on a future Tuesday but horrified by the prospect of it on the day itself. Her (anticipatory) experiences of the dental operation might be described as incoherent. Hortense seems to fail to share “the consciousness...of the bearing that events have upon [her nature]” (forthcoming: 37, fn.28)¹⁴. In other words, she fails to sympathise with herself. This echoes Street's characterisation of her as lacking mercy and compassion. We might say that Hortense fails to construct her own good and fails to achieve affective coherence.

This critique might extend to Indy. Indy was indifferent to agony not just on future Tuesdays but on the Tuesdays themselves. However, Street's illustration involved a “physical” self in revolt against the sensations of the operation whilst Indy's “core emotional” or “innermost” self remained calm. We might say here that Indy's innermost self fails to sympathise with his physical self. So he too might fail to achieve affective coherence.

I will discuss the definition of our notion of affective coherence further in §2.2, where I will apply it to our Tuesday case, *Friendship*.

¹⁴ Permission to reproduce this extract has been granted by Christine Korsgaard.

1.3.3 The social dimension

I will now turn to the final dimension of the story of agency, the social dimension. The story I have told so far is about the actions and experiences of an individual agent. But Korsgaard tells a parallel story about a group of people acting as a unified agent. The account of group agency therefore builds on the account of individual agency. But just as the account of an agent's phenomenology added a new dimension to our understanding of the agent's decision-making, the account of group agency will add a new dimension to our understanding of the individual agent.

In Korsgaard's framework, a group of people can be an agent just as an individual can. The agential perspective is first-personal, but there is a first-person plural as well as a first person-singular, and a group can take a first-person plural perspective (2009: 197). Groups can devise plans to coordinate their actions just as individuals can, and so can also have identities (1996b: 372-3). They make decisions about action by deliberating together, together choosing a rule to govern their actions (2009: 190). It is the fact that they choose a rule together which explains how they have a unified will and that their action counts as shared. This unifies the group, forming one agent out of a collectivity of individuals. This parallels the individual unifying herself, forming one agent out of a mere heap of impulses to act (2009: 67-8). She does this by choosing a rule (2009: 68-9). This constitutes her will: because she is acting in accordance with her representation of a rule (2009: 68), her movement counts as an action that we can attribute to her, the agent (2009: 100-3). The same applies to the group.

Korsgaard's main concern is to justify universal morality, that is, how we should act towards all humans. Therefore, some of her claims about group agency apply more appropriately to the group of all humans (1996c: 120-1;

2009: 209), rather than to particular groups that we can choose to belong to (2009: 200). In chapter 3, I will do the interpretative work necessary for our analysis. Here I will summarise Korsgaard's points, and merely state the additional elements of the picture of particular group agents¹⁵ that I take to be implied by her framework.

As an agent, the group agent must achieve coherence in action-guidance just as the individual agent must. An individual agent's self-conception describes her actions across her life. A group agent's shared self-conception describes the actions of its members, insofar as they interact with each other¹⁶. An individual agent's self-conception describes actions for all possible situations; her rule commits her to the same action whenever the same circumstances apply. A group agent's self-conception describes actions for all its members; its rule commits each of them to the same action if they find themselves in the same circumstances (2009: 191)¹⁷. An individual agent's self-conception must involve consistency between the various particular roles she fulfils; a group agent's self-conception must involve consistency between the roles its members fulfil, since the success of their interaction depends on the success of each of the members' actions (2009: 192).

¹⁵ Although Korsgaard does not use the term, I use "group agent" to mean a group of people that acts as a unified agent.

¹⁶ Cf. how a pair describing itself as a married couple defines the actions they perform together (2009: 187) and how the overall human story has complementary roles for individual humans (2009: 212).

¹⁷ I will apply this below to the notion of universalizability. I characterise universalizability in Korsgaard's framework as ranging over two dimensions: an individual agent's life, and all members of society (cf. 2009: 72-4)

Similarly, if it turns out we can build a notion of affective coherence for an individual agent, we may be able to do the same for a group agent. I suggested that affective coherence might mean success in constructing the good for oneself, done by sharing evaluative experiences with oneself. This locution is more naturally applied to a group, who would presumably construct a shared good by sharing evaluative experiences with each other. In §3.2, I will also discuss the idea of shared valuing capacities. In the individual case, this role was played by psychological mechanisms. I will suggest that the group analogue is a certain kind of social practice.

I have outlined the three dimensions of Korsgaard's story of human agency. The core story involves the individual agent's construction of a practical identity, corresponding to a consistent set of rules prescribing actions over her life. This is what it is to achieve coherence in action-guidance, and I will apply this part of the account to *Friendship* in §2.1. The affective dimension of the story involves the agent's evaluative experiences, underpinned by her psychology. I suggested a possible notion of affective coherence, but we will need to explore this further in §2.2. The social dimension of the story comprises group-agent analogues to parts of the account outlined in the first two dimensions. This suggests how a group agent might achieve coherence, which should explain what it is for an individual member of a group to achieve social coherence. We will explore this in chapter 3.

I have explained my choice of Korsgaard's attitude-dependent account as the example I will use to examine my Tuesday case, *Friendship*. Specifically, I will use it to test the attitude-dependent account's claim that an agent's coherence is enough to ensure that her values are legitimate and her actions justified. Given the interrelation between the three dimensions of the story of agency, it may be difficult to separate the corresponding notions of

coherence completely. Nevertheless, Korsgaard's account is relatively amenable to a progressive analysis introducing one of these notions at a time. We should bear in mind though that the results from our analysis might not necessarily generalise to all attitude-dependent accounts. My hope is that they will be indicative.

As I add dimensions to Korsgaard's story, I will check how *Friendship* needs to be adjusted (if it can be) to maintain the coherence of its central character, Theodore. As I progressively build the picture of Theodore through the thesis, I will check intuitions to see whether, and how, each aspect of coherence contributes to making Theodore's values seem more acceptable. My aim is that, by the end, we should have considered our intuitions carefully, and will be able to give a firm verdict on the complete picture.

Chapter 2: The Ideally Coherent Eccentric's Individual Agency

If someone who ignores someone on Tuesdays whom he usually values, can he really be coherent? If we were not convinced by the coherence of Indy and Hortense in Street's cases of Tuesday indifference, we might be sceptical about the prospects for Theodore. To achieve coherence in action-guidance, he will have to meet strong Kantian requirements, willing maxims as universal laws and acting with complete consistency. If there is such a thing as affective coherence, it will somehow be related to the idea of experiences playing a central role in an organic, Aristotelian-inspired notion of self-maintenance.

In this chapter, I will argue that Theodore achieves both types of coherence without a great deal of change to our initial picture of him. In our discussion of Korsgaard's core story of agency and action-guidance, the need for the agent to unify herself yields two conditions, universality over her life and consistency, which Theodore can satisfy without changing his behaviour as described in *Friendship*. I argue for this in §2.1.

The affective dimension of the story raises the question of whether Theodore can be like us in his evaluative experiences and feelings, or whether affective coherence will demand that he is quite different. But first we will need to work out our notion of affective coherence, which I do by using Korsgaard's account of judgements about goodness. Theodore achieves affective coherence, but at the cost of substituting his friend's feelings for his friend as the object of his valuing. I explain this in §2.2.

2.1 Coherence in Action-Guidance

In this section, I will argue that the core story of the agent's decision-making, a story of how she achieves coherence in how she guides her actions, can quite straightforwardly accommodate Theodore's behaviour as described in *Friendship*. This requirement does nothing to rule him out as a coherent agent with legitimate values, someone whose actions are genuinely justified. The main purposes of this section are to specify Theodore's rule for action, for use in analysis over the rest of the thesis; and to address what may seem obvious objections that his behaviour does not match the Kantian spirit of an account like Korsgaard's.

In §1.3.1, I summarised Korsgaard as saying that an agent achieves coherence in action-guidance if her rules, or equivalently her practical identity, provide her with a coherent guide to action for her life. I mentioned three features: an agent's rules must prescribe actions across the course of her life, for all possible situations, and must prescribe rules that are consistent with each other.

To aid the clarity of our discussion, I will introduce two terms to collect these three features. Firstly, an agent's rules must satisfy the condition of *universality across her life*. This will capture the first two features. Secondly, an agent's rules must satisfy the condition of *consistency*. (In chapter 3, I will introduce the corresponding conditions for group agents, namely *universality across group's membership* and *interpersonal consistency*.) This provides our test: Theodore achieves coherence in action-guidance if and only if his rules satisfy these two conditions.

Let us fill in some detail of how the universality condition works. An agent has a maxim which describes her action (2009: 11). The maxim is a thought

of the form, "I will perform act A₁ in order to perform act A₂" (2009: 68-9)¹⁸. The agent's action counts as an action because her maxim determines what she does, as opposed to some external or internal force causing her to move. A maxim that describes a justified action must have the form of universal law (cf. 2009: 15). A universal law can be followed in any kind of case, by specifying the circumstances in which the action should be performed (2009: 73, 180)¹⁹. In any kind of case, the specification of the circumstances provides a determinate answer about whether the action should be performed. But it is important to note that Korsgaard allows that the law can be provisionally universal instead of absolutely universal: a provisionally universal law allows that there may be good reasons, not yet specified in the maxim, why the action should not be performed as described (2009: 73-4). Thus, provisionally universal laws effectively come with an unstated "everything else equal" clause. This makes the seemingly strong claim that an agent's laws apply across her life (to "all of the conscious inhabitants of her body, present and future", 2009: 198; cf. 1996b: 371-2)²⁰ more plausible. A law that purports to apply in all relevantly similar circumstances can be updated as one proceeds through life. Korsgaard even states that, strictly speaking, we could will a new maxim on each new occasion, as long as the occasion was specified as relevantly

¹⁸ Although Korsgaard elsewhere gives the more standard form, "I will perform act A for the sake of end E" (2009: 11), the form I am using will make our analysis more straightforward.

