Shame is Valuable

A thesis submitted in fulfilment of

the requirements for the degree of

Master of Arts

in the

Department of Philosophy

at

Rhodes University

by

Jonathan George Euvrard

November 2009

Abstract

In this thesis I argue that shame is valuable because it contributes to the moral life by promoting coherence. I start by developing and defending a conception of shame. On my conception, rational shame involves a negative self-assessment, in which I am both the assessor and the object of assessment, and in which the standard of assessment is my own. I then develop a notion of coherence, and apply it to the relationship between values, and the relationship between values and actions. I also tie the notion of coherence to what I call “the moral life”. I then discuss two ways in which shame can work to promote coherence. Firstly, I describe a process of critical reflective self-assessment, and show how this is a particularly effective method of promoting coherence when coupled with shame. Secondly, I discuss the connection between my emotions and my values, arguing that this connection promotes coherence, and that shame works to reinforce this connection and thereby to promote coherence.

Acknowledgements

When you enjoy working with people as much as I do, writing a thesis can feel like a lonely exercise. There is simply no way to avoid seemingly endless hours of struggling alone to understand difficult readings or rewrite clumsy passages. I have been very fortunate to find friends, family and my supervisor always waiting at the other end of these efforts, ready to provide whatever support I needed, whether that was distraction or a sincere discussion of the philosophical problems that confronted me.

My family is exceptional. Living with them again has inspired in me a new appreciation for the beautiful home that they have created (and constantly recreate), the home that created me. My family provide not only limitless love and support, but also a keen and critical interest in my philosophical work, so their influence on the ideas developed here is both direct and indirect, and certainly immense.

Friends from near and far have shared my experience of writing this thesis. Their continuing interest in my research was probably my most significant source of motivation when I felt flat, and our travels all over South Africa frequently put things back in perspective. My thanks must go especially to one little girl who is already well on her way to becoming a philosophical giant.

Dr Samantha Vice is a superb supervisor. She has helped me every step of the way, from insightful discussions and suggestions at our regular meetings to careful and prompt feedback on everything from my first sketches to my final draft. My thanks must also go to other members of the Department of Philosophy for their professional and personal interest in me and my research.

Table of Contents

Introduction

This thesis can be approached from two rather different perspectives, each of which I have adopted at different phases of working on it. On the one hand, you can approach it with specific theoretical questions about shame. What is shame? How is shame related to values and what role does this relationship play in the moral life? How is shame related to the self? How is shame valuable? I have structured the chapters around these questions. My main claim is that shame is valuable because it promotes coherence and thereby contributes to the moral life. The chapters in this thesis expand on the main claim and answer the theoretical questions above in the following way. In chapter one, I develop an account of shame and defend it against some objections and alternative accounts. I shall argue that shame involves a negative *self*-assessment, in which one recognises a mismatch between one's self-conception and one's ideal self-conception. In chapter two, I present a model of values that provides the background against which I develop the notion of coherence. I tie this notion to the moral life, suggesting that the moral life can be understood as a quest for coherence. In chapter three, I describe a process of critical reflective self-assessment, and I develop a model of the self that consists of a reflective self, an ideal self-conception, and a character, while consistently stressing the unity of the self. Finally, in chapter four I argue that shame combined with critical reflective self-assessment enhances our capacity to consciously promote coherence in our lives. I also argue that shame reinforces our emotional connections to our values, and as such can work consciously and unconsciously to promote coherence.

On the other hand, you can approach this thesis with pre-theoretical questions prompted by your experience of shame. For example, you may start with certain intuitions based on your emotional reactions to your own wrongdoing. You can react to your own wrongdoing in various ways. Some of these are captured in the familiar notion of guilt: you can feel bad that you have transgressed, that you have caused harm, that you have hurt somebody, etc. These feelings centre on the wrongdoing and are directed out towards the world and those who share it with you. Appropriate responses to these feelings include similarly outward-directed actions like confession, apology and reparations. However, your emotional reaction may have another aspect that also centres on your wrongdoing but is directed inwards, towards yourself. You may feel that your wrongdoing flows from you in more than simply a causal way, that it says something about you, about the kind of person you are. You may be struck by the painful awareness that the relationship between you and your wrongdoing is not merely coincidental, that there is something wrong with a part of you that led you to act wrongly. What is this emotional reaction? What can you do with it? I shall argue that the concept of shame captures this experience. I shall describe and recommend a

process of critical reflective self-assessment as an appropriate response to shame, and I shall show how together shame and critical reflective self-assessment can play a valuable role in your life.

These two perspectives complement each other. The theoretical questions give the thesis structure, and encourage a more careful and coherent account, while the pre-theoretical questions keep me honest, always challenging me to ensure that my account speaks to the lived experience of shame. By attending to both of these perspectives I hope to provide a rigorous and robust defence of shame, and also to show how you can make sense of your own shame, how you can understand and appreciate the valuable role it plays in your own life. In other words, I aim to show how shame can be valuable to *you*, given the values you already have.

Like a fertile garden, philosophical research has a way of growing quickly out of control as each question leads to another and each answer leads to more questions. Throughout this thesis, I have had to maintain a habit of careful but sometimes drastic pruning. While addressing questions about shame, two big questions towered over me that I could not adequately address. They are "What is an emotion?" and "What is a value?"

I do not develop or subscribe to any specific theory of emotions in this thesis. I take myself to be at least partly excused for this by one of the most influential philosophers working on theories of emotion, Robert C Solomon, who writes:

> The real work will continue to be with regard to particular emotions, and often with specific regard for the particular instance of a particular sort of emotion. Thus a "theory of emotion" might be too much to expect ...[1]

Solomon echoes a widespread scepticism about the prospects for one theory of emotion that would help to make sense of all emotions, from surprise to sympathy.[2] I present my account without explicitly fitting it to any such theory of emotions. I try to be attentive to potential challenges to my theory even when these problems have roots in theories of emotions that I would anyway reject. As such, throughout this thesis I engage with philosophers on particular issues without attempting to provide the theoretical background against which their ideas were originally developed. I hope this broadens the scope of my account of shame, and makes it accessible and interesting to those who are committed to any one of a broad range of theories of emotion, and to those who are not committed to any.

However, there will be some limitations to the range of theories with which my account of shame is compatible. On my account, shame is a painful emotion that always involves a negative self-assessment. I argue that this makes shame both reason-responsive and open to rational criticism. These features do not follow as conclusions from any particular theory of emotions to

[1] Robert C Solomon, 'Emotions, Thoughts, and Feelings', p. 84.
[2] For references to other philosophers who share this scepticism, see Ronald de Sousa, 'Emotion'.

which I am committed. Rather, they reflect my intuitions about shame particularly, considered against examples, objections and rival accounts. However, such intuitions place my account within the broadly "cognitivist" set of theories of emotion. In contrast, "non-cognitivist" theories of emotion would not allow such a central role for assessment in an emotion, nor would they accommodate reason-responsiveness and liability to rational criticism in the way that I suggest. For example, a Jamesian theory of emotions would not look for an assessment or belief to identify a particular emotional episode as one of shame, but rather certain physiological responses like blushing or hiding.[3] As such, my account of shame is probably not acceptable to a person with non-cognitivist commitments.

Similarly, my use of "value" in this thesis is intended to be broad and inclusive. In arguing for the value of shame, I focus on what would commonly be understood as instrumental value: shame is valuable insofar as it promotes coherence, and coherence is valuable insofar as it is required for you to effectively promote[4] the values you already have. So my argument is that shame is valuable to you because of its relation to the values you hold. I am arguing that, if you accept my account of shame, you should see your shame as good *for you*. My argument for the instrumental value of shame is not intended to be exclusive – that is, I do not mean to exclude the possibility that shame is valuable in other ways. I argue only that shame is valuable in at least one way, because it promotes coherence, and I go on to show how it does this, and why coherence is valuable. I also develop a model of values. This is not intended as a theory of value but rather demonstrates a way to map the relationships of coherence and conflict between values. Here I use "value" in the broadest possible sense,[5] in which I may value fairness, but I may also value chocolate. Obviously, these values need not be of equal importance to me, but in the language I am using one cannot say that the latter is not important enough for my relationship to it to be called valuing. I also discuss the formation and revision of one's values, emphasising both one's own role and that of others in these processes. In talking about values so broadly, I again hope to broaden the scope of my account and make it accessible to those with any one of a broad range of theoretical backgrounds.

This thesis forms part of a resurgence of philosophical interest in emotions. As Phil Hutchinson begins his recent book, *Shame and Philosophy*, "Emotions research is 'in'."[6] Yet, despite this interest in emotions, the publisher promoted Hutchinson's book as "the first and only extended

[3] See Ronald de Sousa, 'Emotion'.
[4] Throughout this thesis I shall use the word "promote" in connection with values. I mean "promotion" in the broadest possible way, without any specifically consequentialist connotations. I shall use "abuse" in a similarly broad way to mean the opposite of "promote".
[5] This tracks dictionary definitions of value as a verb and a noun: "To consider of worth or importance" (Oxford English Dictionary) and "any object or quality desirable as a means or as an end in itself" (Dictionary.com).
[6] Phil Hutchinson, *Shame and Philosophy*, 'Preface', p. vi.

philosophical exploration of the concept of shame."[7] I hope that this thesis goes some way to show why shame deserves far more attention. It is such a complex emotion that it seems likely to catch and hold the interest of any philosopher working on emotions. And it is a powerful and puzzling emotion with which most people are both intimately familiar and yet strangely unfamiliar. Both the familiarity and the unfamiliarity stem from shame's close relationship to similarly murky but central concepts like the self. The experience and the study of shame direct our attention towards the self. This redirection is particularly obvious when shame is occasioned by wrongdoing, because we tend to follow our guilt and look outwards from the action towards the world. By turning our attention towards the self, and by highlighting the connection between elements of the self and action, both in theory and in practice, shame leads us back to ourselves. This enhances our ability to change ourselves, and thereby our future actions, and thereby the world.

[7] Synopsis on Amazon.co.uk, accessed 14 May 2008 <http://www.amazon.co.uk/gp/product/product-description/0230542719/ref=dp_proddesc_0?ie=UTF8&n=266239&s=books>.

Chapter One: What is shame?

Introduction

In this thesis, I shall argue that shame is valuable. Before I can begin this argument, I need to be clear about what experiences count as shame, and which of these experiences I shall be defending. In this chapter, I develop and argue for a conception of shame, defending it against possible objections and rival conceptions. My conception enables me to make a further distinction between what I shall call rational and irrational shame. My defence will focus primarily on rational shame. On my conception, rational shame is a painful emotion that involves a negative self-assessment, in which I am both the assessor and the object of assessment, and in which the standard of assessment is my own. However, I also discuss irrational shame, in which my assessment and my standard of assessment come apart, and in chapter four I shall suggest some ways in which even irrational shame can be valuable. I begin, now, by briefly describing some general features of the emotional experience that the concept of shame picks out.

Shame is a painful emotion that involves a negative assessment of oneself. (I shall not address vicarious shame in this thesis.) When I feel shame, I feel bad about myself. Shame, like guilt, may be occasioned by an action. I may feel bad about myself because of something I have done. To take Jean-Paul Sartre's famous example from *Being and Nothingness*:

> I have made an awkward or vulgar gesture. ... But now suddenly I raise my head. Somebody was there and has seen me. Suddenly I realise the vulgarity of my gesture, and I am ashamed.[1]

An "awkward or vulgar gesture" may not occasion very deep shame. Archetypal examples of shame usually involve serious moral wrongdoing. Such wrongdoing is often accompanied by a mixture of guilt and shame. Put simply, "[w]e feel guilty for what we do. We feel shame for what we are."[2] In Sartre's words,

> ... it is a shameful apprehension of something and this something is me. I am ashamed of what I am. Shame therefore realises an intimate relation of myself to myself. Through shame, I have discovered an aspect of my being.[3]

So even when shame is a reaction to a wrongdoing, the object of my shame is not my action but the person performing the action – me. Shame, however, is not only felt in response to actions. I may feel shame for any "aspect of my being". I might be ashamed of just being black. But a person's race,

[1] Jean-Paul Sartre, *Being and Nothingness*, p. 221.
[2] Lewis Smedes, *Shame and Grace: Healing the Shame We Don't Deserve*, p. 10. See also John Rawls, *A Theory of Justice*, p. 445 and Bernard Williams, *Shame and Necessity*, p. 88.
[3] Jean-Paul Sartre, *Being and Nothingness*, p. 221.

if it is a real feature of a person at all, is not a shameful feature, because there's nothing wrong with being black. In the face of widespread racism and other forms of oppression that employ practices of shaming, this raises a question: would black people be better off developing an immunity to shame? Questions like this challenge the main argument of this thesis, in which I shall defend the value of shame as a moral emotion. In order to do this I shall need to work with a conception of shame. My goal in this chapter is to develop this conception.

My conception of shame will be based on the work of John Rawls, Gabriele Taylor and Bernard Williams. I start by briefly sketching my conception of shame, outlining the main elements. Then I shall attend to each element in turn, drawing support from the literature and elaborating on the way I think each element should be understood. Finally, I shall defend my conception against the rival conceptions of John Deigh and Cheshire Calhoun. As we shall see, the strongest challenges to my conception raised by these authors centre on the role of audience in shame, and these discussions about the role of audience lead to discussions about shame and rationality.

I said that shame is a painful emotion that involves a negative assessment of oneself. I take it that all the authors I discuss would agree on this skeletal definition, but there is much room for debate on how to interpret and elaborate on this definition. Shame involves an assessment, so it involves an assessor, an object of assessment, and a standard of assessment against which the object is assessed. Is the assessor the agent, the audience, or only certain members of the audience? The object of assessment is clearly oneself. When I feel shame, I am the one being negatively assessed. But is this an assessment of only an aspect of myself, myself in a certain role, my moral self or my whole self? And what is the standard against which I am assessed? Is it my standard, yours, or ours? Is it a moral standard?

I shall argue that shame is a painful emotion involving a negative *self*-assessment. I am both the assessor and the object of assessment. As Sartre says, "Shame therefore realises an intimate relation of myself to myself."[4] I shall argue that the standard against which I assess myself is my own standard of assessment, which is based on my values. However, while my shame always involves my negative self-assessment, my assessment and my standard of assessment are strongly influenced by the assessments and standards of others.

In contrast, some authors make the assessments and standards of others central to shame. For example, Deigh connects shame to a sense of worth that "comes from knowing one's status or essential nature" and "reflects concern with the congruency between one's conduct or appearance and one's real worth" and this leads him to argue that "we should conceive shame ... as a reaction to ... the threat of demeaning treatment one would invite in giving the appearance of someone of

[4] Jean-Paul Sartre, *Being and Nothingness*, p. 221.

lesser worth."[5] For Deigh, then, shame does not involve my own assessment or my own standard of assessment, while the assessment of others is a necessary element of his conception.

I shall now look at each element of shame in turn: the object of assessment, the standard of assessment, types of assessment and types of shame, and the assessor.

The object of assessment

The object of shame is the self. As Sartre says, "I am ashamed of what I *am*."[6] Williams agrees: "Shame looks to what I am."[7] Rawls says, "Shame springs from a feeling of diminishment of self".[8] In Taylor's words, "a person feeling shame becomes conscious not merely of what he is doing, but becomes conscious also of his self."[9] The assessment is not limited to a particular action or aspect of the self; shame involves the whole self.[10] As Helen Merrell Lynd describes it,

> Shame is an experience that affects and is affected by the whole self... Separate, discrete acts or incidents, including those seemingly most trivial, have importance because of this moment of *self*-consciousness, the self stands revealed. Coming suddenly upon us, experiences of shame throw a flooding light on what and who we are...[11]

Unfortunately, this only adds to the complexity of shame, because it is not at all obvious what exactly the self is. Some philosophers have even denied that such a thing exists.[12] In chapter three, I shall develop a model of the self, and discuss in more detail the way in which shame makes us painfully aware of how our shameful actions flow from our ourselves – from "what and who we are". For now, I merely wish to note the personal and overwhelming nature of shame as a familiar phenomenological feature. My shame is directed very specifically and personally at me.

The standard of assessment

In the next section, I shall argue that shame involves a negative *self*-assessment; my shame is based on *my* belief that I am in some way deficient. This self-assessment implies a standard of assessment. Nevertheless, even if the self-assessment is my own, it is not obvious that the standard of assessment is therefore also my own. I could assess myself according to the standards of others. In this section, I shall argue that in rational shame the standard of assessment against which I

[5] John Deigh, 'Shame and Self-Esteem: A Critique', p. 242.
[6] Jean-Paul Sartre, *Being and Nothingness*, p. 221.
[7] Bernard Williams, *Shame and Necessity*, p. 93.
[8] John Rawls, *A Theory of Justice*, p. 445.
[9] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 59.
[10] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 85.
[11] Helen Merrell Lynd, *On Shame and the Search of Identity*, p. 49. See Lynd for many illuminating examples.
[12] For examples, see Richard Sorabji, *Self*, p. 17, where Sorabji attributes this view to Nietzsche, Hume and others.

measure myself is my own, but I shall also explicitly acknowledge the role others play in the formation of my own assessment.

To make the distinction between assessment and standards of assessment clear, consider this analogy: My friend has asked me to edit an article she has written for the technology section of a local newspaper. The editor's guidelines suggest a minimalist, technical style with careful but enthusiastic use of jargon. I personally suspect that a more relaxed and inclusive style would better entertain and inform the average reader of the newspaper, including myself. Now, I may assess the article and suggest revisions in light of either the editor's guidelines or my preferences. These represent different standards of assessments, and these two standards of assessment would lead me to make rather different assessments, and hence rather different suggestions for amendments to the article. I highlight this distinction between assessment and standards of assessment because it is useful in understanding examples of shame in which these two elements come apart. I discuss such examples below in response to Deigh and Calhoun.

On my conception of shame, my negative self-assessment is based on *my* standard of assessment. Similarly, for Rawls the agent who feels shame "has betrayed a lack of the moral excellences he prizes and to which he aspires."[13] Williams agrees that we feel "ashamed because we have contemptibly fallen short of what we might have hoped of ourselves".[14] As Taylor says, the person feeling shame "feels herself degraded, not the sort of person she believed, assumed, or hoped she was or anyway should be."[15]

One way to think about the standard of self-assessment involved in shame is to speak of an *ideal* self-conception. This ideal may not be fully articulated; I may only be aware of certain hopes or expectations I have of myself in particular situations or on reflection. Such discovery may even signal the moment in which such standards are developed or articulated. But we may think of these hopes and expectations as constituting an ideal self against which I assess myself. Many things may come together in an articulation of an ideal self-conception. Rawls, Taylor and Williams variously speak about excellences, a plan of life, values, goals, ideals, expectations and character, and I suspect that each of these terms picks out a potential element of an ideal self-conception. I shall discuss the idea of an ideal self-conception in detail when I develop a model of the self in chapter three. Here, it should be clear that my standard of self-assessment is connected to these elements that may be thought of as coming together to form an ideal self-conception. Using this we may picture shame as involving a mismatch between my self-conception and my ideal self-conception. In chapter two, I

[13] John Rawls, *A Theory of Justice*, p. 445. Rawls is speaking about moral shame. I reject the distinction between moral and natural shame below.
[14] Bernard Williams, *Shame and Necessity*, p. 92.
[15] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 64.

will show how this mismatch is a instance of incoherence closely related to incoherence in other spheres. In shame, I am strikingly aware of this incoherent relationship between who I am and who I would like to be.

Types of assessment and types of shame

Some philosophers make a distinction between different types of shame based on the types of assessment involved. Both Calhoun and Rawls distinguish "moral shame" from other types of shame. My conception of shame does not identify distinct types of shame. While I acknowledge that shame can occur in a variety of different circumstances, my conception is intended to capture all varieties of shame.

Rawls attempts to draw a distinction between natural shame and moral shame. He writes:

> natural shame is aroused by blemishes in our person, or acts and attributes indicative thereof, that manifest the loss or lack of properties that others as well as ourselves would find it rational for us to have ... it is our plan of life that determines what we feel ashamed of.[16]

> Someone is liable to moral shame when he prizes as excellences of his person those virtues that his plan of life requires and is framed to encourage. He regards the virtues ... as properties that his associates want in him and that he wants in himself.[17]

Rawls also states that "virtues are excellences".[18] So natural and moral shame is occasioned by the awareness of the loss or lack of properties in oneself that it would be rational for us and others to want us to have given our plan of life. The difference is that moral shame further requires that the properties in question have a moral quality. This maps an everyday distinction between different types of fault or failure. On this distinction, I may see my poor performance on the flute last night as a devastating failure, but not a moral failure, whereas the way I take my frustration out on my sister following the performance might strike me as a moral failure.

Williams rejects the distinction between natural and moral shame:

> One thing that a marked contrast between shame and guilt may express is the idea that it is important to distinguish between "moral" and "nonmoral" qualities. Shame itself is neutral on that distinction.[19]

Taylor argues that shame is always a moral emotion, "provided that morality ... is taken to include personal morality, a person's own view of how he ought to live and what he ought to be."[20] She further says that

[16] John Rawls, *A Theory of Justice*, p. 444.
[17] John Rawls, *A Theory of Justice*, p. 444.
[18] John Rawls, *A Theory of Justice*, p. 445.
[19] Bernard Williams, *Shame and Necessity*, p. 91.

> Shame can be seen as a moral emotion, then, not because sometimes or even often it is felt when the person believes himself to have done something morally wrong, but rather because the capacity for feeling shame is so closely related to the possession of self-respect and thereby to the agent's values.[21]

For Rawls, natural and moral shame is heavily value-laden. I agree with Taylor that when we are talking about values that relate to how I live my life, we are talking about moral values. So I take all varieties of shame to be moral, whether or not they involve an assessment that might commonly be thought of as a moral assessment. However, what is important here is simply that my conception of shame does not *require* distinctions between different types of shame. I aim to capture all shame in one conception. If my conception is successful, the onus falls on those who distinguish different types of shame to justify their distinctions by showing what extra work they do.

The assessor

Shame involves a negative self-assessment. In shame, I am the assessor, and I make a negative assessment of myself. However, the most common examples of shame involve an audience and their negative assessment of me. In Sartre's example (quoted above), it seems to be the audience's assessment that is doing the work. It is only on realising that he has been seen that the agent feels shame. So there are intuitive grounds for understanding shame as primarily a reaction to an audience's negative assessment, as acknowledged by Williams when he says, "The basic experience of shame is that of being seen, inappropriately, by the wrong people, in the wrong condition."[22] In Taylor's words, "The person feeling shame feels exposed: he thinks of himself as being seen through the eyes of another. The case of shame introduces an observer or audience."[23]

As noted earlier, Deigh makes the audience's assessment central to shame. For Deigh, when I feel shame this implies a thought about how I might appear to an audience as someone of lesser worth, and how such appearance would invite demeaning treatment of me. So the agent must at least be imagining the reaction of an audience, though an audience need not be actually present. Similarly, on Calhoun's conception,

> Moral shortcomings must first be exposed to public view before they can be the source of shame; or at the very least, the contempt that others would show us were our shortcoming exposed must be clearly imaginable.[24]

---

[20] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 76.
[21] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 84.
[22] Bernard Williams, *Shame and Necessity*, p. 78.
[23] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 57.
[24] Cheshire Calhoun, *An Apology for Moral Shame*, p. 131.

Calhoun says that the audience must be an "actual group".[25] In other words, in the absence of an audience, the audience's reaction may be imagined, but even the imaginary reaction must be attached to an actual audience. Hence, as Calhoun says, "One can feel ashamed imagining what one's beloved grandmother would have thought were she still living."[26]

One need not be confused by talk of the actual or imagined reactions of actual or imagined audiences. I take it that the point of these labels is simply to accommodate the various levels of independence authors believe is required for the assessment in shame. The imagined reaction of my dead grandmother still possesses a certain degree of independence from me because I do not necessarily project my own assessment onto her, but rather I imagine how she would have assessed me were she present. Thus Calhoun accepts my dead grandmother's imagined reaction as sufficiently independent for shame, while she rejects the imagined reaction of a purely imagined audience.[27]

Neither Williams nor Taylor takes the assessment of an actual audience to be central to shame. As Taylor points out, "One may feel shame when quite alone and knowing this to be so."[28] However, Williams takes at least the imagined assessment of an imagined audience to be necessary for shame. He says,

> Even if shame and its motivations always involve in some way or other an idea of the gaze of another, it is important that for many of its operations the imagined gaze of an imagined other will do.[29]

Williams refers to an audience in the form of an "internalised other":

> the other need not be a particular individual or, again, merely the representative of some socially identified group. The other may be identified in ethical terms. He ... is conceived as one whose reactions I would respect; equally, he is conceived as someone who would respect those same reactions if they were appropriately directed to him.[30]

As Williams acknowledges, the obvious worry here is that the internalised other is not an *other* at all, because his assessments merely reflect my own – he does not seem to have a life of his own in any interesting sense. Williams claims that it would be a mistake to reduce the internalised other to merely an "echo chamber for my solitary moral voice".[31] He says, "The internalised other is indeed abstracted and generalised and idealised, but he is potentially somebody rather than nobody,

[25] Cheshire Calhoun, *An Apology for Moral Shame*, p. 131.
[26] Cheshire Calhoun, *An Apology for Moral Shame*, p. 131.
[27] See discussion of the "Moral Pioneer" in Cheshire Calhoun, *An Apology for Moral Shame*, pp. 129-132.
[28] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 58.
[29] Bernard Williams, *Shame and Necessity*, p. 82.
[30] Bernard Williams, *Shame and Necessity*, p. 82.
[31] Bernard Williams, *Shame and Necessity*, p. 84.

somebody other than me."[32] I remain unconvinced that the internalised other is robust enough to constitute an audience. "Audience" must mean more here than simply "assessor", or it would be trivially true that any emotion of assessment involves an audience.