¹⁹ In line with Kant, Korsgaard interprets universalizability as meaning the law should be followed by any agent who finds herself in the circumstances specified (see also §1.3.3, footnote 17). I will discuss this point further, in the context of universalising over membership of a group, in §3.1.

²⁰ Permission to reproduce this extract has been granted by Oxford University Press.

different, and this would count as universal willing (2009: 73). Thus, a maxim can be highly specific yet have the form of a universal law.

Now consider Theodore's actions as described in *Friendship*. On days other than Tuesdays, if Susan is sad, he asks her about it, listens carefully, and acts to help solve her problems if he can. On Tuesdays, if Susan is sad, he just smiles sympathetically as he would with any other member of society. On any day, if he notices a problem that he anticipates will lead to Susan being sad on a day other than Tuesday, he takes action as a friend. But if he only anticipates that it will lead to Susan being sad on a Tuesday, he ignores it (unless it endangers her). The maxim that describes his action might be as follows:

"I will, whenever Susan is sad or is likely to be sad on a day other than a Tuesday, ask her what is wrong, listen carefully to her problems, and take action if I can solve them, in order to restore her mood except on Tuesdays."

This is a thought of the form "I will perform act A_1 in order to perform act A_2 ". It may seem strange to think of it as having the form of a universal law. But it does specify circumstances in which the action should be performed, in such a way that Theodore can determine what to do in any case he finds himself. It is just that those circumstances are themselves strange: they include all occurrences of a problem making Susan sad or risking making Susan sad, except for those that are making her sad or will make her sad on a Tuesday. Theodore can keep following this rule as his life progresses.

It therefore seems that Theodore's rule satisfies the condition of universality across his life. It might be objected that this maxim is unlikely to be the one that Theodore is sincerely using to describe his action. This is an objection

that Fabian Freyenhagen discusses when considering the related problem of supposed “empty formalism” for Kantian ethics, using an example of Eric Cantona insincerely specifying a maxim of universal form that supposedly prescribes he kick a named Crystal Palace fan on a certain day in 1995 (2011: 110)²¹. Similarly, Onora O’Neill dismisses, as an irrelevant objection to Kantian ethics, the claim that one could formulate *some* maxim to justify any act (1989: 86-7). If Theodore were using the above maxim as an excuse to avoid dealing with Susan on a particular Tuesday out of selfishness, say, then a similar objection could be used to excuse Korsgaard’s framework from having to explain *Friendship*. But we are interested in the case where Theodore is sincere. This is in the spirit of Parfit’s original case of future Tuesday indifference regarding one’s own agony and Street’s development of that case. So this objection is not relevant to us.

In a discussion of his original Tuesday case, Parfit considers and dismisses the objection that his character’s desires rely on an arbitrary distinction (Parfit 2011: 79). We can translate that objection, and Parfit’s response, into our terms. An opponent of Tuesday cases might object that a maxim cannot make arbitrary distinctions and still satisfy universality. But the distinction between Tuesdays and other days is arbitrary, on an attitude-dependent account, only if there is no attitude of Theodore’s on which to ground the distinction. Theodore has such an attitude, namely, seeing his action as justified. We cannot appeal to any attitude-independent facts about Tuesdays to prove him wrong.

²¹ Tuesday cases can be seen as a special case of Hegel’s objection that Kantian ethics suffers from empty formalism: the claim would be that Kantian ethics lacks the substantive content necessary to exclude Tuesday cases.

I have argued that Theodore's rule satisfies the condition of universality across his life. Now let us consider the second condition, consistency.

Korsgaard tells us that the task of choosing rules is a task of unifying ourselves to be the cause of our ends, which is how we count as agents (2009: 68-9). We can think of the condition of universality as ensuring that the agent is unified in one sense, being bound together across her life. The condition of consistency refers to unification in a different sense, overcoming the conflict that our different roles might lead us into. When discussing what happens when we are faced with decisions, Korsgaard reminds us that we may have different roles and relationships, suggesting various actions (2009: 126). In some cases, we cannot perform all the actions that are suggested. To act, we must choose. Our roles and relationships offer various conceptions of our practical identities. But these conceptions do not cohere into one practical identity. They do not offer one set of rules that we could follow, because, as mentioned, it is impossible to perform all the actions suggested. I therefore define a consistent set of rules as one that does prescribe a set of actions that we could follow.

I should note that Korsgaard does not make the distinction between universality and consistency that I do here. But we will find it helpful for our analysis, and will also apply it in the case of group agency. The two conditions offer two distinct tests for an agent's rule. A set of rules might be consistent and prescribe one set of actions that it is possible to follow. But their maxims might not all have the form of a universal law: they might not all describe an action to be performed in all relevantly similar circumstances

across a life²². Similarly, a rule that satisfies the universality test might not be consistent with the agent's other rules.

Let us consider Theodore again. Remember, the maxim that describes his action is:

"I will, whenever Susan is sad or is likely to be sad on a day other than a Tuesday, ask her what is wrong, listen carefully to her problems, and take action if I can solve them, in order to restore her mood except on Tuesdays."

If Theodore wills this maxim as the rule to govern his action, could that rule be consistent with his other rules? It might seem that such a rule would surely clash with Theodore's other rules. After all, Theodore is Susan's friend, so he will be acting towards Susan in other friend-like ways on Tuesdays, and on other days in ways which might have results on Tuesdays. Surely some of these actions would be inconsistent with his ignoring her sadness on Tuesdays. If this is so, Theodore's rule does not pass the consistency test, and there is no need for Korsgaard's account to explain his behaviour.

I respond to this objection as follows. Note that there are friendships of a less involved sort. Imagine Susan has a friend – call him Ted – who shares common interests with her, enjoys engaging in shared activities, and so on, but is not into that fluffy emotional stuff. He does not want to have to deal with any of her negative emotions. Now notice that his interaction with

²² Strictly speaking, this would imply that the set of rules could only be followed in a subset of occasions across the agent's life. But they would still be consistent when applied within that subset of occasions.

Susan may have the effect of reducing her negative emotions. For instance, by having an interesting conversation with her, he may happen to reduce her temporary misanthropy. But he does not speak with her in order to reduce her misanthropy. So the following facts are consistent: he has no rule prescribing that he act in order to reduce her negative emotions; and his rules prescribe actions that happen to reduce her negative emotions.

So our first point is this. A set of rules can prescribe actions that bring about an unintended, though not necessarily unforeseen, effect (reducing negative emotions). It may lack any rules prescribing actions that intentionally bring about such effects. Yet it is still consistent. This point is relevant, because it is not that Theodore has a rule prescribing he act to ensure Susan stay sad on Tuesdays, it is that he has no rule prescribing that he act to reduce her sadness on Tuesdays. (Compare Indy, in Street's example of indifference to one's pain on Tuesdays, who does not have a desire to seek out pain on Tuesdays.)

Now let us suppose that Ted adds rules prescribing that he acts to reduce all of Susan's negative emotions except sadness. He does not mind that fluffy emotional stuff after all, but he does not like dealing with sadness. Now it might be, again, that these actions have the unintentional, though not unforeseen, effect of reducing Susan's sadness. Again, there is no inconsistency. Now Ted adds a rule prescribing he act to reduce Susan's sadness on Wednesdays. If this has the unintentional effect that Susan's sadness on other days is reduced, again there is no inconsistency. We can proceed in this manner until Ted has such rules for the other days of the week, except Tuesdays. There is still no inconsistency. But of course, now Ted turns out to be Theodore.

I have argued that Theodore's actions towards Susan pass the consistency test, and the test of universality across his life. He therefore has a set of rules which can guide his actions. He has achieved coherence in action-guidance. The core component of Korsgaard's story of agency cannot therefore exclude Theodore as a possible case of an agent who has genuine values, and whose actions are justified. It is possible for someone to act in the way that Theodore does. The next question is then whether someone could act this way yet feel coherently, if there is indeed such a thing as feeling coherently. In the next section, we will add the affective dimension to Korsgaard's story of agency, and consider how the notion of affective coherence might aid our analysis.

2.2 Affective Coherence

In the last section, I set out a maxim describing Theodore's Tuesday-excepting behaviour, which he could will as a universal rule and as part of a consistent set of rules. We will now investigate whether there is a requirement on how he experiences his interactions with Susan, which he would need to satisfy in order to be a coherent agent with Tuesday indifference. In §1.3.2, I raised the possibility that affective incoherence might have something to do with the failure to share evaluative experiences with oneself. What this notion might mean is unclear. One thing we do not want to include in it is the kind of conflict between our emotions and desires that we feel regularly. In §2.2.1, I define a notion of affective coherence. I start by unpacking Korsgaard's account of judgements about goodness to suggest that they have the potential to produce beliefs. I then use a case of everyday affective conflict to relate these beliefs to judgements about rightness. In §2.2.2, I apply this notion of affective coherence to Theodore. I conclude that Theodore must have emotions and desires in line with actions

in order to be affectively coherent, such that he does not feel worried or pained by Susan's sadness on Tuesdays. On the other hand, this discrepancy with how he normally feels about Susan's sadness does not make him incoherent.

2.2.1 Defining a notion of affective coherence

Recall that, in Korsgaard's framework, a judgement about goodness is an evaluative experience (forthcoming: 37). More specifically, if an agent judges that X is good for her, then she "shares" a positive evaluative experience of X with herself. Korsgaard tells us that the function of such judgements is not to describe the world. Instead their function is to sympathise with oneself, and thus construct a view of what is good for one.

However, I now advance the claim that these "judgements" have at least a potential cognitive component, given how Korsgaard defines them, and given other features of her account of goodness. She tells us that when we, and the other animals, experience something as good, we are experiencing its contribution to our self-maintenance (forthcoming: 35-6)²³. Such an experience does not imply a thought or belief about the object, if it is an experience that other animals can have. However, we humans are reflective, and are therefore capable of becoming aware *that* we are experiencing something as contributing to our self-maintenance. Through our reflection on our evaluative experience we can gain a belief.

²³ Korsgaard uses the term "well-functioning" in this passage, but remember that Korsgaard takes the Aristotelian position that the function of a living being is to maintain itself (2009: 27; forthcoming: 17).

Let us consider an example. Ursula judges that a sandwich is good. That is to say, she has a positive evaluative experience of the sandwich. We would count this as her judgement about its being good. Her evaluative experience involves perceiving it and having welcome experiential responses to it. For instance, its look, smell and taste are pleasant to her. As Korsgaard tells us, this is Ursula's way of experiencing its contribution to her self-maintenance. Her experiential responses are such that they prompt her to eat it, which would maintain her in her practical identity as a healthy person. She can do this without having any explicit thought about how the sandwich maintains her. However, since Ursula is capable of reflection, she is certainly capable of having such a thought.

I will now use my claim, that judgements about goodness can produce beliefs, in order to develop a notion of affective coherence. In §1.3.2, I suggested that affective incoherence might involve a failure to share one's evaluative experiences with oneself. However, we found from considering Street's Tuesday cases (in §1.1.2) that it was hard to pin down what exactly this means. For instance, Street's character Indy seems to experience affective conflict. His "physical self" has feelings of hatred and resistance during the agony of his dental operation on a Tuesday. At the same time, he somehow (at his "core") feels detached and curious. Perhaps this counts as a failure to share evaluative experiences with oneself – for Indy, a failure to share between his physical and core selves. But we should not be too hasty to conclude this, because having conflicting emotions is an everyday experience. We do not want to say that every affective conflict involves affective incoherence.

Let us set up Ursula as a paradigm case of everyday affective conflict. Suppose she has an obligation to meet an imminent work deadline, but is

tempted to eat a sandwich instead. She will have some feelings pulling her towards her work. She experiences the work as attractive, either because she finds it intrinsically enjoyable or interesting, or because she values some extrinsic reward like money or praise. But she also experiences what Korsgaard calls necessitation, the psychological force that reminds her of her obligation (2009: 3), and which would have emotional concomitants in case of her breaking it, such as guilt, regret, repentance and remorse (1996c: 104).

I will now set up a notion of affective coherence that admits Ursula. To do this, I will bring in Korsgaard's notion of judgements about rightness. If Ursula has an obligation to work to meet her deadlines, then she judges that it is right to work. Remember from §1.3.1 that she makes judgements about rightness in virtue of having constructed a rule. That rule might be expressed by the maxim:

"I will work, in order to meet my deadlines"

This is equivalent to her having this action as part of her practical identity. In this case, suppose Ursula includes in her practical identity a conception of herself as a diligent worker. She defines working in order to meet deadlines as part of what it is to be a diligent worker²⁴.

As mentioned, alongside her rule, Ursula also has feelings pulling her towards her work. Equivalently, she has a positive evaluative experience of her work. This implies that she experiences her work as contributing to her self-maintenance. In this specific example, she experiences working, in

²⁴ "Defining" a self-conception here can mean receiving a self-conception from society and endorsing it.

order to meet deadlines, as contributing to maintaining her in her identity as a diligent worker.

Now suppose Ursula were to reflect on her evaluative experience of meeting her work deadline. That is, she reflects on these feelings of satisfaction, and also the feeling of “necessitation”. Through reflection, she can gain (or recall) the belief that working in order to meet her deadline contributes to maintaining her in her identity as a diligent worker. And this belief is correct. She has endorsed the practical identity of being a diligent worker, and in particular, the rule she has endorsed prescribes working in order to meet deadlines. Given her actual practical identity, implemented in her decisions and actions, it is true that working in order to meet her deadline contributes to maintaining her in her identity as a diligent worker.

Now consider her feelings towards the sandwich. I stated earlier that Ursula experiences it as contributing to her self-maintenance. More precisely, Ursula experiences eating sandwiches when she is hungry, in order to assuage her hunger, as contributing to maintaining her in her identity as a healthy person. Her feelings towards this still uneaten sandwich are an extension of that experience, because she anticipates that experience with respect to the uneaten sandwich.

As before, she can reflect on this experience and acquire the corresponding belief. And this belief would be correct too, assuming being a healthy person is part of her practical identity, and that she has a corresponding rule expressed by the maxim:

“I will eat sandwiches when I am hungry, in order to assuage my hunger”

We have established that the beliefs that Ursula could acquire through reflecting on her evaluative experiences would be correct, given her rules. To put this differently, the beliefs corresponding to her judgements about goodness are confirmed by her judgements about rightness. Where does her affective conflict fit into this picture?

Ursula is in conflict because her feelings are prompting her both to work *now* to meet her deadline, and to eat the sandwich *now*. These actions are incompatible. But there is an important difference between these actions. To be a diligent worker, Ursula must work now to meet her deadline. Otherwise she will not meet it. So we should restate her maxim:

“I will work when my deadlines are imminent, in order to meet my deadlines”

It will make things clearer here if we give the corresponding judgement about wrongness, not the judgement about rightness. Ursula judges that it is wrong for her to stop working when her deadlines are imminent. Putting this in terms of her practical identity again, stopping working when her deadlines are imminent would inhibit her maintaining herself in her identity as a diligent worker.

Similarly, her positive evaluative experience of her work will have its negative counterpart, her negative evaluative experience of not working. That is, she experiences not working when her deadlines are imminent as inhibiting her maintaining herself in her identity as a diligent worker. Once again, if she reflects on this negative experience, she will acquire a belief, and that belief is confirmed by her judgement about wrongness.

By contrast, it is not necessary for her to eat the sandwich now to continue being a healthy person. That is, there is not a more specific restatement of her maxim, describing her eating sandwiches when deadlines are imminent. But recall from §2.1 that Korsgaard sees all maxims as having a suppressed “other things equal” clause. So Ursula’s maxim is really:

“I will eat sandwiches when I am hungry (other things equal), in order to assuage my hunger”

When deadlines are imminent, other things are not equal. This is why Ursula does not judge that it is wrong not to eat the sandwich, even though she is hungry. If she did not have this “other things equal” clause, her rule would be “absolutely universal”, and would prescribe eating sandwiches on every occasion she felt hungry, regardless of other circumstances. Fortunately, that is not the case. In terms of her practical identity, not eating sandwiches on every occasion she feels hungry does not inhibit her maintaining herself in her identity as a healthy person.

Her positive evaluative experience of sandwiches is as before. Its negative counterpart is the negative evaluative experience of not eating sandwiches when she is hungry. She experiences not eating sandwiches when she is hungry as inhibiting her maintaining herself in her identity as a healthy person. When she reflects on this experience, she acquires the belief that not eating sandwiches when she is hungry inhibits her maintaining herself in her identity as a healthy person.

How does this relate to her judgements about rightness? We might expect that there is some conflict with the conclusions we just drew from her “other things equal” sandwich-eating rule. But the only implication of that rule was that she *lacks* a judgement about the wrongness of not eating a sandwich

on this occasion. So there is no conflict here. Ursula does have a general, “other things equal” judgement that it is right to eat sandwiches when she is hungry. The corresponding judgement about wrongness is that it is wrong²⁵ *never* to eat sandwiches when she is hungry. It is this judgement that confirms her belief mentioned above, about not eating sandwiches inhibiting her maintaining herself as a healthy person.

Through my discussion of Ursula, I have implicitly defined a notion of affective coherence which does not rule out affective conflict. Let us make it explicit. An agent is affectively coherent when the beliefs that would be produced by reflection on her judgements about goodness are confirmed by her judgements about rightness.

Let us also give a more general characterisation of why Ursula’s case of affective conflict is not a case of affective incoherence. Ursula’s case involves a general belief about the contribution of sandwich-eating to her self-maintenance, a belief produced by reflection on her judgements about goodness. This general belief is correct, because it is confirmed by her general judgement about the rightness of sandwich-eating. That general belief does not commit her to believing anything about this specific occasion. Since (a) she lacks a specific judgement about rightness confirming that she *should* eat the sandwich on this occasion, (b) she has a specific judgement about rightness confirming that she should work, and (c) working and eating are incompatible actions, the conclusion is that she should not eat the sandwich on this occasion. But this does not mean that she has a false belief.

²⁵ In case this sounds odd, remember that we are not talking about moral wrongness here, but rather about what it takes for Ursula to stay true to her identity. Also note that the “other things equal” clause could apply to other food being available, such that she can sometimes choose other food over sandwiches.

This notion of affective coherence does not involve sharing evaluative experiences with oneself in quite the way we expected. Rather, the sharing is between judgements about goodness and judgements about rightness. Let us now apply our worked-out notion to *Friendship*.

2.2.2 Application to Friendship

Let us start by supposing that Theodore's feelings towards Susan are the same as most people's feelings towards their friends. It is painful and worrying for him to perceive Susan's sadness or likely sadness. He feels the desire to act, in the ways summarised in the exposition of *Friendship*: to ask her what is wrong and to try to help. He has these kinds of responses with respect to Tuesdays as for any other day.

Theodore's feelings constitute a certain judgement about the goodness of Susan for him. He experiences helping her when she is sad or likely to be sad, in order to restore her mood, as contributing to maintaining himself in his identity as her friend. If he reflects on his experience, he would acquire the belief that helping Susan when she is sad, in order to restore her mood, contributes to maintaining himself in his identity as her friend. Like Ursula and the sandwich, Theodore's is a general belief.