Taylor's strategy is to grant audience a *causal* role in shame. Once the audience's assessment has played a causal role, it is no longer important. She says,

> A person feeling shame judges herself adversely. This judgement is brought about by the realisation of how her position is or may be seen from an observer's point of view. But there is no reference to such a point of view in her final self-directed judgement.[33]

Elsewhere she says that the "notion of audience ... has a role to play in the explanation of the self-directed judgement".[34] In Sartre's example, my shame follows my sudden awareness that someone has seen me. Now, imagine that I subsequently realise that the person is blind and has not seen me. This realisation is very likely to have no effect on my shame, because the assessment I attributed to the audience has already prompted a self-assessment, and that self-assessment may persist despite the now apparent absence of an audience in the relevant sense.

On my conception, I do not take the assessment of an audience, real or imagined, to be causally or conceptually necessary for shame. This is partly because shame may be felt in the absence of, or even in spite of, the audience's assessment. Let me introduce an example that makes this clear, in which an audience is not even believed to be present. Imagine that I read an entry from my private journal which points to a persisting flaw in my character. Perhaps it reports an occasion on which I lied to get out of trouble, and reminds me of a similar situation more recently. No one else knows the truth about either incident. Now, we might be tempted say that the journal has enabled me to occupy the position of an audience, by giving me a critical distance from my actions and decisions. So let us remove the journal. Imagine now that I am simply lying in bed thinking about my day and remember the incidents and feel shame in exactly the same way. I am not considering how someone else might judge the incidents I am remembering, but I still feel crushed with shame, I blush, and I shut my eyes and fervently hope not to wake up in the morning.

If this example is plausible, it seems unnecessary to say that an audience is required to play any role in shame. Indeed, the example explicitly denies even the thought of a (real or imagined) audience's (real or imagined) assessment. It may be necessary for *me* to adopt the position of an audience, in some minimal sense, towards myself, but insofar as this is necessary for any form of self-assessment, it is not a distinctive element of our concept of shame. We must not be led astray

[32] Bernard Williams, *Shame and Necessity*, p. 82.
[33] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 68.
[34] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 64.

by the dramatic metaphors we sometimes use to describe self-assessment and self-confrontation.[35] The above example and similar cases of shame show us that an audience's reaction is not necessary for shame.

Shame may be felt not only in the absence of an audience's negative assessment, but even contrary to the audience's assessment. As Williams notes, "people can be ashamed of being admired by the wrong audience in the wrong way."[36] In *Shame and Philosophy*, Phil Hutchinson provides the example of Léopard. Sitting in a Rwandan prison for crimes committed as part of the genocide there, surrounded by comrades who feel no remorse or shame, Léopard acknowledges his crime and feels a deep shame, "despite the mockery of his comrades who have yet to, or who do not, acknowledge their crime".[37] As Hutchinson points out, Léopard's audience is constituted by his friends and comrades, and his shame is felt in spite of their assessment. His negative self-assessment is formed and maintained even though it is contrary to that of his audience. Again, this fits well with my conception, on which we need refer only to Léopard's own negative self-assessment to make sense of his shame.

On my conception, my shame always involves my own negative self-assessment. I have argued that this negative self-assessment is based on my own standard of assessment, which is based on my values. I hope to have shown how the audience's assessment and my shame may come apart by discussing how shame can occur in the absence of, and contrary to, the audience's assessment. However, I do not want to paint a solipsistic picture of shame, nor do I want to imply that audience cannot play an important role in shame. In my discussion, I have made a distinction between assessment and standards of assessment. Audience sometimes plays important roles in both of these elements.

Firstly, as I have mentioned, an audience's negative assessment can, and often does, cause shame. An audience's negative assessment can prompt me to make a negative assessment of myself and thereby cause me to feel shame. On the conception of shame I have been describing, my own standard of assessment is central, and my standard of assessment is based on my values. However, it is true that in some cases I would not have felt shame in the absence of an audience's negative assessment. There are many reasons for this. Most obviously, self-deception, habit or complacency often help me to avoid acknowledging negative aspects of myself. In the face of an audience's negative assessment, I may find myself pushed towards a new level of self-confrontation. I have said that an audience's assessment is not necessary in all cases of shame, but in many cases of shame the audience's assessment plays a very important role by prompting my own self-assessment.

[35] Phil Hutchinson, *Shame and Philosophy,* p. 143.
[36] Bernard Williams, *Shame and Necessity*, p. 82.
[37] Phil Hutchinson, *Shame and Philosophy,* p. 142.

Secondly, the audience is important in explaining my standard of assessment, and the values that ground it. My values are developed with and informed by others. The standards against which I assess myself probably differ in some respects from the standards against which others assess me. However, our standards are likely to have much more in common, because me learn them from each other. While we can and often do defend both our values and our standards when they differ from those of others, we should not underestimate the ongoing influence others have on our values and standards of assessment.

Audience can play direct and indirect roles in shame. At the moment of feeling shame, I may adopt directly from the audience both their negative assessment of me and their standard of assessment on which this is based. Indirectly, over my entire life my standards of assessment and values are developed with and informed by others. What I have tried to show is that the audience's negative assessment does not play a direct role in every case of shame and it is therefore not a necessary constituent of shame.

Rationality

A careful discussion of the rationality of emotions is far beyond the scope of this thesis. However, it will be useful to touch very briefly on a few points here. Firstly, emotions are reason-responsive. Of course, most emotions are not immediately or fully responsive to reason. In a moment of gut wrenching fear, I am unlikely to calm down immediately upon deciding to do so. However, I can probably calm myself down little by little, and over time I may even learn to control my fear in similar situations. Secondly, emotions are liable to rational criticism. We regularly assess our emotional responses to situations and circumstances. And thirdly, we try to train and develop our emotional responses in various ways.

Oddly enough, the rationality of emotions is particularly perspicuous in irrational emotions. Consider this example of irrational fear from Joe Simpson's *This Game of Ghosts*, when he recounts the experience of Yves, a seasoned mountain rescuer:

> Swaged wire cable six millimetres in diameter appears as thin as a silver silk thread when a man hangs from the end of sixty feet of it. Yves pushed from his mind any thought of it snapping. He knew the breaking strain of the wire, he knew it wouldn't snap ... *It will not snap. I have done this many times. It will not snap.* He repeated it like a mantra, vocalising it in his head, pushing the fearful thoughts away, convincing himself.[38]

I cannot discuss all the complexities of irrational emotions here. All I wish to do is point to some common features of irrational shame and other irrational emotions. Yves's fear is open to rational criticism. We call it "irrational" because fear is "the emotion of pain or uneasiness caused by the

[38] Joe Simpson, *This Game of Ghosts*, pp. 168-169.

sense of impending danger",[39] but in Yves's case the impending danger is absent and Yves knows this, he knows that the cable will not snap. A satisfactory conception of fear will attribute to the agent a belief that there is impending danger. In irrational cases, however, this belief may be absent. Similarly, on my conception of shame, I attribute to the agent a negative self-assessment based on the agent's own standard of assessment. In irrational cases, however, the self-assessment and the standard of assessment may come apart. A person experiencing irrational shame will be unable to explain what it is they are ashamed of in the light of their values. It is also important to note how Yves's understanding of his fear as irrational helps him to deal with it. He knows that fear is usually based on a sense of impending danger, and so he constantly reinforces his belief that he is not in danger, and reinforces the connection between this belief and his emotions. Yves's fear is reason-responsive to some degree; his irrational fear diminishes as he convinces himself that he is not in danger. And over time, Yves is likely to feel less irrational fear while performing rescue operations, even if he is never able to overcome it entirely.

Similarly, I may feel irrational shame, particularly in overwhelming situations. I may feel shame and adopt a negative self-assessment that does not correspond to my standard of assessment. Imagine, for example, that I feel shame occasioned by my failure at some display of macho male bonding, but on reflection I realise that macho male bonding has no place in my ideal self-conception, that I place little or no value on it. On my conception, my shame is open to rational criticism, it is reason-responsive and it can be developed and trained, which is not to say that any of these will be easy. I can recognise the mismatch between my negative self-assessment and my standard of assessment, and acknowledge that my shame is irrational. I can think about my values, about my ideal self-conception. Both of these may work to relieve my shame, and regularly responding in this way to similar episodes of shame may lead me to be less ashamed in similar situations in the future.

Examples of irrational shame take centre stage in papers by John Deigh and Cheshire Calhoun, who argue against conceptions of shame, like the one I have presented, which ground shame in the agent's negative self-assessment rather than that of the audience. One strategy employed by both of them is to discuss examples of shame in which one or more of the elements of my conception appears to be absent. So they discuss examples in which the connections between my shame, my self-assessment and my standard of assessment have come apart. I take their papers to be the most formidable challenges to my conception, so I shall spend some time engaging with them with reference to examples that are sympathetic to their arguments. I have presented a

[39] Oxford English Dictionary.

conception that describes rational shame, and I shall argue that the examples below represent cases of irrational shame, and as such do not threaten my conception.

I am aware of the danger of arguing circularly here: that the examples do not fit my conception is unproblematic because they are irrational, and they are irrational because they do not fit my conception. However, I hope to show that we are right to understand the examples as irrational for independent intuitive and pragmatic reasons. Firstly, I believe the examples do strike us intuitively (or pre-philosophically) as irrational. Deigh explicitly acknowledges this and defends his use of irrational examples.[40] Calhoun acknowledges that this intuition is shared by the authors of her examples.[41] Secondly, as I shall illustrate, understanding the shame in these examples as irrational provides the agent with the resources they need to deal with their shame in these situations. In this way I hope to address Calhoun's complaint that my conception forces us to "discount as irrational or immature much of the shame suffered by socially disesteemed populations—racial minorities, women, the poor, lesbians and gay men."[42] To call an instance of shame irrational is not dismiss it, or the person feeling it, as ridiculous. On the contrary, I hope to show that my conception provides the agent with the best resources to deal with the shame they experience.

Arnica's irrational shame

In 'Shame and Self-Esteem', John Deigh takes an example from André Gide's *Lafcadio's Adventures* of a young girl who feels shame over her humorous surname:

> Arnica Péterat [which might be translated as "Miss Fartwell"] – guileless and helpless – had never until that moment suspected that there might be anything laughable in her name; on her first day at school its ridicule came upon her as a sudden revelation; she bowed her head, like some sluggish waterweed, to the jeers that flowed over her; she turned red; she turned pale; she wept.[43]

On Deigh's conception, Arnica's shame is a reaction to her appearance as someone of lesser worth, and the demeaning treatment that such appearance invites. Arnica's self-conception, her conviction of her own real worth, is not threatened and does not diminish. Arnica's shame is connected to her awareness of her classmates' negative assessment of her directly, and not via her own self-assessment. She does not adopt their negative assessment, and she does not form a negative self-assessment. She simply becomes aware of their negative assessment and feels shame.

On my conception, Arnica's shame implies a negative self-assessment. She must feel that her own real worth is diminished. And Arnica's self-assessment is based on her standard of assessment.

[40] John Deigh, 'Shame and Self-Esteem: A Critique', p. 237.
[41] For example, by Sandra Lee Bartky. See Cheshire Calhoun, *An Apology for Moral Shame*, p. 136.
[42] Cheshire Calhoun, *An Apology for Moral Shame*, p. 135.
[43] Quoted in John Deigh, 'Shame and Self-Esteem: A Critique', p. 232.

But this makes Arnica sound silly, because her real worth has nothing to do with her surname's vulgar connotation, and she should know that. Deigh says that "one could not regard it as a fault in oneself".[44] This leads him to say that Arnica feels shame over something that lies outside of her self-conception. This is why Deigh takes this to be a direct counterexample to conceptions of shame like mine, because my conception posits a negative self-assessment, based on Arnica's standard of assessment, which appears to be absent here. So Arnica's shame does not appear to fit my conception of shame.

I could simply reject this example: Gide does not mention anywhere that it is shame that Arnica feels. She "turned red", but blushing is compatible with a number of emotions, and in this particular situation it seems more likely to indicate embarrassment than shame. More charitably, though, I can accept this as a case of shame, but insist that if it is shame that Arnica experiences then she does make a negative self-assessment. In this case, Arnica's negative assessment might have come apart from her standard of assessment. In this case, her shame is irrational, but understandable. She simply adopts her classmates' negative assessment of her without adopting their standard of assessment. Arnica is a young girl attending school for the first time. Until now, she may have been surrounded by people whose assessments and standards of assessment largely mirror her own. We may thus imagine that Arnica accepts her classmates' criticism in good faith, without considering what standards of assessment it is based on. Faced with overwhelming criticism in an already overwhelming situation, it is understandable that Arnica makes a negative self-assessment of herself and feels shame. She adopts their negative assessment of her even though, on careful reflection, she might well reject this assessment and the standards of assessment on which it rests.