But remember from §2.1 that Theodore's maxim is:

"I will, whenever Susan is sad or is likely to be sad on a day other than a Tuesday, ask her what is wrong, listen carefully to her problems, and take action if I can solve them, in order to restore her mood except on Tuesdays."

This expresses Theodore's rule, which corresponds to the practical identity "Susan's friend" that he has endorsed. According to his actual practical identity, only helping Susan when she is sad or likely to be sad on a day other than a Tuesday, in order to restore her mood except on Tuesdays, contributes to maintaining him in his identity as Susan's friend. So his general belief is not confirmed by his judgement about rightness. Theodore is affectively incoherent.

Notice that there is a misleading structural similarity between Ursula's and Theodore's situations. It is wrong for Ursula to eat the sandwich now. It is not required for Theodore to help Susan on Tuesdays. Ursula has a general feeling in favour of eating sandwiches. Theodore has a general feeling in favour of helping Susan. So what is the difference?

Although it is wrong for Ursula to eat the sandwich now, her judgement that it is good to eat sandwiches is borne out by her judgement about the rightness of eating sandwiches. Neither of those judgements make reference to time, because they are both general. What makes it wrong to eat the sandwich now is that it is required for her to work now. We can spell this out in terms of *pro tanto* reasons. Ursula has a *pro tanto* reason to eat the sandwich, but it is outweighed by her reason to work.

By contrast, Theodore's judgement that it is good to help Susan is not borne out by his judgement about the rightness of helping Susan. His judgement about rightness has a day restriction, so that he is not required to help her on Tuesdays, even though his evaluative experience is such that that is how he experiences things. Theodore does not even have a *pro tanto* reason to help Susan on a Tuesday. Even if he had nothing else to do, he would not be required to help her. If he did go ahead and act on his emotions and desires,

his decision would be arbitrary from the point of view of his practical identity.

Theodore, under our current description, is affectively incoherent, according to a reasonable interpretation of Korsgaard's account. If we want to keep our Tuesday case to assess that account, we will have to change our description of Theodore. Let us therefore assume that Theodore is not like most people as regards his feelings towards his friend. If he notices her sadness on a Tuesday, or notices problems that will lead to her sadness on a Tuesday, he does not have emotions and desires pulling him towards helping her.

It might be objected here that, whilst we have eliminated affective incoherence as we have defined it so far, we have simply introduced affective incoherence of another sort. Don't Theodore's evaluative experiences of Susan conflict with each other, in the same kind of way we described Hortense's as doing in §1.3.2? There, we noted that Hortense is indifferent to the agony of an unanaesthetised dental operation on a future Tuesday, but horrified by the prospect of it when the day arrives. She seemed to fail to share evaluative experiences with herself between days. Is this not precisely what Theodore is doing?

This objection assumes that the object of Theodore's evaluative experience is indeed Susan²⁶. Suppose that Theodore takes Susan's sadness itself as the object which he is acting towards, where such action maintains him as the kind of agent that he is. Suppose also that Theodore sees "Susan's sadness

²⁶ There is another way of addressing this objection: Theodore could think of "Susan except on Tuesdays" and "Susan on Tuesdays" as two separate objects. See the discussion at the end of §3.3.

except on Tuesdays” and “Susan’s sadness on Tuesdays” as different objects. He experiences the former as bad, whilst he has no experiential response to the latter. Then we cannot say that his evaluative experiences conflict, because there is no common object over which they could be conflicting.

Is it up to Theodore to define the object of value like this? One person who will have something to say about this is Susan. This suggests that we should move on to the social dimension of Korsgaard’s story of agency. The affective dimension will reappear in §3.2.

We have considered how Theodore as an individual agent can display Tuesday indifference to his friend Susan. I set up two conditions for an individual to achieve coherence in action-guidance, the conditions of universality across his life and of consistency. I suggested a maxim describing Theodore’s action which would allow him to satisfy both those conditions. I proposed a notion of affective coherence which does not rule out everyday affective conflict. According to this notion, Theodore would be affectively coherent if he lacked emotions and desires pushing him towards addressing Susan’s sadness on Tuesdays. But we are dealing with a case of friendship. The social dimension of the story of agency will therefore likely change the way we understand the two notions of coherence we have introduced so far. In the next chapter, we will see how our analysis changes.

Chapter 3: Interacting with the Ideally Coherent Eccentric

What happens when the ideally coherent eccentric enters society? Up until now, we have been analysing how Theodore treats Susan as an object of value. The question was how he acted and felt towards her, given his Tuesday indifference. It seems that Tuesday indifference does not stop an individual agent from having values. But the fact that Susan is herself an agent, and shares activity and experience with Theodore in their friendship, may change our case considerably.

In this chapter, I show that the social dimension of agency changes our analysis, but the final results are quite similar. The constraints of cooperating with Susan as one of a pair of friends do not bind on Theodore very tightly. The two new conditions, universality across group membership and interpersonal consistency, mean that Theodore must understand his actions differently and to some extent adjust them. However, he does not lose the ability to make his Tuesday exception, as I show in §3.1.

Adding a social dimension to the notion of affective coherence does bring about a major change. Here I construct a component to play the analogous role to valuing capacities, the psychological underpinning of evaluative experience, in the individual case. I draw on David Velleman's (2000) account of shared intention, to propose that group agents need shared practices of representing objects to each other and eliciting shared responses. These enable group agents to have shared evaluative experiences. But these shared evaluative experiences presuppose sameness in individual evaluative experiences, as I argue in §3.2

When we apply this notion of social affective coherence to Susan and Theodore, we introduce this strong condition of sameness in individual evaluative experiences to their friendship. This could block Theodore from his Tuesday indifference. But it also allows the possibility that Susan and Theodore both display Tuesday indifference. At this endpoint of the analysis, where both agents act coherently, have coherent evaluative experiences, interact coherently and have a coherent shared evaluative outlook, we seem to have exhausted the possibilities of ruling out our Tuesday case. But can there be anything wrong now, when everything lines up so well? I suggest that the root of our intuitive resistance is an ethical commitment to the integrity of the individual, at the end of §3.3.

3.1 Social Coherence in Action-Guidance

In this section, I will argue that Korsgaard's account of how groups guide their shared actions does not necessarily exclude Theodore's Tuesday-exception behaviour. I will explain two conditions on a group agent's rules implied by Korsgaard's account, which I term *universality across the group's membership* and *interpersonal consistency* (§3.1.1). These conditions have implications for the rules a member of a group can choose for her own actions. I will then consider whether this restriction on the individual group member's rules stops Theodore from deciding to ignore Susan's sadness on Tuesdays (§3.1.2). I will conclude that we can set up *Friendship* so that the restriction does not affect him, as long as Susan accommodates his behaviour.

3.1.1 The universality and interpersonal consistency conditions

In §2.1, I interpreted Korsgaard's account of individual rule-construction as implying two conditions on the individual agent's rules. These were

universality across a life and consistency. I used these as the test of an individual agent's coherence in action-guidance. As mentioned in §1.3.3, Korsgaard gives parallel accounts of group agency and individual agency. I therefore define two conditions on the group agent's rules parallel to those just mentioned. These are *universality across group membership* and *interpersonal consistency*. By analogy, these are the test of a group agent's coherence in action-guidance. The implications of these conditions for an individual member of the group will give a test of her "social coherence", or more specifically, the social dimension of her coherence in action-guidance.

I will first explain universality across group membership. As mentioned in §2.1, according to Korsgaard, an individual agent's action is justified only if her maxim has the form of a universal law (2009: 15). An agent's rule is universalizable if she wills that *any* agent who finds herself in the specified circumstances should act as her rule prescribes (2009: 74, 191). On my interpretation, universalizability is implied by the demands of coherence in action-guidance. The need for coherence in action-guidance over the agent's own life implies that she should act as her rule prescribes whenever *she* finds herself in the specified circumstances. Whose need for coherence in action-guidance is served by requiring that she will that every other agent also act in line with her rule? It must be the group of all agents, understood as itself a group agent.

My interpretative claim here is supported by Korsgaard's claim that each individual's practical identity should be endorsable from the point of view of the human identity, so that her story fits into the larger human story (1996c: 121; 2009: 212). Korsgaard makes a related claim that we invite others (i.e. all other humans) to share the values we create (1996a: 290; 2009: 209). This is how she argues for morality (1996c: 123; 2009: 206). But in our

Tuesday case, we are interested in the rules and values shared by a pair of people, not by all humans. Since Korsgaard's focus is to explain universal morality, her discussion of interaction tends not to distinguish cooperation within a group one *may* join (2009: 200) clearly from cooperation within the group of all human agents, which one *must* be part of (1996c: 120-1; 2009: 209)²⁷. I want to keep the two separate, and to focus on the first. I will therefore need to adapt her claims about interaction.

Korsgaard claims that the individual agent's rule should be universalizable across all agents. We are interested in interaction within a particular group. So we can adapt her claim to state that the individual agent's rule should be universalizable across all the members of the group, if the rule applies to an action relevant to the group²⁸.

Where does this claim come from? The universalizability requirement on the individual corresponds to a requirement on the group agent she is a member of. As an analogue, consider the case of individual agency. There the individual needs a rule applying to all possible circumstances, so that she is unified as one agent, rather than being a mere heap of impulses to act in various ways (2009: 67-8). Analogously, in the case of group agency, the group needs a rule applying to all its members so that it is unified as one

²⁷ Korsgaard at one point notes that thinking of ourselves as sharing an identity with the rest of humanity may only have come about during the Enlightenment (1996c: 117). Thus, though we must *now* belong to this group agent of all reflective agents, this may be a historical contingency.

²⁸ I restrict to "actions relevant to the group" because not everything a group member does is under the jurisdiction of the group. For instance, in areas of my life which have nothing to do with my friend, I do not need to consider my friend in how I decide my actions.

group agent, rather than being a collectivity of individuals acting in various ways²⁹. My interpretation here is confirmed by Korsgaard's claim that it is the shared choice of a rule that constitutes the group's unified will (2009: 190). If the group has such a rule for its shared action³⁰, then this implies that the individual members will have rules for their individual actions which are universalizable in the way that Korsgaard describes.

Take Korsgaard's example of someone wanting to climb a mountain and being helped by someone else (2009: 198). We can presume that these two are friends, and they have a general maxim describing their shared action to help each other, such as:

"We will help each other realise our ambitions, in order to promote a shared sense of achievement"

Then the corresponding individual maxim for each of them might be:

"If my friend's ambition is to climb a mountain, I will drop him off at the start of the trail, in order to help him realise his ambition"³¹

²⁹ Cf. Helm's distinction between a collectivity of agents pursuing a goal, and a group agent, a distinction he characterises by appealing to the presence of a "shared evaluative perspective" in the latter case (2010: 262-4).

³⁰ I will refer to actions performed by the group agent as "shared actions"

³¹ Throughout this chapter, I will assume that the purpose mentioned in the individual's maxim is to help bring about the shared act mentioned at the beginning of the group's shared maxim. If the individual's maxim has the form "I will A_I in order to A_J" and the group's has the form "I will A_G in order to A_H", then the individual's A_J-ing is her contribution to the group's A_G-ing.

Suppose you are the one with the ambition to climb a mountain. Then you can have your own rule prescribing that you do the actual climbing; that action is not shared, so you do not need my help with that. However, to achieve our shared action, I will act according to the individual rule for helping, and drop you off at the start of the trail. The individual rule for helping does not apply to you at this moment, because you do not find yourself in circumstances where you have a friend with the ambition to climb a mountain. But it is still universalizable, because if you were in such circumstances, and I had the ambition, you would help me.

I have argued that Korsgaard's claim about the universalizability of an agent's rule across all agents can be adapted to be a claim about the universalizability of her rule across all members of her group. I argued that this claim implies a prior claim about the group agent's shared rule. That claim is that the shared rule applies to all its members; that is, it passes the test of universality across group membership. Let us now turn to our second condition on the group agent's rule, interpersonal consistency.

In §2.1, I explained that the consistency test for an individual agent's rules relates to the need to overcome the conflict of her roles and relationships pulling her in different directions. These roles may demand actions which cannot all be performed. Choosing roles that do not conflict is equivalent to choosing a consistent set of rules. For the group agent, the issue is that group members' actions pull in different directions. If they each perform those individual actions, the group will not achieve its shared action.

Consider Korsgaard's example of a teacher and student trying to schedule an appointment (2009: 192). If they choose a time which is impossible for one of them, they will fail to meet. Again, let us spell this out in terms of

their shared and individual maxims. Their task is to choose a shared action, described by a maxim of the form:

“We will meet at time X, in order to discuss our subject”

That shared action implies individual actions for each of them, described by maxims of the form:

“I will arrive at the meeting-point at time X, in order to facilitate our meeting at X”

Suppose the teacher proposes 3pm, which is impossible for the student. This is because the student has some other rule prescribing that she be somewhere else at 3pm. If she and her teacher agree on 3pm, they will have a shared rule prescribing they meet at 3pm, which would imply an individual rule prescribing she arrive at the meeting-point at 3pm. This individual rule would conflict with her existing rule. So her individual set of rules would fail the consistency test. Assuming she does not change her existing rule, she cannot accept the proposed shared rule.

The shared rule fails what I am calling the interpersonal consistency test. Given the group members' existing set of rules, it implies additional individual rules for them that one or more of them cannot follow. Thus, they fail as a group to find a shared action, and fail to unify themselves into a group agent. There would be two ways to avoid this failure. One would be to find a different shared rule. The other would be for the group members to adapt their set of existing rules.

3.1.2 Application to Friendship

I have set out the two conditions to test a group agent's coherence in action-guidance. Let us now apply these to *Friendship*.

If Susan and Theodore are friends, then they constitute a group agent. They carry out a number of shared actions together. In the original statement of *Friendship*, Theodore's unproblematic responses to Susan (the way he addressed her sadness on days other than Tuesdays) were described as individual actions: he asks her what is wrong, he listens to her problems, he takes action if he can solve them. But these individual actions are only performed as part of shared actions: he asks her questions and listens to her responses as part of a discussion, and (at least in a friendship) he only takes action given some kind of agreement or shared understanding that he will do so³². Similarly, but at a more general level, his responding to any particular instance of her sadness is (at least in a friendship) only performed as part of a shared enterprise of addressing the problems underlying that sadness together with her. In an established friendship at least, one friend (Susan) is not surprised to find the other (Theodore) responding. She understands that *they* do this.

Remember that Theodore's individual action was described by the maxim:

"I will, whenever Susan is sad or is likely to be sad on a day other than a Tuesday, ask her what is wrong, listen carefully to her

³² Of course, it would be possible for these to be purely individual actions. He asks her what is wrong; she ignores him. She tells someone else her problems; he walks up and listens. He unilaterally acts to solve them. But this is not friendship, and is not what was implied by the original description of the scenario.

problems, and take action if I can solve them, in order to restore her mood except on Tuesdays.”

It might seem that, for this to be universalizable, the friendship would have to be such that Susan had a symmetric individual maxim describing her responses to Theodore’s sadness. Since Susan does not have Tuesday indifference, we could then rule out Theodore’s rule as non-universalizable.

However, this assumes that their corresponding shared rule is something like:

“We will restore each other’s mood except on Tuesdays, in order to promote our general wellbeing except on Tuesdays”

But a project shared by a pair of friends need not be symmetric like this. Perhaps Theodore needs no support from Susan to address his sadness, because he gains other things from their friendship like engaging with common interests. Or perhaps Susan feels it important to address Theodore’s sadness whenever it arises, regardless of the fact that he does not fully reciprocate. This is not entirely implausible. If the friendship is otherwise valuable enough to Susan, she may be willing to put up with this asymmetry.

That is, we can put the question of Theodore’s sadness to one side. As long as a shared rule specifies the conditions under which any group member should perform the prescribed action, it fulfils the condition of universality across group membership. The project of managing Susan’s sadness could be specified in such a way that this is true.

So we can focus on the shared project of managing Susan's sadness. Given the description of their shared actions above, their shared maxim would be something like:

"We will, whenever Susan is sad or is likely to be sad on a day other than a Tuesday, discuss her problems, and enact a shared plan of actions to solve them, in order to restore her mood except on Tuesdays"

The shared rule would imply individual rules for Theodore and for Susan. Theodore's rule now needs rewording. Let us try the following maxims for Theodore:

"I will, whenever Susan is sad or is likely to be sad on a day other than a Tuesday, ask her what is wrong, listen carefully to her problems, and take the agreed actions to solve them, in order to facilitate our discussion of her problems, and to fulfil our enactment of our shared plan of actions to solve them"

and for Susan:

"I will, whenever I am sad or am likely to be sad on a day other than a Tuesday, respond to Theodore's questions about what is wrong, and take the agreed actions to solve them, in order to facilitate our discussion of my problems, and to fulfil our enactment of our shared plan of actions to solve them"

As stated, these individual maxims are not universalizable. It is not clear how Susan could follow Theodore's rule if she were in the "same" circumstances as him, and vice versa. The issue is that Susan cannot ask

herself about her problems, and Theodore cannot respond to his own questions. Also, some of the pronouns and proper names used make it impossible for the other friend to be in the same circumstances.

But we can resolve this with some tidying. Consider the following maxim:

“I will, whenever Susan is sad or is likely to be sad on a day other than a Tuesday, learn what I need to and provide any necessary information about her problems, and take the agreed actions to solve them, in order to facilitate our discussion of her problems, and to fulfil our enactment of our shared plan of actions to solve them”

Now, given that Theodore will tend to be in circumstances where he needs to learn about Susan’s problems, he will tend to do the learning. In practice, he will learn through asking and listening. But if Susan *were* in the same circumstances, she could do some learning about herself. (In practice, this might involve an internal dialogue, but it might involve other methods.)

Similarly, Susan will tend to be in circumstances where she has the necessary information, so she will do the information providing. But Theodore could well be in those circumstances, if he notices something about Susan that she does not, and then he could provide her with information.

It seems then, that the pair of friends could define a shared rule which would allow Theodore to carry out his Tuesday-exception behaviour. The condition of universality across group membership does not rule it out. Let us turn to the interpersonal consistency condition.

The shared rule we set out above implies individual rules for each of Theodore and Susan. The condition of interpersonal consistency demands that, for each group member, the individual rule implied fits into a consistent set with her other rules. The question here is whether Susan can accept the rule that has been implied for her.

Presumably, Susan's sadness matters to her on any day of the week. She will therefore be performing actions to address it, including her sadness or potential sadness on Tuesdays. Will she be unable to do this if she accepts the rule above? There seems no reason why she should be. She will simply treat her project of managing her sadness on Tuesdays as something that Theodore is uninterested in. Of course, to her, there is no distinct project of managing Tuesday-sadness as such. But we can put it like this: Theodore discusses and helps with some aspects of her sadness management, but not all. After all, this is true of most friendships, where some matters will seem too private, or too trivial, to discuss. If this does not cause a problem in those friendships, it should not in Susan's and Theodore's.