However, Deigh insisted that Arnica could not make a negative self-assessment. Later, he says that

> one is not responsible for one's parentage and thus ought not to judge oneself according to facts wholly determined by it. Inasmuch as shame in these cases reflects such judgement, they exemplify experiences to which one ought not be liable.[45]

There is an important difference between what Deigh says here and what he said earlier. Here he admits that Arnica *ought* not to judge herself over a pun on her surname; earlier he stated that she *could* not. I am suggesting that Arnica *does* make a negative self-assessment, whether she ought to or not. In the light of this example, Deigh's conception has intuitive appeal because it seems obvious that Arnica ought not to tie her real worth to whether or not her surname is liable to a silly pun, and on Deigh's conception Arnica does not. However, when we consider Arnica's situation, it is hardly

[44] John Deigh, 'Shame and Self-Esteem: A Critique', p. 232.
[45] John Deigh, 'Shame and Self-Esteem: A Critique', p. 237.

surprising that Arnica's shame turns out to be irrational, because the link between Arnica's self-assessment and her standard of assessment has broken down.

As irrational examples of fear do not fit the standard conception of fear, so Arnica's shame does not fit my conception of shame neatly. Arnica's self-assessment is not based on her standard of assessment. This does not count against my conception if the example is clearly irrational and I can explain why the example deviates from my conception. I hope to have shown that such an explanation is easily available.

My conception also provides us with resources that Deigh's conception cannot provide. Imagine now that, on returning home, Arnica still feels ashamed and relates her experience to her parents. How should they respond? My conception suggests that they should talk to Arnica about her values, about what is important to her, about how a silly pun on her surname does not define her worth. And my conception explains why we expect Arnica to be less vulnerable to this particular shaming criticism as she grows up. We can picture the adult Arnica being quite unperturbed by people laughing at her name. Deigh's conception seems to suggest that Arnica should avoid future shame by hiding or changing her name. His conception does not have the resources to explain the development of Arnica's capacity for shame.

Arnica's case fits Deigh's picture of shame very neatly: Arnica feels shame because her classmates make a negative assessment of her. My conception of shame led me to say that Arnica's shame is irrational because her negative self-assessment is not based on her standard of assessment. This means that on my conception, Arnica's shame requires a further story to explain how her self-assessment and her standard of assessment can come apart. I have tried to show that this is not a weakness but a strength of my conception. It yields a more complicated explanation, but this is right because shame is a complex emotion. Deigh's neat explanation fails to capture this complexity. I also argued that this complexity provides us with further resources. On my conception, we have the resources to explain to Arnica how her shame relates to her self-conception, and how this is (or can be) related to her values. We can encourage her to develop these links as she grows up, and potentially not feel shame about her name in future. Deigh's conception cannot provide us with these resources.

Jennifer's irrational shame

In 'An Apology for Moral Shame', Cheshire Calhoun criticises conceptions like mine for failing to give audience a sufficiently central role in shame. She says,

> Grounding the mature agent's shame in her own self-criticism not only fails to explain the social character of shame; it forces us to discount as irrational or immature common shame experiences."[46]

Calhoun also uses cases of shame that I would call "irrational" to make this point. She is particularly concerned with "recalcitrant shame": shame (which I would call "irrational") which persists even after careful reflection. Arnica's shame seems unlikely to be recalcitrant – she might only feel it when her peers jeer at her. Recalcitrant shame in mature, reflective people provides a tougher challenge to my conception than Arnica's shame because the additional story it requires must explain why the connection between the agent's self-assessment and standard of assessment fails persistently, not just in an obviously overwhelming situation like Arnica's.

So let me introduce an example of recalcitrant shame. Imagine that Jennifer is an intelligent, reflective philosophy student who has had an abortion. She feels shame whenever the subject comes up in the presence of her mother, with whom she still lives, who knows about and vocally disapproves of the abortion. However, Jennifer believes strongly that the abortion was not wrong.

On my conception, Jennifer's shame implies that she makes a negative self-assessment. A story needs to be told about how and why Jennifer feels shame over the abortion when she strongly believes that it was not wrong. There are many possibilities, but it seems most likely that Jennifer simply fails in some way to maintain her autonomous judgement about abortion under the shaming criticism of her mother with whom she has developed her standard of assessment and on whom so many of her values depend.

Calhoun does not explicitly develop a positive account of shame, but it appears that on her conception Jennifer's shame is a reaction to her mother's negative assessment of her. For Calhoun, "The power to shame is a function of our sharing a moral practice with the shamer and recognizing that the shamer's opinion expresses a representative viewpoint within that practice."[47] Jennifer's mother criticises her because she believes abortion is wrong, and Jennifer feels shame because she accepts her mother's viewpoint as representative. It is important to see the difference here from Deigh's conception, on which Jennifer might entirely reject her mother's criticism, leaving her sense of real worth unaffected. For Calhoun, Jennifer feels that her real worth is diminished, because she gives her mother's criticism "practical weight".[48] I return to this below.

Calhoun's conception of shame challenges mine on two levels. Firstly, recalcitrant shame like Jennifer's challenges my conception in a similar way to Arnica's. It does not immediately fit my conception of shame, because there appears to be a breakdown between the agent's shame, self-assessment and standard of assessment. Calhoun's conception does not draw a connection between

[46] Cheshire Calhoun, *An Apology for Moral Shame*, p. 135.
[47] Cheshire Calhoun, *An Apology for Moral Shame*, p. 141.
[48] Cheshire Calhoun, *An Apology for Moral Shame*, p. 139.

these three elements and therefore does not need to provide a special explanation of why such a connection breaks down in Jennifer's case. Secondly, on my conception Jennifer's shame is irrational. So insofar as Jennifer strives to be rational, she *should not* feel shame as a result of her mother's criticism, because she believes that she has nothing to be ashamed of. Calhoun's conception of shame leads us to believe that Jennifer *should* feel shame in this situation, and Calhoun argues that this is a merit of her conception and failure of mine.

These two challenges work together. I have said that, in cases of irrational shame, we should not be surprised to see a deviation from my conception of shame. As with familiar conceptions of other emotions, my conception describes rational shame. However, if Calhoun can reverse our intuition that Jennifer's shame is irrational, then the *prima facie* lack of fit between Jennifer's shame and my conception will be a real problem for me. So everything rests on Calhoun's claim that "even if in one's own view one has nothing to be ashamed *of*, one may nevertheless have *reason* to feel ashamed."[49]

Calhoun's central point is that we ought to give "weight" to the criticism of others, even if that criticism is indefensible.[50] The problem is how we give "weight" to moral criticism of ourselves that is without epistemic merit. I would suggest that we understand this notion of "weight" as implying that we have at least some reason to accept that the negative assessment of ourselves might be accurate. Jennifer has good reason to suspect that her mother's criticism is accurate. Jennifer got many of her values from her mother and they continue to share many values. She should not be too readily dismissive of her mother's shaming criticism. She should consider it carefully. However, after careful reflection, Jennifer may come to the conclusion that this specific shaming criticism of her mother's is unfounded and entirely reject it. For Calhoun, Jennifer would still have failed to give her mother's criticism "weight".

So Calhoun would reject my suggestion, because she is adamant that the "weight" she refers to does not imply any reason to accept as accurate. Jennifer need not attach any truth whatsoever to the negative assessment grounding her mother's shaming criticism if it goes against her best judgement. Rather, she should give it "practical weight":

> Moral criticism has practical weight when we see it as issuing from those who are to be taken seriously because they are co-participants with us in some shared social practice of morality.[51]

Jennifer and her mother are certainly "co-participants ... in some shared social practice of morality", but it is difficult to understand exactly what this "practical weight" amounts to. What can it mean for

[49] Cheshire Calhoun, *An Apology for Moral Shame*, p. 143.
[50] Cheshire Calhoun, *An Apology for Moral Shame*, p. 135.
[51] Cheshire Calhoun, *An Apology for Moral Shame*, p. 139.

me to endorse my shame while believing that the negative assessment which grounds it is entirely false? Shame is linked to conceptions of self. It seems that the background to Calhoun's unhappiness with my conception of shame is, as it was for Deigh, a problem with the account of worth and identity that underlies that conception. On Calhoun's account, we should not look to Jennifer's standard of assessment to explain the self-assessment underlying her shame. She says,

> Who I am, morally, is who I am interpretable and identifiable by others as being. That I fancy myself (even with what I take to be the best reasons) to be one kind of person rather than another does not give me an identity as that kind of person.[52]

So Jennifer should see herself as somehow morally deficient under her mother's shaming criticism despite her own best judgement to the contrary, because her mother's assessment of her should outweigh her own self-assessment.

This seems wrong to me. Calhoun accuses defenders of conceptions like mine of failing to take seriously the social practice of morality. But it seems that she fails to take seriously what it is to have values, and what it means to live out these values. A small development of our example will help to bring this out. Imagine that Jennifer is white and is regularly subjected to shaming criticism from her mother over her relationship with her black boyfriend. On Calhoun's account, Jennifer should not (even on reflection) reject this criticism nor develop an invulnerability to it. However, this is bound to negatively affect Jennifer's relationship with her boyfriend. Insofar as Jennifer actively allows her mother's criticism to interfere with her relationship, Jennifer might be failing to live out her values by letting her boyfriend's race influence how she treats him.

Calhoun is right to remind us of the important role others play in the revision of my values and self-conception, but I must also take ownership of them and of their development. While it is important that my values and my self-conception are open to change, there are also times when I must defend them. This is part of what it means to value our values. I shall say more about this in chapter two.

On Calhoun's conception, Jennifer's shame is caused by the audience's negative assessment of her, to which she gives "practical weight". On my conception, the audience's negative assessment of Jennifer causes shame by prompting Jennifer's own negative self-assessment. Jennifer has somehow failed to maintain her autonomous judgement. She has allowed her own self-conception to be diminished despite her beliefs and values. This might seem counterintuitive when we have described Jennifer as an intelligent, reflective philosophy student. But this is precisely where it *is* important to take audience seriously: our self-conceptions and our values grow and develop, as we do, with other people in a social world. We are responsive to the criticism of others. One of the most

[52] Cheshire Calhoun, *An Apology for Moral Shame*, p. 146.

destructive consequences of bigoted shaming criticism is that it leads socially disesteemed people to see themselves negatively. On my conception, shame felt by such people can highlight these destructive self-assessments, and thereby enhance our ability to reject both the assessments and the irrational shame that accompanies them.

The strength of my conception is that it provides us with the resources to explain why we should respond in different ways to different instances of shaming criticism – sometimes we should accept it, and sometimes we should reject it (even though this may be difficult). I said earlier that it is not obvious that I should ignore the standards of my community in my assessments, and I have not suggested that they can or should always be ignored. However, indefensible shaming criticism should be ignored. Jennifer may, after careful reflection, decide that her mother's shaming criticism is indefensible. Under the shaming gaze of her mother she may, similar to Yves, need to keep repeating to herself, *abortion is not wrong, abortion is not wrong*. Calhoun's conception, like Deigh's, does not have the resources to explain the various ways in which we respond to shaming criticism.

Rationality and normativity

Calhoun states that shame "does not second the critical normative judgments that we reach as autonomous, reflective individuals."[53] In fact, on the conceptions of Deigh and Calhoun, shame cannot do this. I have tried to show how shame and our critical normative judgements can come apart on my conception. More importantly, though, my conception shows that shame can second the critical normative judgements that we reach as autonomous, reflective individuals, and shows us how this can happen. This is particularly helpful, as it offers us real insight into how we can and should deal with shame as we experience it in our lives.

In this chapter, I have developed and defended a conception of shame. As yet, I have discussed rationality only insofar as it was necessary for dealing with challenges from Deigh and Calhoun. I do not mean to have implied that rationality and normativity collapse into one another. On the one hand, irrational shame occurs when there is a breakdown between one's values, as they come together in one's standard of assessment, and one's negative self-assessment. On the other hand, inappropriate shame is simply shame that one should not feel. Ultimately, what determines whether one *should* feel shame is whether the object of one's shame is, in fact, shameful. Arnica ought not to feel ashamed of her surname. Jennifer ought not to feel ashamed of her black boyfriend. But Arnica and Jennifer cannot see the objects of their shame *sub specie aeternitatis*. Jennifer cannot know for sure whether her abortion is shameful. Our best moral knowledge is contained in our values. When we feel shame, it is to these values, and our standard of assessment

[53] Cheshire Calhoun, *An Apology for Moral Shame*, p. 145.

based on these values, that we must turn to determine whether our shame is appropriate. If the object of our shame is not shameful according to our standard of assessment, then our shame is irrational. This does not yet mean that our shame is inappropriate; it may be our standard of assessment that is in need of revision.

On my conception, shame is always the site of conflict. In the case of rational shame, something about myself or my life as I am living it is in conflict with my values. In the case of irrational shame, my emotional response to something about myself is in conflict with my values. In either case, something needs to change in order to resolve the conflict and improve coherence. I must change the way I live my life, change something about myself, change my emotional response or change my values. None of these is an easy task, nor is it easy to know which change is the right one. This is what makes shame not only painful but also confusing. As I hope to have shown, the strength of my conception is that it captures this complexity, and gives us some idea of how to deal with these conflicts.