It might seem though that Theodore's cutoff when it comes to Susan's sadness will inhibit their discussion and resolution of her problems. But remember that a similar issue came up in §2.1, and we addressed it. Let us replicate that reasoning here. Theodore's participation in the discussion and resolution of Susan's problems might have, to him, the unintentional but not unforeseen effect of reducing her sadness on Tuesdays. This is not an infringement of his rules. He does not have a rule prohibiting his doing anything to reduce her sadness on Tuesdays. His Tuesday indifference only consists in the fact that he has no rule prescribing that he do anything. So Theodore is free to take part in these discussions and shared plans. But if

Susan tries to get him to talk about or help with something specifically to do with Tuesday, he will not be forthcoming.

I have argued that Theodore and Susan can set up a shared rule which satisfies the condition of interpersonal consistency, and the condition of universality across group membership, and which allows him to engage in Tuesday-exception behaviour. These are aspects of their shared decision-making process which ensure that, as a group agent, they have coherence in action-guidance. They can have a coherent set of shared actions, and Theodore can still ignore Susan's sadness on Tuesdays. However, we have not yet introduced the affective dimension of the story of group agency. I turn to this in the next section.

3.2 The Argument for Affective Social Coherence

In the previous section, I explained those aspects of the social dimension of Korsgaard's story of agency which focus on guiding shared action. From that part of the story alone, it seemed legitimate to conclude that Susan and Theodore could define their shared rules and shared actions so that Theodore could ignore Susan's sadness on Tuesdays. In this section, I make an argument to fill in the affective component of the social dimension. I build on Korsgaard's notions of valuing capacities, evaluative awareness and judgements about goodness, using David Velleman's account of shared intention (2000) as an analogy to suggest how her notions might translate to the case of group agency. The key notion will be that of a social practice of representing objects that relies on shared experiential responses to those objects. I will then apply this component to *Friendship* in the following section.

Korsgaard does not give a full account of how groups make judgements about what is good for *them*³³, and how they have shared evaluative experiences of objects, although various claims she makes are indicative. These include her discussion of our experience of others' pain (1996c: 148-9, 152-3); the deep psychological roots (1996c: 140-1) and inescapability (2009: 1996c: 144; 2009: 202) of our susceptibility to others' reasons; and the origin of all value in human psychology (2009: 209). However, these comments are quite scattered. Additionally, apart from a few references to the nature of love (2009: 202) and how values might relate to particular relationships and communities (1996a: 290-1), her comments are less relevant to the case of particular group agents. I will therefore not give an interpretative argument from the text at this point. We have enough of the rest of Korsgaard's framework to extrapolate to the final component.

I will argue for the claim that, for a group agent to have shared judgements about what is good for them as a group, its members must have the same judgements about what is good for them as individuals, when we consider those objects which are relevant to maintaining the group. We might say that for the group agent to achieve affective coherence, each member must show "affective social coherence". I support this by the *Argument for Affective Social Coherence*. It can be summarised as follows:

ASC-1) A group agent has a coherent set of judgements about goodness only if it has shared valuing capacities

ASC-2) A group agent has shared valuing capacities only if, for all things that contribute to maintaining them as that kind of

³³ Although, as mentioned in §1.3.2, she does explain how one individual might share judgements about what is good for *another individual* with her (forthcoming: 37, fn.28).

group agent, its members have the same individual evaluative experiences

Therefore

ASC-3) A group agent has a coherent set of judgements about goodness only if, for all things that contribute to maintaining them as that kind of group agent, its members have the same individual evaluative experiences

ASC-4) Individuals have the same individual evaluative experiences towards an object if and only if they have the same judgements about how that object is good for them as individuals

Therefore

ASC-5) A group agent has a coherent set of judgements about goodness only if, for all things that contribute to maintaining them as that kind of group agent, its members have the same judgements about how those things are good for them as individuals

I will now explain the premises of the argument.

As set out in §1.3.2 and discussed in §2.2, in Korsgaard's account, an individual agent has judgements about goodness in virtue of the evaluative experiences she has of objects (forthcoming: 37). For her to judge that X is good is equivalent to her experiencing it as welcome. These evaluative experiences are the result of the operation of her valuing capacities (forthcoming: 10). A valuing capacity is a capacity to have an evaluative awareness of objects, such that the agent is called to respond in a particular

way (forthcoming: 10; 2009: 111). More precisely, for an agent to have such an evaluative awareness is for there to be a representation of the object to the agent's consciousness, a sense of the response as called for by the representation, and an established connection between the representation and this sense (2009: 111). By extension, it should be that group agents make judgements about goodness as the result of the operation of their shared valuing capacities. This would establish premise (ASC-1).

But what is a shared valuing capacity? It would have to be the source of shared evaluative awareness of objects. This might seem a troubling notion, if shared awareness implies the existence of a group mind, a position denied by authors who have discussed what it is to share an intention, such as John Searle (1990: 404) and Michael Bratman (1993: 98) (also cf. Helm 2010: 264-5). We can break this problem down by considering the components of evaluative awareness mentioned above. Are there such things as shared representations? Are there shared senses of a response being called for? Is so, how are there established connections between these? If we can provide an account along these lines, we will have shown that there is such a thing as shared evaluative awareness.

David Velleman's contribution to the shared intention literature offers a parallel that might be helpful to us (2000: 200-20). Velleman, interpreting Searle, takes an individual's intention to be a mental representation that causes her action, by representing itself as causing it. Thus, if I intend to take a walk, I represent my intention as being what settles the question of whether I go for a walk, and indeed it does.

Velleman argues that this can be extended to the group case. First, note that a representation can be public, taking the form of an utterance, an

inscription, or depiction. Second, note that representations can sometimes be acts (for instance, asserting that X) rather than states (for instance, the belief that X). His example is two people telling each other that they will go for a walk if the other will. Each person's statement is a representation of what she will do, conditional on the other doing it. The pair of statements is, collectively, a representation that causes the two people to take a walk, and does this by representing itself (the pair) as being what settles the question of whether they do. Each person's statement on its own does not play this causal role, because neither person will take a walk without the other if the other does not make the corresponding utterance.

We asked whether there were such things as shared representations. For Velleman, there are public representations, the pairs of statements just mentioned. Now Velleman's example is of two people making a decision. So it is more comparable to the group agent's decision-making process that we discussed in §3.1, that is, how a group constructs a set of rules. But we can use it as a model for how a group constructs a picture of what is good³⁴. Instead of representing actions to each other, people can represent objects in their environment. This can be done through language, but also through body language and expressive movements³⁵. If Vivek and Wei are walking and Vivek spots a lion, his tensing up with fear is not just a form of evaluative awareness of something dangerous to him (thus playing a role in his individual agency). It is also his representing the presence of the lion to Wei. Such "acts" are often involuntary, but they are still ways we represent

³⁴ In fact, if the following logic is applied to Velleman's example, we can interpret Vivek and Wei as sharing evaluative awareness of a particular kind of object, namely the act of walking itself. I will not discuss this in further detail.

³⁵ In fact the list is probably longer and would include art, rituals, dress code, etc.

things to each other. In the terms I defined above, Vivek's tensing up is a shared representation of the lion to Wei.

But Vivek does not just convey to Wei that there is a lion. He also conveys that the lion is dangerous. That is, his representation to Wei causes her to have the sense that a response is called for, namely, running away. How does this happen?

Again, let us compare with Velleman's example. There, the pair of statements has a causal impact on Vivek and Wei, getting them to go for a walk. Consider Vivek's perspective. Suppose it is he who says, "I'll go for a walk if you will", and Wei who says, "Then I will". As he hears Wei's response, Vivek experiences the call to go for a walk, which Velleman ascribes to his desire not to have spoken falsely (2000: 216), and implicitly, to social pressure (2000: 214)³⁶. We can tell a similar story for Wei. Now, it is not that there is a shared sense of a response being called for, that is, some single entity which both Vivek and Wei are related to. Rather, they each have a token experience of the same type. But they have that experience as a result of the same causal process. There is a social practice, a set of expectations and norms around the use of language to announce one's conditional intentions (cf. Velleman 2000: 217). This social practice plays the role of the established connection between Vivek's representation to Wei and each of them feeling called to respond. And this social practice is indeed one entity that they share, by participating in it.

³⁶ Of course, Vivek may also have the desire to go walking because he likes it, etc., but that was present before Wei's response. It may be true that one particular desire, the desire to go walking with Wei, becomes more salient after her response, but this complication does not detract from my point about social practices.

Once again, this model extends. There are such social practices governing linguistic expression of emotions, body language and expressive movements. These are the established connections between shared representations of an object and the participants' sense of a response being called for. Since she is well-versed in the social practice of reading body language, Wei reacts to Vivek's tensing up by feeling the call to run away, just as he does. This is how Wei and Vivek have a shared evaluative awareness, or experience, of the lion. By having this awareness, they, as a pair, judge that it is dangerous.

Now we can explain what it is for Vivek and Wei to have a shared valuing capacity. It is the mechanism that produces this shared evaluative experience. In the individual case, that mechanism was psychological, and was equivalent to the established connection between representations and one's sense of being called to respond. In the group case, again the mechanism is equivalent to the established connection. That means it is the social practice itself. Thus, Vivek and Wei's shared valuing capacity in this case is the social practice of using and understanding body language.

I have set out what shared valuing capacities are, a concept we needed for the claim that they are necessary for a group to have a coherent set of judgements about goodness, in (ASC-1). Let us turn to (ASC-2). This is the claim that a group has shared valuing capacities only if its members share their individual evaluative experiences, for all those things that contribute to maintaining them as that kind of group agent.

Consider Vivek, Wei and the lion again, and suppose the negation of this claim. Suppose Vivek and Wei participate in the social practice of using and understanding body language. But Wei, as an individual, does not

experience the lion as dangerous. That means she does not feel the call to run away on perceiving the lion. But we said earlier that Wei reads Vivek's body language in virtue of her feeling a call to run away on seeing him tense up. Then, as long as Wei sees both Vivek and the lion, she both feels the call to run away and does not feel the call to run away. This is a contradiction.

In general, the kind of social practice we are discussing operates along the following lines. One group member has an individual evaluative experience of an object. She expresses her experience in a way that other group members perceive. Their perception then causes them to have the same evaluative experience. Thus, if any of the other group members has, as an individual, a contradictory evaluative experience, the social practice breaks down for her. Thus, the group members must share their individual evaluative experiences. However, this requirement does not apply to objects not covered by the social practice. Further, since the group only needs the shared valuing capacity to maintain it as the kind of group agent that it is, it only needs that part of the social practice which covers objects that help maintain it. This establishes (ASC-2).

At this point it is worth noting that "social practices" can also include group-specific practices. Thus, a smaller group can tailor and adapt a social practice to share more specific evaluative experiences with each other. This will be relevant to our analysis of Susan and Theodore.

(ASC-1) and (ASC-2) together imply (ASC-3), the claim that a group agent has a coherent set of judgements about goodness only if its members have the same individual evaluative experiences towards all those things that contribute to maintaining them as that kind of group agent.

I have included (ASC-4) in the argument so that we can rephrase the conclusion we have reached so far. (ASC-4) tells us that individuals' evaluative experiences towards objects are the same as their judgements about the goodness of those objects for them. We know this from the exposition of the affective dimension of Korsgaard's account (forthcoming: 37), set out in §1.3.2.

This gives us the conclusion of the *Argument for Affective Social Coherence*, namely that a group agent has a coherent set of judgements about goodness only if, for all things that contribute to maintaining them as that kind of group agent, its members have the same judgements about how those things are good for them as individuals³⁷. My reasoning relied on the idea of a shared valuing capacity as a social practice governing how participants share representations of objects. In the next section, I will apply these ideas to Susan and Theodore.

3.3 A Final Picture: Shared Tuesday Indifference

In the last section, I argued that members of a group agent must share judgements about goodness, concerning those objects relevant to their maintenance as the kind of group agent that they are. As part of this argument, I filled in the detail of how a group can share experiences through representing objects to each other, and of how the shared valuing capacities which underpin that representation are related to social practices. In this section, I will apply that detailed interpretation of Korsgaard's framework to our Tuesday case, *Friendship*. Susan and Theodore must end up sharing emotions towards the problems that afflict her on Tuesdays. However, this

³⁷ Cf. with Helm's framework, where individuals participating in a group can share emotions as part of the group agent, yet have different emotions as individuals (2010: 269-70)

does not rule out Theodore's problematic Tuesday behaviour. We can end up in a scenario where both Susan and Theodore stop caring about her (supposed) problems on Tuesday. Intuitively, I will argue, this is still unacceptable. I conclude that what underlies our intuitions of unacceptability is an ethical commitment to integrity. *Friendship*, at least when analysed using Korsgaard's account, allows the friends to treat Susan as separate from her projects in their actions and experiences, or even not treat Susan as an individual at all. I relate this to our discussion of Street's characters, Hortense and Indy, in chapter 1. I finish by indicating some possible implications of this analysis for attitude-dependent accounts in general.

We know from §3.1 that Susan and Theodore have a shared project of managing Susan's sadness on Tuesdays. The maxim describing this shared action is:

"We will, whenever Susan is sad or is likely to be sad on a day other than a Tuesday, discuss her problems, and enact a shared plan of actions to solve them, in order to restore her mood except on Tuesdays"

This maxim, along with their others, characterises the kind of friends they are, since their complete set of rules is equivalent to their shared practical identity. But let us focus on this shared action alone. Susan and Theodore are the kind of friends who have this shared project. That is the kind of group agent that they are.

What is the object that Susan's and Theodore's project centres on? In §2.2, I supposed it was Susan herself, until we were forced to conclude that Theodore could only achieve affective coherence if his judgements of

goodness (or value) were about Susan's sadness. However, since we are about to discuss both Susan's and Theodore's evaluative experiences, it would be better here to refer to Susan's underlying problems. (Susan's sadness is her evaluative experience of her problems.)

As a group agent, Susan and Theodore have a coherent set of attitudes towards Susan's problems, if we mean coherence in action-guidance. They see action to resolve those problems as justified if those problems will lead to sadness for Susan on a Tuesday. This allows them to guide their actions towards them. It is true that Susan has her own individual attitudes about when her individual actions towards her problems are justified. But that is a separate matter, since she is able to get on with guiding her own actions, and there is no conflict with her shared actions with Theodore, as we concluded in §3.1.

At the end of chapter 2, I raised concerns about whether it was possible to assess Theodore's affective coherence without bringing in the social dimension of Korsgaard's story of agency. Now we have the full picture, and can reintroduce that question. Do Susan and Theodore as a group agent achieve affective coherence? Do they succeed in having a coherent set of judgements about goodness?

According to the *Argument for Affective Social Coherence*, they do not. Susan's individual judgement of the badness of her problems is that they are saddening. As we concluded in §2.2, Theodore experiences Susan's problems (sometimes indirectly, by witnessing her sadness) as painful and worrying, but only if they do not lead to her sadness on Tuesdays. So Susan and Theodore do not have the same judgements of the badness of the object, namely Susan's problems, that contributes to maintaining them as the kind

of group agent that they are, namely a pair of friends with the project of managing Susan's sadness on Tuesdays. But this is a necessary condition for them, as a group agent, to have a coherent set of judgements about goodness. They each fail to achieve affective social coherence, so they as a pair of friends fail to achieve affective coherence.

Bringing in the notion of shared valuing capacities, discussed in §3.2, shows why this is so troubling. The sharing of individual evaluative experiences is necessary for having a shared valuing capacity, and the shared valuing capacity is necessary for generating a coherent set of judgements about goodness for the group agent. As we found, the shared valuing capacity is equivalent to a social practice. Specifically, the group should share a practice of representing objects to each other, such that they prompt each other to feel the call to respond in the same way. But Theodore and Susan do not share such a practice. At least sometimes, Susan will express her reaction to an object of concern, namely one of her problems, verbally, through body language, perhaps through tone of voice...and Theodore will not share her reaction. This would be reasonable if it concerned something outside the scope of their friendship. Friends do not need to care about exactly the same things. But we have seen that Susan's and Theodore's friendship is partly characterised by how they react to her problems. Although they may have a coherent way of responding to those problems together in action, their lack of the kind of shared practice I have defined here shows a failure to share their lives, a central ethical concern (cf. Korsgaard 1996a: 275).

It seems then that Korsgaard's account implies a notion of affective social coherence that Theodore's behaviour causes him and Susan to fail to achieve. If that is so, then we can explain our intuitive opposition to the

picture of Theodore we have built up on Korsgaard's own terms. But our investigation is not over yet. We can adjust our Tuesday case yet again. Suppose Susan does share Theodore's evaluative experience of her problems. That is, she does not feel sad on Tuesdays at all. The "problems" that would have led to her feeling sad on Tuesdays no longer have that effect. In fact, we might expect her not to see them as problems any more, since she is no longer prompted by her sadness to address them. So she might adjust her individual rules for action as well.

Well, doesn't this make the picture *less* intuitively objectionable? Theodore is no longer ignoring Susan's sadness at all. He is simply going along with Susan's feelings. Or perhaps he has persuaded her to accept his devil-may-care attitude to Tuesdays. Either way, the friends are in complete harmony, and all is well.

Or is it? Again, remember that until the end of §2.2, I was assuming that the object of value was Susan herself. When I made the shift to thinking of Susan's sadness itself as the object of Theodore's concern, I pointed out that this would probably be unacceptable from Susan's perspective. Although we are now talking about the aspects of her life that affect Susan, rather than her sadness, the same point stands. Yet now Susan herself is accepting a friendship that puts the things that affect her, rather than her herself, as the focus of their shared activity and their shared evaluations.

The worry here is not about how special Susan should be. It is rather about integrity. In his critique of utilitarianism, Bernard Williams argues for the ethical importance of recognising the identification of a person with his deeply held projects and attitudes, and against an ethical theory that allows these to be valued and weighed in isolation from him (1973). If a friendship

is an ethical project for two, there is something deeply worrying about Theodore and Susan detaching the aspects of her life that supposedly matter to her (indeed let us refer to them as projects), and making these the material for their activity and feelings. This is true even if they subscribe to universal morality. Indeed, imagine that all members of society take up a Tuesday-excepting stance to each other, thus reinforcing Theodore's and Susan's definitions of value. Then we end up with a moral theory much like that which Williams is attacking³⁸.

There was in fact an alternative route for Susan and Theodore, but it leads us more quickly to the same endpoint. It was not necessary to shift from Susan as the object of value to Susan's problems. They could instead have split Susan into "Susan on Tuesdays" and "Susan except on Tuesdays", so that their shared activity and experiences were focused on the latter and not the former. Once again, this highlights the threat to integrity.

We can also see this problem emerging if we return to Street's own Tuesday cases. Indy's case of consistent indifference to his agonising dental appointment on Tuesdays resembles the first attack on integrity I mention.

³⁸ It may seem surprising that Williams, invoked by Street as the forerunner to her conception of the attitude-dependent account, should appear here as a champion of an ethical claim that an attitude-dependent account is falling foul of. This may be because my analysis of Korsgaard does not apply to other accounts, though given my final discussion of Street's cases, I am not so sure. Or it may be that Williams' is not an attitude-dependent account in quite the way that Street thinks. It may also seem surprising that a challenge from Parfit has led us to this point by Williams, given his own consequentialism and his stance on personal integrity (1984; Korsgaard 1996b). But Parfit may have his own ethical issues with integrity; see his discussion of the Consent Principle (2011: 177-211).