Conclusion

In this chapter, I have provided a conception of shame. My main claim in this thesis is that shame is valuable, so my first priority has been to clarify precisely what it is that I am talking about. I have argued that shame is a painful emotion involving a negative *self*-assessment, in which the agent is both the assessor and the object. This places the self on both sides of the emotion, and raises the challenge of defining the self, to which I shall return in chapter three. I maintained that the self-assessment in shame is connected to the agent's standard of assessment and thereby to the agent's values, while also highlighting that our assessments, our standards of assessment and our values are informed by others. Then I responded to criticism largely based on examples of shame in which the agent's self-assessment and the agent's standard of assessment come apart. I hope to have shown that this kind of breakdown occurs in irrational shame, and that this is a common feature of irrational emotions.

While discussing my conception of shame, I introduced some concepts that I shall clarify further in later chapters. I have referred to values, standards of assessment, and an ideal self-conception, and I have mentioned how conflict between these can occasion shame. In chapter two, I shall contrast this conflict with coherence, and tie this to what I call "the moral life". In chapter three, I shall discuss how values come together in an ideal self-conception, and how this acts as a standard of assessment in shame. I shall do this in the course of developing a model of the self, and describing a process that I shall call critical reflective self-assessment. Finally, in chapter four, I shall have all the pieces I need to describe and defend the value of shame.

Chapter Two: Coherence and the moral life

Introduction

In this thesis I am arguing that, given that we are engaged in a certain kind of overarching project that I shall call "the moral life", shame can make sense and be valuable to us through its contribution to that project by promoting coherence. In this chapter, I shall develop the fundamental notion that grounds the moral life: coherence. I aim to show how the moral life can be understood as a quest for coherence. In doing so, I shall argue for the conceptual richness and practical importance of coherence to our moral lives. This chapter is divided into two sections. In the first section, I develop a coherence model of values, and in the second section, I argue for the practical importance of coherence through a discussion of the familiar and overlapping concept of integrity. But first, I shall briefly describe my use of "coherence" and "the moral life", because these concepts are crucial to my argument that shame plays a valuable role in the moral life by promoting coherence.

The "moral life" is a project of trying to be or to become a certain kind of person who lives life in a certain way. As I said in chapter one, Taylor argues that shame is always a *moral* emotion "provided that morality ... is taken to include personal morality, a person's own view of how he ought to live and what he ought to be."[1] This is the sense of "moral" that I mean to pick out when I refer to the moral life. I take it that just about everyone is engaged in the moral life (with the possible exclusion of very young children and psychopaths). Each person has an idea about the kind of person they would like to be and attempts to be or to become that person. In chapter one, I called this the ideal self-conception and, in chapter three, I shall develop this idea. I am not claiming that the moral life is one's whole life, that it does or should occupy all one's time and energy.[2] By talking about the moral life, I mean to focus on an important aspect of just about anyone's life, rather than a special kind of life that only some of us are living. However, I am happy to acknowledge that this thesis might not be relevant to someone to whom the idea of trying to be or to become a certain kind of person who lives life in a certain way is utterly unfamiliar or unappealing.

"Coherence" describes a relationship between two or more things. One thing can be coherent only insofar as it forms coherent relationships with other elements, or it consists of two or more parts that form coherent relationships with each other. I shall be employing coherence in both ways and concerning two kinds of things: values and actions. Firstly, I aim to show the importance of coherence between values. The moral life is partly concerned with achieving coherence between

[1] Gabriele Taylor, *Pride, Shame, and Guilt*, p. 76.
[2] Indeed, in chapter four I provide reasons to think that such exclusive focus is self-defeating.

values, bringing values together to form a coherent whole. In chapter three, I shall discuss this process of bringing values together into a coherent whole in terms of forming a coherent ideal self-conception. This aspect of the moral life would be practically meaningless if it were not for a second coherence relationship, that between our values and our actions. This can be seen as coherence between a particular value and a particular action, or coherence between our ideal self-conception and our actions. A coherent relationship can exist between a single value and a single action. However, given that any two values can conflict, and conflicting values will often move us towards conflicting actions, there is no guarantee that coherence between any discrete value-action pair will promote overall coherence. What is important is overall coherence between action and our values taken as a whole, and this is what I mean to refer to when I speak about coherence between values and action. So coherence between values and action depends on coherence between values. The moral life can therefore be seen a quest for coherence between values, and between values and action.

Throughout this thesis, I emphasise the reflective nature of the moral life, because I shall be arguing that one of the ways shame promotes coherence is by bringing a new awareness to our moral lives. This happens especially in the way shame highlights incoherence between our values and our actions, but shame also leads us to discover incoherence between our values by prompting critical reflective self-assessment, as I shall discuss in chapter three. An adequate model of value will help us to understand and work with relationships of coherence and conflict between values, and thus provide a valuable resource for better understanding and engaging in critical reflective self-assessment and our moral lives. I turn now to develop such a model.

A coherence model of value

My aim in this section is to develop a model of value. I take the most important utility of any model of value to be its capacity to map the different relationships between values. The relationships between values display varying degrees of coherence or conflict. A model of value should provide a framework for understanding these relationships and resources for promoting coherence and resolving conflicts. In order to allow us to work with values in these ways, a model of value must allow for infinite revisability without encountering the problem of infinite regress. That is, the model must leave every value open to being revised, but at the same time locate a non-arbitrary standard against which such revision can be carried out. There is an intuitive plausibility to a hierarchical model of value, so I shall begin by sketching what such a model would look like. I take Harry Frankfurt's model of the will, as first developed in his landmark paper, 'Freedom of the Will and the Concept of a Person', to be the most familiar hierarchical model in the philosophical

literature so I use this as my starting point. I then examine the capacity of such a model to perform the functions listed above. Finally, I propose an alternative, non-hierarchical model of value and argue that this alternative model is better equipped to perform these functions.

Frankfurt provides a model of the will, not one of values, but his model provides an interesting starting point for thinking about what a hierarchical model of value would look like. The goal of Frankfurt's paper is to present the conditions for personhood, one of which is the ability to have "second-order volitions".[3] On Frankfurt's model, a first-order desire is a desire to perform an action. A second-order desire is a desire to desire to perform an action, or a desire not to desire to do so. And a second-order volition is an effective second-order desire, "one that moves (or will or would move) a person all the way to action".[4] For example, I may have a first-order desire to punch somebody in the face. I may have a second-order desire not to desire to punch somebody in the face. If the second-order desire is effective – if it is a volition – then it will override my desire to punch and I will not do so.

Frankfurt states that it is the ability to form desires and volitions of higher orders that manifests our "capacity for reflective self-evaluation".[5] I think this is why hierarchical models are intuitively appealing, because they take as a starting point our ability to take a step back from our desires and evaluate them. As Stephen L White puts it, "Intuitively, it is clear that not all one's desires are on a par. Some desires are desires we desire not to have."[6] We do not simply fulfil all our first-order desires, nor do those first-order desires that feel strongest always win out against those that feel weaker. We often evaluate our first-order desires, especially those in conflict, and form a second-order desire for one to lead to action rather than another. On a hierarchical model of value fashioned after Frankfurt's model of the will, the capacity for reflective self-evaluation could be mapped in a similar way. Different values move us towards different actions that are frequently in competition. I may value the vengeful satisfaction I would get from punching somebody in the face, without placing any value on the value I place on this kind of satisfaction. It is a value I would rather not have. On a hierarchical model of value, then, conflict at a particular level is resolved at a higher level. In other words, when we metaphorically "step back" to evaluate, this is captured by moving up a level on the hierarchical model.

Despite this intuitive appeal, I want to suggest that a hierarchical model of value fails to capture the way in which we actually engage in reflective self-evaluation. Consider the following example. Mary wants to punch Martha in the face, because Martha has just punched Mary in the

[3] Harry G Frankfurt, 'Freedom of the Will and the Concept of a person' p. 10.
[4] Harry G Frankfurt, 'Freedom of the Will and the Concept of a person' p. 8.
[5] Harry G Frankfurt, 'Freedom of the Will and the Concept of a person' p. 7.
[6] Stephen L White, *The Unity of the Self*, p. 296.

face. Mary places value on the vengeful satisfaction she would get from punching Martha. When Mary steps back and engages in reflective self-evaluation, a hierarchical model of value would suggest that a second-order value develops or becomes apparent to her – in this case, Mary places a disvalue on her vengeful satisfaction. This story makes sense, but it is silent on precisely what we are trying to capture: reflective self-evaluation. How and why does Mary come to disvalue her vengeful satisfaction?

When Mary steps back and engages in reflective self-evaluation, she does not simply form a second-order disvalue of her vengeful satisfaction. The moment she steps back, she will most likely be in a position to acknowledge quite a number of other values bearing on the situation. Mary may be committed to pacifism. She may value actions in which she asserts herself. She may have a deeply ingrained sense of retributive justice. When Mary engages in reflective self-evaluation, these other values will become relevant to her because of the relationships between these values and those that were already explicit on the hierarchical model. My point is not that a hierarchical model cannot allow for the existence of these other values and the relationships of coherence and conflict between them and the values already accounted for; the problem is rather that the hierarchical model has nothing to say about these other values and, more importantly, about the *relationships* between those values already accounted for and these other values..

This is problematic both theoretically and practically. Theoretically, a hierarchical model of value allows for our "capacity for reflective self-evaluation", but fails to capture the way people often actually engage in such a process. The model captures only the result of Mary's reflective self-evaluation, but fails to tell us how or why she arrives at that particular higher-order evaluative attitude. And practically, the model does not prompt Mary to look for other potentially relevant values when she engages in reflective self-evaluation and it does not help her to understand the relationships between these values and those that were already obvious to her.

The hierarchical model employs what Marilyn Friedman has called a "top-down" approach.[7] As I said above, conflict at any level is resolved at a higher level. This means that elements at higher levels of the hierarchy have more authority, because they overrule elements at lower levels. However, as Friedman points out, it is unclear what gives evaluations at higher levels their special authority. She is particularly concerned about a potential evaluative regress in hierarchical models. The potential for an infinite evaluative regress emerges because there is nothing to stop unresolved conflicts at any level from bringing about unresolved conflict at a higher level.[8] Inevitably, the elements on the highest level of the hierarchy, the level at which the regress is arbitrarily brought to an end, cannot themselves be evaluated because there is nothing higher to which one can appeal or

[7] Marilyn A Friedman, "Autonomy and the Split-Level Self".
[8] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 602.

in the light of which one can make such an evaluation without resuming the regress. This model thus fails to allow for what I called infinite revisability without infinite regress, because it inevitably leads us either to appeal to endless higher levels or to terminate the regress arbitrarily, thereby placing some values beyond revision.

While there have been attempts to address these problems with hierarchical accounts, I shall not assess those here.[9] Rather, I take these problems to provide a good reason to explore the potential utility of another type of model to which they do not apply. Friedman proposes that we drop the assumption that "critical reflection must be exclusively a 'top-down' affair",[10] and I agree that this is the right place to start. As White suggests, "One's decisions are motivated by the fact that there are relations of support among one's ... desires."[11] I shall propose a model of value on which values can be assessed in the light of all other values in a multidirectional way, and I shall pay particularly close attention to these "relations of support". The main features of my model are drawn from Laura Waddell Ekstrom's work on autonomy. The elements of Ekstrom's model are not values but "preferences" and "acceptances".[12] For Ekstrom, one's character system is the set of propositions that one accepts (acceptances) and one's preferences. One of the key differences between Ekstrom's model and that of Frankfurt lies in her use of preferences rather than second-order volitions. Ekstrom states that

> A preference ... is formed because one finds a certain first-level desire to be good, either in itself or as a means in a particular instance to realizing one's general conception of the good.[13]

A preference is thus an evaluative attitude formed in the light of one's conception of the good. Frankfurt's model, as I briefly described it earlier, consists of desires and volitions, neither of which is essentially related to one's conception of the good. Ekstrom's introduction of this evaluative priority goes a long way to avoiding the problems of regress raised by hierarchical accounts, because "acceptances and preferences are formed based on an agent's subjective conceptions of what is good and true...".[14] Thus, Ekstrom's model provides an answer to Friedman's worry about the source of the special authority of elements that function like Frankfurt's second-order volitions. Ekstrom's preferences gain their special authority from their essential connection to one's conception of the good.

---

[9] See Harry G Frankfurt, *Necessity, Volition and Love*, especially "On the Necessity of Ideals", pp. 108 – 128.
[10] Marilyn A Friedman, "Autonomy and the Split-Level Self", p. 30.
[11] Stephen L White, *The Unity of the Self*, p. 297.
[12] Ekstrom takes the notion of "acceptances" from the epistemological literature on coherence. The notion is not important for my discussion.
[13] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 604.
[14] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 608. I have dropped her italics.

Ekstrom does not make it clear whether or not one's conception of the good is simply constituted by one's preferences and acceptances. If it is, then this is a feature of her model that I wish to borrow, and to which I shall return below. If not, then it seems that one's conception of the good would lie somewhere beyond the scope of the model, and this would be a significant limitation because it would mean that the model cannot have much to say about revising that conception. What kind of elements would then form one's conception of the good, and how are these evaluated? I require a model for making sense of coherence between values, and a model of value cannot appeal to a conception of the good which lies outside of the model, because one's conception of the good is constituted by values. So if we understand Ekstrom to place the conception of the good somehow outside of her model, then a model of value fashioned after hers will not work for me.

The central feature of Ekstrom's model that I wish to borrow is the notion of coherence. Ekstrom proposes that "we take an agent's *true* or *most central self* to be a subset of these acceptances and preferences, namely, those that *cohere* together".[15] She describes various features of coherent elements, and I shall go through each of these features in turn. First, though, it is worth reiterating that I am pulling out features of the notion of coherence she uses for a model of acceptances and preferences and redeploying them with the goal of developing a model of values. So I am evaluating their appropriateness for my model of value, not for Ekstrom's model of autonomy.

Ekstrom says that cohering elements "fit together; they hold together firmly, displaying consistency and mutual support".[16] One can develop this thought with a variety of metaphors or similes. The most helpful image is perhaps that of spherical magnets of various sizes, strengths and polarities.[17] Some elements will repel each other, while others will group together, often with many smaller elements coming together around bigger elements, forming mutually supportive clusters. It is difficult to give a general answer to why some groups of values form mutually supportive coherent relationships while others do not, but I take it to be a common observation that this does in fact happen. For example, sympathy and kindness easily form a coherent relationship, whereas vengefulness is likely to conflict with both. Conflict also happens in various ways, with perhaps the most obvious being a logical conflict, but values can come into practical conflict too. Ekstrom explains the long-lasting nature of core elements by referring to their support by good reasons. I think this is right, if we understand "good reasons" broadly. In chapter four, I shall describe an emotional connection that one can have with one's values. In certain contexts, this connection may

[15] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 608.
[16] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 608.
[17] This image is even more helpful if we imagine magnets to come with more than two possible polarities.

count as a "good reason" for retaining some values or removing others. Coherence itself can count as a "good reason": an element that is supported by other elements in a cluster may be hard to remove or revise without engaging with the other elements in the cluster. Of course, that something counts as a "good reason" does not imply that it counts as a sufficient reason.

I am hesitant to adopt Ekstrom's characterisation of coherent elements as those that are "*fully defensible* against external challenges"[18] because I am not sure what this means. Would full defensibility of a value require that one could give enough "good reasons" such that anyone should adopt the value insofar as they are reasonable? This seems far too demanding, and I would rather say that elements that are supported by good reasons would be more likely to survive a process of critical reflective evaluation. This is something less than "fully" defensible because, as I said above, good reasons need not be sufficient reasons. And it seems possible that what counts as a good reason for me may not count as such for others, from whom the "external" challenge would presumably come. So I would say that coherent elements are defensible against challenges, from myself or others, rather than "*fully defensible* against external challenges".

Ekstrom states that coherent elements will be "one's deepest attitudes about the world, where deeply held attitudes may be taken to be those to which one firmly and fervently clings in the face of challenge".[19] The elements of my model are values interpreted very broadly rather than Ekstrom's carefully defined preferences and acceptances, so my model will contain much else besides one's deepest attitudes about the world. This is not to suggest that all values have similar importance to us, or that coherence between any two pairs (say) of values is of equal importance. Rather, as my metaphor of magnets above suggests, values should be pictured as coming in a variety of sizes and strengths, and some will have a far bigger influence than others on overall coherence. As for clinging "firmly and fervently", this can be taken in two ways. On the one hand, we defend coherent elements against challenges, as I discussed above. On the other hand, we develop an emotional connection to our values, as I shall discuss in chapter four.

Ekstrom's description of cohering elements as those that "one is *comfortable owning*"[20] seems open to misinterpretation. One may find oneself in situations where one is made distinctly uncomfortable even by those values that fit most coherently with one's overall set of values. The most obvious case is that in which one finds oneself to be greatly at odds with those around one with regard to some central value(s). This may well occasion what I described in chapter one as "irrational shame". Still, while "comfort" is not the most precise word here, we can talk metaphorically about one being "comfortable" with that value. Ekstrom goes on to say that one "is

[18] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 608.
[19] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 608.
[20] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 608.

not conflicted."[21] This is more helpful, although also potentially misleading. Ultimately, if one had an entirely coherent set of (acceptances and preferences or) values then one would never be (seriously)[22] conflicted. But clearly talk of an entirely coherent set is only interesting theoretically and not likely in practice. An element that is broadly coherent with the other elements in the set will, however, create far less conflict, and this observation is important.

Thus far, I have limited my discussion to coherence between values. Coherence between actions and values depends largely on coherence between values. If one has two values that conflict, then an action that promotes one value is likely to undermine the other. So what is important is coherence between actions and one's values taken all together, and clearly this is easier to achieve if one's values come together in a coherent way. Psychic conflict can be paralysing and present a serious obstacle for living a moral life as I have described it, and perhaps for living any life at all. Frankfurt acknowledges this point explicitly.[23] The utility of a model of value is its capacity to help us understand the different relationships of conflict and coherence between values and thereby provide us with resources to resolve conflicts and promote coherence. I have provided some reasons to think that a hierarchical model will fail to do this adequately. I am proposing an alternative model with a far less rigid structure. This model describes the kinds of structures of values we are likely to find and the kinds of relationships between values that hold these structures together (or apart). However, each of us must examine our actual values to discover the structures and relationships that actually exist within our set of values. My model of values therefore exists in two forms. On the one hand, it provides a theoretical framework for understanding values and relationships between values in abstraction, and on the other hand, it can be used to map an individual's values to form a working model of those values and the relationships between them.

The structures that emerge on a working model of a particular person's values will be unique to that person, as a result of the mutually supportive bonds that develop between actually cohering values and the repelling fields that form between conflicting values. This working model will more closely resemble clusters or perhaps even galaxies in a cosmos than a neatly tiered hierarchy; it will be complicated and always changing; but it will more accurately map the complicated nature of conflict and coherence between our evaluative attitudes. Moreover, I believe that this more accurate model provides us with a far better resource for resolving conflicts and promoting coherence. It is an improvement on the hierarchical model because it can more easily accommodate the various types and strengths of relationships between different values, and it leaves values open

[21] Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 608.
[22] On Ekstrom's model, it would still be possible to have conflicting desires, but not conflicting "authorized preferences".
[23] Harry G Frankfurt, 'Freedom of the Will and the Concept of a person', p. 16.

to critical reflection at any time, not only when they come into conflict by moving us towards alternative actions. Most importantly, it simultaneously allows for infinite revisability without encountering the problem of an infinite regress.

On my model, then, we can see how values like sympathy and kindness are likely to be found together in the same cluster, while vengefulness is unlikely to be part of this cluster. However, vengefulness might be found in another cluster together with other values such as righteousness and retributivism. Reconsidering a value like vengefulness may lead us to reconsider these other values, which may in turn lead us to reconsider the relationship between these values and those in the sympathy-kindness cluster.

On the hierarchical model, conflict is resolved by appealing to a higher level on the hierarchy. One of my complaints against Ekstrom was that she was unclear about where the conception of the good was located. On the model of values I have proposed, we can locate my conception of the good in my ideal self-conception as constituted by my values.[24] Conflict is resolved by appealing to other relevant values which cohere or conflict with those values undergoing evaluation. My model therefore avoids evaluative regress by locating a standard of assessment that is internal to our value set. And it allows for infinite revisability because all values, including those that are most fundamental, are always open to revision, at least in principle.[25] So it responds to Friedman's call to avoid an exclusively top-down approach to critical reflection. Some values are more fundamental than others and, as such, they have a greater effect on a broad range of other values, such that the adoption, revision or rejection of a more fundamental value may lead to the adoption, revision or rejection of many other values. However, by calling these values "more fundamental" I mean only to point to the fundamental role played with regards to other values. Values that are more fundamental are not of a different type or class to other values, and they have no special status that places them beyond the scope of assessment and potential adoption, revision or rejection.

This model of values therefore provides a valuable resource for better understanding and engaging in our moral lives. It also helps us to better understand the role of shame within our moral lives. In chapter one, I said that rational shame involves a mismatch between my self-conception and my ideal self-conception. In the case of shame occasioned by an action, the action informs my self-conception. In chapter three, I shall describe how values come together to form my ideal self-conception. So, in such cases, shame involves a mismatch between my values and my actions. It

[24] I discuss this in far more detail in chapter three. This could replace Ekstrom's "subjective conception" of the good. See my discussion above and Laura Waddell Ekstrom, 'A Coherence Theory of Autonomy', p. 608.
[25] I say "at least in principle" to acknowledge that one may find it psychologically impossible to revise some values, or at least to do so without seriously jeopardising one's sense of identity.

should be clear, following my discussion in this section, that this mismatch can be understood as a lack of coherence between my *values* and my *actions*, and that this is often caused by a lack of coherence between my *values*.

Coherence and integrity

On the model that I have described above, coherence might seem like a technical notion limited to describing a certain formal relationship between one's values, leaving the moral significance of coherence unclear. In this section, I aim to show how the formal constraints of coherence can be seen to generate the substantive constraints that define our concept of integrity. As Nancy Schauber points out, "Attributing integrity to a person is arguably the most respectful praise we can offer."[26] We commonly take integrity to be importantly valuable both in ourselves and in other people. I take it that showing the overlap between the concepts of coherence and integrity will go a long way towards showing the value and practical importance of coherence in our moral lives.

Integrity thus forms an important link in the central argument of this thesis. I am arguing that shame is valuable because it promotes coherence, and promoting coherence is valuable because it constitutes an important aspect of our lives. My discussion of integrity bears on the second step. Integrity is a familiar moral concept and my discussion of integrity in terms of coherence is meant as an extended example of the importance of coherence in our lives. There is also a more direct relationship between shame and integrity. In chapter one, I defined shame as a mismatch between my ideal self-conception and my actual self-conception. I also suggested there, and I shall discuss in more detail in chapter three, how we may see my values coming together to form my ideal self-conception. Actions that occasion shame are likely to be those that do not relate coherently to my values as they come together in my ideal self-conception. It should become clear that these actions are also those that reveal a lack of integrity. So actions that reveal a lack of integrity are likely to occasion shame. Shame can therefore act as a guardian of integrity, both alerting us to its loss and motivating us to maintain it.

Demonstrating the conceptual overlap between coherence and integrity is made difficult by a central tension that exists in our general concept of integrity and that threatens to tear it apart. The challenge for any conception of integrity, including the conception I shall offer based on coherence, is to resolve this tension by accommodating both of the elements that are in tension without losing the intuition behind either of them. Lynn McFall identifies this tension very clearly and goes some way towards developing a coherence-based conception, so I shall follow her paper

[26] Nancy Schauber, 'Integrity, Commitment and the Concept of a Person', p. 119.

on integrity closely. She argues, however, that the formal conditions for coherence cannot fully capture our concept of integrity, and that we need to add explicit substantive constraints. I shall argue that we do not. I propose that we understand integrity to require coherence between values and coherence between values and actions.

Webster's dictionary captures both elements of the tension in its definition of integrity, which is:

1. an uncompromising adherence to a code of moral, artistic, or other values;
2. utter sincerity, honesty, and candor; avoidance of deception, expediency, artificiality, or shallowness of any kind.[27]

McFall begins her paper by observing the tension between these two elements. The first element ties integrity to an agent's adherence to their own values, while the second element ties integrity to certain conventional moral principles. It cannot be taken for granted that these coincide. Rather, the definition includes both these elements precisely to accommodate conflicting intuitions about the ascription of integrity in certain cases. Some philosophers, like Cox, La Caze and Levine, argue that it is the second element, "utter sincerity, honesty, and candor; avoidance of deception, expediency, artificiality, or shallowness of any kind" which is "conceptually prior and most central to understanding integrity",[28] while McFall appears to prioritise the first element,[29] "an uncompromising adherence to a code of moral, artistic, or other values". However, McFall explicitly declares her aim to "give an account that does justice to both senses as well as clarifying the relation between them",[30] and this is also my aim.

I shall argue that McFall is right to prioritise the first element. She argues that we can ground adherence to a code of values in the formal conditions for coherence. However, she seems to feel that in order to account for the second element of the definition – conformity to certain conventional moral standards – it is necessary to add the condition of importance and what she calls the "Olaf Principle" to her conception. I shall argue that neither of these is necessary, because what she intends to capture using these added conditions is either external to the concept of integrity, or it is already covered by the conditions for coherence. I shall show how an account of integrity based on the formal conditions for coherence can accommodate all of our common intuitions about integrity. This is especially challenging when, as I have already suggested, these intuitions may come into conflict. Part of my strategy will be to highlight the way in which some of these intuitions are more obviously linked conceptually to coherence, while others are associated with our practices of

[27] Webster's Third New International Dictionary.
[28] Cox, La Caze & Levine, *Integrity and the Fragile Self*, p. 1.
[29] Lynne McFall, 'Integrity', p. 5.
[30] Lynne McFall, 'Integrity', p. 6.

ascribing integrity as a form of praise. It will become clear that it is often very difficult to make an ascription of integrity with justifiable confidence, and so we often make inferences based on conditions we associate with coherence. But these conditions are only linked to integrity via their contingent association with coherence, so they are defeasible and not independent conditions for integrity.