Just as Susan's projects are split from her, Indy splits his physical wellbeing from himself. He views it with detachment and curiosity. This may be an admirable ability which could be put to use in heroic circumstances. And it may well make sense in the context of his home planet, where physical suffering on Tuesdays is prevalent. But as a stance to take when arranging dental appointments, it is troubling. And even if Indy consents, should the *dentist* be going along with it?

Hortense's case of future indifference to agony combined with Tuesday horror, meanwhile, resembles the second attack on integrity. Just as Susan is split into multiple objects of concern, Hortense splits herself. The fact that agony is involved hides this point. Imagine instead that Hortense has some ongoing project, but only acts to ensure its fruits come to her on days of the week other than Tuesday. Here the issue would not be moral, as Street suggests, but she would still be frustrating part of herself, and failing to recognise that that part belongs to her.

It might be objected here that I have distorted our discussion by changing Street's cases so that the objects of value are Indy and Hortense themselves, people, rather than their pain. Wasn't Parfit's original point that some *things* are intrinsically valuable? But my point can generalise. If we value something, we should recognise its integrity. Otherwise we are not really valuing *it* at all.

How does this relate to the debate between attitude-dependent and attitude-independent accounts? First, I must reiterate that an analysis of one account, Korsgaard's, does not necessarily generalise to all. Given Korsgaard's prominence, these conclusions are still interesting. Korsgaard sees herself as advocating an autonomy-based ethics (1996b: 5), where integrity of the agent

is central (2009: xii-xiii). But these conceptions of autonomy and integrity are *practical*, a matter of successfully constructing the guides to action and the evaluative experience that we have discussed. This means that the only criterion is what works: if Susan and Theodore can really put their friendship into practice, and maybe get us all alongside in a wider moral theory, then there would be nothing stopping them.

But the brief remarks above about Street's Tuesday cases suggest that it may be worth exploring the extension of this critique to other attitude-dependent accounts. I can only offer a brief thought here. It is possible that accounts that prioritise attitudes will always risk threatening the integrity of the individual, because recognising the integrity of the individual is partly a matter of recognising the independence of the individual from our attitudes. Here we might return to the realist I quoted at the outset, Iris Murdoch, who suggests that "a just and loving gaze directed upon an individual reality...[is] the characteristic mark of the active moral agent" (1970: 33).

Conclusion

I have argued that Korsgaard's attitude-dependent account cannot rule out Theodore's actions in *Friendship* as justified according to its own criteria. Neither can the notions of coherence it delivers dispel our intuitive concern about this Tuesday case. However, the purpose of this thesis has not been simply to vindicate Parfit's objection to attitude-dependent accounts of value. As we saw in chapter 1, Street's response to Parfit showed that there are number of issues at play in his case of future Tuesday indifference. I have attempted to set some of these out methodically over the course of the thesis.

At the end of chapter 1, I stated that our investigation would help us see how different aspects of coherence contribute to making Theodore's actions seem more acceptable. Whilst the demands of coherence have not blocked the Tuesday case, they have offered some constraint on how Theodore's Tuesday-exception behaviour could manifest itself. Perhaps more clearly, they have constrained the evaluative attitudes available to him.

For instance, in §2.1, we saw that Theodore's judgements about rightness had to meet conditions of universality and consistency, due to his need to achieve the tasks of constructing rules and a practical identity, tasks central to Korsgaard's main story of agency. Whilst this did not affect his action towards his friend Susan, the constraints on the maxims he could use to describe that action represent constraints on how he could justify that action. This may be a useful point in thinking about how other attitude-dependent accounts incorporate the notion of coherence in action-guidance.

In §2.2, we obtained a strong result about affective coherence, that Theodore's judgements about goodness (that is, his felt evaluative

experiences like emotions and desires) had to match his judgements about rightness. Specifically, the beliefs that could be produced through reflection on his judgements about goodness had to be confirmed by his judgements about rightness. We adjusted the case to allow for this, but how easily characters like Street's Indy could meet this condition is less clear.

In §3.1, Theodore's individual actions had to be reinterpreted in light of his cooperation with Susan as part of a "group agent", that group being them as a pair of friends. We had to adjust Theodore's maxim describing his individual action, to allow for the possibility of Susan finding herself in his circumstances and having to perform his action. Thus, universalizability of his rule across the group – the friendship – was a constraint with some bite. Again, we adjusted the scenario so his behaviour could continue. But if Susan had to be able to apply his rule, it had to make sense to her, at least to some extent (even if she does not fully understand why he would have such a rule). In addition, Susan had to agree, or at least go along, with his actions.

In §3.2, we obtained perhaps our strongest result. In order for the pair to have shared evaluative experiences, and have the shared judgements about the value of things which were relevant to their friendship, both individuals had to have the same individual experiential responses to those things. This put a strong constraint on what the relationship between Susan and Theodore would have to be like to allow his Tuesday behaviour. In the end, Susan had to buy into the way that Theodore sees the world. This may support Street's suggestion, when discussing her character Hortense, that the real issue at play is a moral one.

To put this differently, a major issue with the Tuesday cases is that *we*, the readers, cannot buy into the way that these characters see the world. Hortense seems cruel, a tendency we disapprove of. Theodore makes

decisions based on Tuesdays, which makes no sense to *us*. Indeed, it may be objected that many of my claims in this thesis about intuitions are driven by the fact that I personally do not happen to agree with Theodore.

However, I have not just insisted on an intuitive verdict on the case. I have, sketchily, offered a diagnosis. I will repeat that to conclude the thesis. What Korsgaard's attitude-dependent account allows is great freedom in carving up the world, as long as certain formal constraints are met, even granting that these formal constraints are demanding in their own way. For instance, one must have universal rules, but very specific and bizarre rules can be specified as universal rules. A similar point can be made about her account of evaluative experiences: since agents can reshape their practical identities in many ways, the experiences towards objects that help maintain them in those identities are quite malleable. The tricky question then becomes whether there are things that should not be carved up. I suggested that the individual might be one, and also ventured the claim that any object of value needs to be seen as having integrity. An attitude-dependent account might be able to explain why this is the case, and one founded on the idea of individual autonomy and agency might be supposed to be especially well-equipped to do this. But the idea that value needs to be responded to and respected, which motivates Parfit's case and was mentioned at the outset of this thesis, is one that seems vulnerable in an account like Korsgaard's. The value of the individual is perhaps a particularly pressing case. Concerns which stem from this idea about value – and the Tuesday indifference cases may be an expression of this concern – need careful treatment and attention.

References

Anscombe, G.E.M. (1958), 'Modern Moral Philosophy', *The Journal of the Royal Institute of Philosophy* 33, 1-19

Bratman, Michael (1993), 'Shared Intention', *Ethics* 104, 97-113

Freyenhagen, Fabian (2011), 'Empty, Useless, and Dangerous? Recent Kantian Replies to the Empty Formalism Objection', *Bulletin of the Hegel Society of Great Britain* 63, 95-118

Helm, Bennett (2010), *Love, Friendship, & the Self: Intimacy, Identification, & the Social Nature of Persons* (Oxford: Oxford University Press)

——— (2014), 'Emotional Communities of Respect', in Christian von Scheve and Mikko Salmela (eds.), *Collective Emotions* (Oxford: Oxford University Press)

——— (forthcoming), 'Rationality, Authority, and Bindingness: An Account of Communal Norms', in David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility* (Oxford: Oxford University Press)

Korsgaard, Christine (1996a), 'The Reasons We Can Share: An Attack on the Distinction between Agent-Relative and Agent-Neutral Values', in *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press)

——— (1996b), 'Personal Identity and the Unity of Agency: A Kantian Response to Parfit', in *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press)

——— (1996c), *The Sources of Normativity* (Cambridge: Cambridge University Press)

——— (2009), *Self-Constitution* (Oxford: Oxford University Press)

——— (forthcoming), 'The Origin of the Good and Our Animal Nature', in Bastian Reichardt (ed.), *Problems of Goodness: New Essays on Metaethics* (Available at <http://www.people.fas.harvard.edu/~korsgaard/Essays.htm>, retrieved 19/12/2013)

Murdoch, Iris (1970), *The Sovereignty of Good* (London: Routledge and Kegan Paul Ltd.)

O'Neill, Onora (1989), 'Consistency in Action', in *Constructions of Reason* (Cambridge: Cambridge University Press)

Parfit, Derek (1984), *Reasons and Persons* (Oxford: Oxford University Press)

——— (2011), *On What Matters*, Vol. 1 (Oxford: Oxford University Press)

Searle, John (1990), 'Collective Intentions and Actions', in Phillip R. Cohen, Martha E. Pollack and Jerry L. Morgan (eds.), *Intentions in Communication* (Cambridge, MA: MIT Press)

Strawson, P.F. (1974), 'Freedom and Resentment', in *Freedom and Resentment and Other Essays* (London: Methuen)

Street, Sharon (2008), 'Constructivism about Reasons', in Russ Shafer-Landau (ed.) *Oxford Studies in Metaethics*, vol. 3 (Oxford: Clarendon Press)

——— (2009), 'In Defense of Future Tuesday Indifference: Ideally Coherent Eccentrics and the Contingency of What Matters', *Philosophical Issues* 19, 273-298

——— (2010), 'What is Constructivism in Ethics and Metaethics?', *Philosophy Compass* 5, 363-384

Velleman, David (2000), 'How to Share an Intention', in *The Possibility of Practical Reason* (Ann Arbor: Scholarly Publishing Office, University of Michigan Library)

Williams, Bernard (1973), 'Integrity', in J.J.C. Smart and Bernard Williams, *Utilitarianism: For and Against* (Cambridge: Cambridge University Press)

——— (1981), 'Internal and External Reasons', in *Moral Luck* (Cambridge: Cambridge University Press)

——— (1985), *Ethics and the Limits of Philosophy* (Abingdon: Routledge)