This practice is neither unreasonable nor uncommon, and we have similar practices concerning other features of people. For example, we would not normally think that a young child could be wise. Age is not part of the concept of wisdom, nor is it impossible for a young child to be wise. Nevertheless, knowing what we do about wisdom and people and the ways people actually become wise, it is reasonable to treat age as a defeasible condition in our practice of ascribing wisdom to people. Similarly, I shall argue that it is not impossible for someone to have integrity without satisfying a condition like consistency of values over time but that, given what we know about people and values and the way people actually adopt and revise values, it is reasonable to treat consistency of values over time as a defeasible condition when ascribing integrity to people. However, it is not an independent conceptual condition for integrity.

McFall lists four formal conditions for personal integrity:

> personal integrity requires that an agent
> (1) subscribe to some consistent set of principles or commitments and
> (2), in the face of temptation or challenge,
> (3) uphold these principles or commitments,
> (4) for what the agent takes to be the right reasons.[31]

It is worth noting that McFall uses the terms "principles and commitments" whereas I have used the term "values" in this chapter. I shall adopt her terms for the duration of this section. McFall also uses the term "consistent" in condition 1 where I would use "coherent", but she describes consistency as coherence applied to principles, and contrasts it with conflict, so I shall treat the two as synonymous.[32] I shall argue that conditions 1 and 3 should be accepted while conditions 2 and 4 should be rejected.

The first condition requires an agent to subscribe to some consistent set of principles or commitments. It should be obvious that no one has a perfectly consistent set of principles or commitments; we are bound to have some that conflict. Does this imply that no one has integrity? McFall states that one "cannot maintain one's integrity if one has unconditional commitments that conflict".[33] By "unconditional commitments", I take McFall to mean those commitments that must

[31] Lynne McFall, 'Integrity', p. 9.
[32] Lynne McFall, 'Integrity', p. 7. She says, "One kind of coherence is simple consistency: consistency within one's set of principles or commitments."
[33] Lynne McFall, 'Integrity', p. 7.

be upheld in any circumstances. But limiting the condition to unconditional commitments does not help. Firstly, I am sympathetic to Cox, La Caze and Levine's view that

> Requiring consistency among one's unconditional commitments is irrelevant because, despite what we may like to believe, few of us have unconditional commitments. There may be things we would not do under just about any circumstances, but this is different from having unconditional commitments.[34]

Secondly (and perhaps it is the same point) I think the many weird and wonderful thought experiments described in the ethical literature show us how any two commitments (unconditional or not) can be brought into conflict. So, limiting the consistency requirement to unconditional commitments is unhelpful. We should rather understand the first condition to require a certain degree of consistency between one's principles or commitments. Indeed, it seems clear that integrity will always be a matter of degree, rather than something one either does or does not have. When we speak about someone having or not having integrity, this is shorthand for ascribing to them a degree of integrity that does or does not meet some (probably vague and subjective) minimum expectation we have.

The second condition, that integrity requires temptation or challenge, should be rejected as an independent conceptual condition for integrity. This may seem like a moot point because any principles or commitments we have are bound to be challenged,[35] and this makes it difficult to think of a plausible counterexample to this condition. Still, let us imagine that a (heterosexual) nun committed to sexual abstinence moves from a mixed community to live in an isolated nunnery where there are no men. She would not lose integrity during the move, she does not have less integrity when she arrives at the nunnery, and it seems obvious that she would not lose integrity after staying at the nunnery for some time. One does not somehow lose integrity because one's principles go unchallenged for a certain length of time. So we must take condition 2 as implying, rather, that the question of integrity simply does not arise in the absence of temptation or challenge, that this is not a situation in which it is relevant to talk about integrity. Here I would like to point to the distinction I suggested earlier between the conceptual conditions of integrity and conditions linked to the ascription of integrity. A lack of temptation or challenge does not lead one to lack integrity, but we praise people who maintain their integrity in the face of temptation or challenge because it is more difficult and shows strength of character. Temptation or challenge is therefore not a condition for integrity and not a necessary element of our conception.

[34] Cox, La Caze & Levine, *Integrity and the Fragile Self*, p. 10.
[35] Cox, La Caze & Levine, *Integrity and the Fragile Self*, p. 8.

It might be suggested that we should understand condition 2 to require only a disposition towards upholding principles or commitments in the face of temptation of challenge.[36] This would allow for the nun to have integrity, if it is true that she would not break her vow of abstinence in the face of temptation, even if this temptation never actually presented itself. This condition can be taken in two ways. Firstly, the disposition might itself be part of the concept of integrity. In other words, part of what it means to have integrity is to have a disposition towards satisfying conditions 1 and 3 above. This would imply that if the nun succumbs to temptation and breaks her vow, then this proves that she did not have the relevant disposition and therefore did not have integrity in the first place. On this interpretation, condition 2 constitutes an independent condition that cannot be directly accounted for by the conditions of coherence. I would reject condition 2 on this interpretation. On my conception, the nun maintains integrity until she breaks her vow. This tracks our use of the broader concept of integrity as it is applied to other things. Cox, La Caze and Levine note how we would say that a "database maintains its integrity as long as it remains uncorrupted by error."[37] Similarly, I think, people maintain their integrity until they lose it – that is, until coherence is corrupted.

On a second interpretation, we can see condition 2 as a defeasible condition that captures our practice of associating coherence with the disposition towards upholding principles and commitments, rather than taking this disposition as an independent condition for integrity. As such, it captures an intuition that may lead us to withdraw the ascription of integrity retrospectively: we might say that when the nun succumbs in the face of challenge, it shows that there was in fact less overall coherence between her principles and commitments than we had previously assumed. Greater coherence will generally lead to less conflict, and a less conflicted person will generally be less likely to succumb in the face of challenge. So we may take the nun's actual failure to uphold her principle to tell us something about the degree of coherence between her principles and commitments. On this interpretation, condition 2 is a defeasible condition which we may reasonably use in ascribing integrity, but it is not a conceptual requirement for integrity because it relies on the formal conditions for coherence that are already accounted for on my conception.

Condition 3 requires that an agent uphold their principles or commitments. McFall also describes this as "coherence between principle and action".[38] If our nun breaks her vow, she would not have integrity insofar as she did not act in accordance with her principle of abstinence. Of course, we can invent a ridiculous situation where the nun could save someone's life by having sex. In such a situation, would it be possible for the nun to maintain her integrity? I shall return to this

[36] Articulating condition 2 as a disposition was suggested to me by Thad Metz.
[37] Damian Cox, *et al*, 'Integrity'.
[38] Lynne McFall, 'Integrity', p. 8.

issue below, but it seems plausible that integrity requires coherence between one's actions and one's set of principles or commitments taken as a whole. So the nun could, in that ridiculous situation, have sex and maintain her integrity.

The fourth condition follows from the third but adds a further requirement: my actions must not only be consistent with my principles, my actions must also flow from my principles in the sense that to act with integrity I must act because of my principles. This seems to imply that I cannot act with integrity if I am lucky enough to find myself in a situation where my principles and actions are conveniently aligned. To call on my nun once more, imagine now that she simply has no libido. She has taken a vow of sexual abstinence and is fully committed to the principles that ground that vow. But, frankly, she finds even the thought of sex so off-putting that it never comes down to a defence of her principles. Condition 4 shares the problems of condition 2. In the absence of evidence of internal struggle, the nun might not be praised for her integrity, but she does have integrity. Again, if we imagine that the nun has developed this distaste for sexual activity, it seem ridiculous to suggest that by bringing her tastes into line with her principles she has lost integrity. Such a principle would counterintuitively put integrity out of reach of some moral exemplars like Aristotle's fully virtuous person, who has trained their tastes to coincide with, and reinforce, their principles. And again, like condition 2, condition 4 seems more closely linked to the ascription of integrity than to the conditions essential to the concept. If we were to observe the nun restraining sexual urges in favour of her principles, we would have evidence on which to infer a certain degree of coherence. In the absence of such evidence, we may withhold the ascription of integrity because we simply do not know one way or the other. However, this does not mean that the nun lacks integrity, and McFall's fourth condition should therefore be rejected.

I have argued that conditions 1 and 3 should be accepted and understood to flow from coherence, while conditions 2 and 4 should be rejected. This leaves the following formal conditions: integrity requires that an agent subscribe to some consistent set of principles or commitments and uphold these principles or commitments. These conditions can also be described as requiring coherence between principles, and coherence between these principles and action. I shall now discuss how these formal conditions for coherence generate substantive conditions for integrity.

Integrity does not require that an agent get their principles right, and an ascription of integrity does not imply approval of an agent's principles. Cox, La Caze and Levine note that when we make an ascription of integrity

> It is often to ameliorate criticism of another's moral judgement. For example, we may disagree strongly with the Pope's views of the role of women in the Church, take this to be a significant moral criticism of him, and yet admit that he is a man of integrity.[39]

However, there do appear to be limits to the kinds of principles a person of integrity can have. McFall asks us to consider the following statement:

> John was a man of uncommon integrity. He let nothing – not friendship, not justice, not truth – stand in the way of his amassment of wealth.[40]

McFall argues that John is not a candidate for integrity because there is no possibility of conflict in which integrity might be lost, and "[w]here there is no possibility of its loss, integrity cannot exist".[41] This suggests that certain kinds of principle are incompatible with integrity because they fail to meet the formal constraints discussed above.[42] Here, condition 2 (which requires the presence of temptation or challenge) cannot be met. I have rejected that condition, so if I wish to reject John as a candidate for integrity, I must look elsewhere.

I think McFall is right to argue that formal conditions for integrity give us reason to suspect that certain kinds of principles will be incompatible with integrity. Thanks to the Enron scandal, real life examples of people like John exist in our collective consciousness. Former Enron CEO, Jeffrey Skilling, was convicted of conspiracy, insider trading, securities fraud and making false statements to auditors. The challenge to my conception of integrity is to explain why we withhold an ascription of integrity from a real person like Skilling. My conception cannot rule out the logical possibility of someone like John having integrity,[43] but it can explain why we have good reasons to suspect that real people living in the real world cannot live with integrity while financial greed, for example, is their highest or only principle.

As I suggested earlier, it is usually very hard to know whether a person has integrity. We have to make assumptions about the principles a person has. I have argued that integrity requires coherence between principles, and between principles and action. John does not let friendship, justice or truth stand in the way of his amassment of wealth. But, firstly, it is simply hard to imagine a real person claiming in all sincerity that their highest principle is wealth amassment. And secondly, it is hard to imagine that a real person does not have important, interconnected principles related to friendship, justice and truth, among other things. And it seems likely that these principles and a relentless pursuit of wealth will be in constant and deep conflict, which will present a serious

[39] Damian Cox, *et al*, 'Integrity'.
[40] Lynne McFall, 'Integrity', p. 9.
[41] Lynne McFall, 'Integrity', p. 9.
[42] Damian Cox, *et al*, 'Integrity'.
[43] But Bernard Mayo might have done so, arguing that "rational egoism is an incoherent notion." See Bernard Mayo, 'Moral Integrity', p. 37.

obstacle for the development of overall coherence. This is even more clearly the case with Skilling, who has a wife and friends and a real human life. In practice, we rarely consider carefully what assumptions are behind our ascriptions of integrity, and how these are related to coherence. I suspect that we often use a deliberative shortcut. We all strive for some level of coherence.[44] If, in our experience, we find that certain principles appear particularly inclined to produce overall incoherence, we will be less inclined to ascribe integrity to people who have those principles. This is why the second element in the dictionary definition of integrity lists certain conventional moral principles. These are principles that most people have found to be conducive to coherence, given the other principles and commitments that they have.

Let me be clear that I am not claiming that if one has a coherent set of principles then one will also have the right principles. This is where the second element in the dictionary definition gets it wrong, tying integrity directly to conventional moral principles, and generating the tension that I highlighted at the beginning of this section. I am not ruling out the possibility that an agent can have entirely the wrong principles and still have integrity. I am merely pointing to the implausibility of that situation given the kind of creature we are and the kind of lives we live. There have been two threads to my argument: firstly, I have tried to explain the way people actually make ascriptions of integrity; and secondly, I have tried to show that this practice is reasonable on my conception of integrity. The strength of my conception is thus that it both explains and justifies the strong but defeasible connection between our ascriptions of integrity and conventional moral principles while avoiding conceptually unnecessary conditions. However, even if my justification of our practice in ascribing integrity fails, it may yet be an accurate description of our practice. Even if it is unreasonable of us to assume that everyone has enough of the kinds of principles and commitments that will cohere better with conventional moral principles like sincerity and honesty, this does not mean that we do not make these assumptions or that these assumptions do not explain our practice.

I hope to have shown that our concept of integrity can be captured using only the formal requirements of coherence between principles, and coherence between these principles and action. McFall argues that we need to add two further elements to her earlier formal conditions for integrity that I believe are unnecessary, because the conceptual work assigned to them can already be done by coherence. These extra elements are "importance" and the "Olaf Principle".[45] McFall states that the Olaf Principle "requires that some of one's commitments be unconditional".[46] I have already raised doubts about unconditional commitments, so I shall not discuss the Olaf Principle here. The

[44] All of us, that is, who are engaged in the "moral life". See my discussion in the introduction to this chapter.
[45] Lynne McFall, 'Integrity', p. 5. The name comes from an e. e. cummings poem.
[46] Lynne McFall, 'Integrity', p. 11.

importance condition requires not just that principles are important to the agent, but that they are principles a reasonable agent could take to be important. So we would not praise someone for their integrity in choosing to stay up late at night in spite of their tiredness to count blades of grass, even if blade counting is an important commitment to *them*.

The importance requirement in practice collapses into the restrictions I have already discussed. In ascribing integrity we place limits on the kinds of principles an agent can have due to assumptions about their compatibility with overall coherence. To say that a principle is important to me is to say that I adopt and endorse that principle. And to say that some principles are more important to me than others implies that I will strive to act in accordance with those more important principles when they conflict with less important principles, because judgements of importance just are value judgements.[47] If our grass blade counter is a brain surgeon who paralyses a patient the following day performing very sleepy surgery, she either made the wrong value judgement by placing more importance on her blade counting than her responsibilities as a surgeon, or she has failed to act on her judgement to the contrary. The limits we place on what kinds of principles an agent can have in order for us to ascribe integrity to them apply also to the ordering of those principles for the same reasons.

I turn now to two interesting and connected questions with which McFall does not engage. In what circumstances can an agent abuse their principles and retain their integrity? And in what circumstances can an agent reject their own principles and retain their integrity? It seems that any plausible account of integrity, while requiring a consistent commitment to principles over time, must also allow for principles to change.

Integrity requires coherence. No one has an entirely coherent set of principles and so, as one strives for coherence, one will need to change one's principles, rejecting or refining some while adopting others. As Hayden Ramsay says, "A vital element of integrity is having the flexibility to transcend our current commitments for the sake of overall coherence and moral excellence."[48] A conception of integrity that does not allow for this process would be deeply counterintuitive. However, a person who regularly makes radical rash changes to their principles and commitments does not strike us as an exemplar of integrity. Again, we need an account that explains why we place this restriction on integrity, without making that restriction too strong. It may seem that integrity must not only allow for change, but require it. Cox, La Caze and Levine take this view, arguing that integrity "is as much, if not more, about when to break certain commitments as it is about when to

[47] Damian Cox, *et al*, 'Integrity', *Stanford Encyclopedia of Philosophy*
[48] Hayden Ramsay, *Beyond Virtue*, p. 105.

keep them."[49] I disagree. Integrity must allow for principles to change, and in certain situations one may be required to change in order to maintain integrity, but change is not an independent requirement for integrity.

The characters of Valjean and Javert from Victor Hugo's *Les Misérables* provide interesting test cases. After many years in jail for a minor crime, Valjean's principles undergo quite radical change in his quest for redemption. Javert, the police inspector who obsessively pursues Valjean, is unwaveringly committed to the principles of the harsh French penal system and its underlying ethic of just deserts. Finally faced with Valjean's goodness, Javert commits suicide, unable to reconcile this revelation with his rigid principles. It seems to me that both Valjean and Javert display integrity. Valjean's principles change, but this change brings about greater coherence. Most of us would say that Javert should change his principles (that he has the wrong principles) but he is incapable of doing so. However, Javert is able to maintain broad coherence until very near the end, when "[a]ll the axioms which had been the supports of his whole life crumbled away".[50] So integrity may require us to change our principles. Through our lives, we may adopt new principles that conflict with old principles, and in order to maintain integrity we must refine or reject either the new or the old principles. However, integrity does not require us to adopt new principles in the first place. What the example of Javert shows us is that integrity is not everything. We should resist the temptation to turn integrity into virtue *per se,* or to think that a lack of integrity is the only serious charge we could make against a person.

In general, then, I want to argue that integrity will require a certain level of consistency over time under normal circumstances, but integrity also allows for changing principles in some circumstances. I do not want to introduce any special constraints on the reason or motivation for a change of principles. The most obvious circumstance in which principles can change without a loss of integrity is where they are found to be inconsistent with other principles. Similarly, though, when new principles are adopted which conflict with older principles, it is not obvious that keeping the old principles will always lead to greater coherence. However, it does seem likely that the balance will often lie in favour of more established and fundamental principles given their greater impact on overall coherence.[51]

I shall now turn to the question about action. In what circumstances may I act against my principles but maintain my integrity? I suspect that there are no such circumstances. In many

[49] Cox, La Caze & Levine, *Integrity and the Fragile Self*, p. 2. They continue on the following page: "A lack of certainty in just about all walks of life is often generated by experience, reflection and maturation, as well as by self-deception, rationalization and other psychological mechanisms that serve, variously, to both undermine and preserve psychological well-being in a plethora of ways."
[50] Victor Hugo, *Les Miserables*, p. 890.
[51] See my discussion of more fundamental values in the previous section.

circumstances, I may act in accordance with some of my principles, and against others. However, as I have already discussed, integrity requires overall coherence between action and my principles taken as a whole. When I act in accordance with an unimportant principle but against other more important principles, I do not act with integrity. In order to perform some actions and maintain coherence, a huge shift in my set of principles and commitments may be necessary. The example of Javert shows us that change on this scale might even be impossible, at least psychologically. This explains the intuition behind McFall's "Olaf Principle", that integrity requires that there are "some things one is not prepared to do".[52] It may seem that, theoretically, I can perform any action and maintain coherence, simply by first changing my principles, but in practice this can be difficult, if not impossible.[53] I hope to have shown again that the strength of my conception is that it shows why some constraints, like consistency over time, are important but also defeasible, while other constraints, like coherence between principle and action, are not defeasible. My conception also accounts for the intuition behind the "Olaf Principle" without making it a strict requirement for integrity.

I have argued that integrity should be understood to require coherence between principles or commitments, and coherence between these principles or commitments and action. I have shown that these formal conditions generate substantive conditions when applied to real people in the real world. Our ascriptions of integrity rest on assumptions about what kinds of principles are generally more conducive to coherence, and these assumptions are often based on our own experience as agents seeking greater coherence for ourselves. However, these assumptions may be wrong, so the substantive conditions generated by my conception are defeasible, they are not fixed restrictions and they are not essential to the concept of integrity. My conception cannot rule out the possibility that John has integrity, but it can explain why most of us would withhold an ascription of integrity from Jeffrey Skilling. In sum, I hope to have shown that my conception can make sense of all of our common intuitions about integrity without including any *ad hoc* or unnecessary conceptual baggage.

This discussion should go a long way towards showing the relevance of coherence for our moral lives. Where I have been using McFall's terms, "principles" and "commitments", I would prefer to talk about values. McFall clearly uses these terms quite broadly, and everything I have said about coherence between principles and commitments in the above discussion applies to coherence between values. As I proposed at the beginning of this section, a discussion of the relationship between the concepts of coherence and integrity shows not only that a certain level of coherence is a prerequisite for integrity, but furthermore that the formal conditions of coherence can account for

[52] Lynne McFall, 'Integrity', p. 11.
[53] I suspect that there may also be theoretical problems with this, such as constraints internal to the concepts of holding a value or subscribing to a principle, but I shall not explore these here.

all of our common intuitions about integrity. There is significant overlap between the two concepts. In the previous section, I showed how a coherence model of values provides valuable resources for understanding the relationships between our values. These two sections together point to the conceptual richness and practical importance of coherence and constitute a strong argument for employing the concept in this thesis and in ethical discussion in general.

Conclusion

I am arguing that shame can make sense and be valuable to us through its role in promoting coherence and thereby contributing to our moral lives. In this chapter, I have developed a notion of coherence and claimed that the moral life can be understood as a quest for coherence. I developed a model of values on which we can best understand the relationships of coherence and conflict between values. I then demonstrated an application of the notion of coherence between values and action through a discussion of integrity.

In chapter three, I apply the notion of coherence to the relationships between the ideal self-conception and character. I have already mentioned very briefly, here and in chapter one, how we can understand my values to constitute my ideal self-conception. I shall discuss this in more depth, and argue that we can similarly understand my actions to inform my character. The focus of chapter three is the process of critical reflective self-assessment, in which I reflect on the relationship between my ideal self-conception and my character.

So the notion of coherence is central to my whole project. Where previously I spoke about rational shame involving a mismatch between my self-conception and my ideal self-conception, it should now be clear that this "mismatch" is a relationship that lacks coherence. It should also be clear that incoherence between any two values could contribute to incoherence within my self-conception, and in turn to incoherence between this and my ideal self-conception. So incoherence anywhere can ultimately lead to shame. However, I hope to show that the connection between shame and coherence is what makes shame valuable, because through shame we become painfully aware of incoherence and this enables and even motivates us to develop greater coherence. As such, it plays a valuable role in our moral lives.[54]

[54] J David Velleman sometimes links integrity and coherence in interestingly similar ways. In *Self to Self,* p. 4, he says "Finally, a person's self-image is the criterion of his integrity, because it presents how his various characteristics cohere into a unified personality, with which he must be consistent in order to be self-consistent. Failures of integrity threaten to introduce incoherence into the person's conception of who he is; and in losing a coherent conception of who he is, the person may feel that he has lost his sense of self or sense of identity."

Chapter Three: Critical reflective self-assessment

Introduction

Throughout this thesis, I have stressed the relationship between shame and the self. In chapter one, I described shame as a painful emotion involving a negative self-assessment, in which one is both the assessor and the object of assessment. I also alluded, there, to the complicated task of developing a conception of the self which can make sense of the self playing all three of these roles simultaneously in shame. In this chapter, I endeavour to develop such a model in the course of describing a process of critical reflective self-assessment. This is a type of self-assessment in which one thinks especially deeply and seriously about who one is and who one wants to be. I do not want to deny the existence or importance of other types of self-assessment. Indeed, the self-assessment that accompanies shame is often neither critical nor reflective in the ways I describe in this chapter. However, I believe that critical reflective self-assessment can and often should accompany shame. Firstly, it is not true that shame always involves only unmediated, unreflective self-assessment. (In chapter one, I hinted that this is perhaps a mistake which plays a central role in mistaken conceptions of shame such as the one I attribute to John Deigh. The example of shame I described in that chapter involving a diary is intended as a corrective example.) Secondly, even cases of shame that do initially involve an unmediated, unreflective self-assessment often lead to the more critical and reflective self-assessment I shall be discussing here. In chapter four, I shall point to the different ways in which different kinds of self-assessment can be valuable, but here I wish to focus on critical reflective self-assessment.

I said in chapter one that self-assessment involves an assessor, an object of assessment, and a standard of assessment against which the object is assessed. In this chapter, I shall attribute these three elements of self-assessment to three facets of the self. Critical reflective self-assessment exemplifies my ability to reflect on who I am and who I would like to be: I am doing the assessment, I am being assessed, and I am using some ideal conception of myself as the standard against which I assess myself. For clarity, I shall discuss each of these facets of the self in turn, but it is the interaction between them in critical reflective self-assessment that is particularly important, and I shall argue that each facet of the self can only be properly understood in relation to each other. I begin with the facet of the self that acts as the standard of assessment: the ideal self-conception.

Ideal self-conception

Everyone has some idea of who they would like to be, of what kind of person they would like to become. In chapter one, I spoke about this in terms of an ideal self-conception, and I emphasised

the important role that values play in this self-conception. Values are closely tied to importance. To say that I value honesty more than table manners means that I take honesty to be more important than table manners. Of course, valuing honesty *more* than table manners is perfectly compatible with valuing table manners too. As I discussed in the Introduction, I am happy to talk about values in a very broad way, which would allow me to talk about valuing my table manners without implying that they are very important to me or that they should be very important to anyone else. On this understanding, to value something is simply to place some positive value on it, and I place varying amounts of value on a huge variety of things, from honesty to table manners to someone pronouncing my name correctly.

However, it is also true that when I say in an everyday discussion that honesty is *one of my values* it seems to imply more than that I simply place some positive value on honesty. Rather, I am saying that I place a high value on honesty, that honesty is a very important value to me. And in the context of thinking about the kind of person I want to be – an honest person, for example – this may seem like the more obvious use of the term "value". Now, if we understand value in this way, as identifying only our most important values, then our ideal self-conception seems to involve more than just values, because a fully rounded self (ideal or not) consists of more than just our most centrally important values.

If you ask children what they want to be when they grow up, they might tell you that they want to be a farmer, a fire fighter, or a chef. Some discussion will often reveal values behind their choice: a child may see fire fighters as brave heroes. Children also frequently admit that they want to be like someone in particular, like a parent or some other actual or fictional hero. While there will no doubt be important values embedded in these choices, too, a child takes on board more than these when appropriating these figures as ideal (and idealised) self-conceptions. This is obvious in the way children mimic not just the fire fighter's bravery, but also her clothing.

Children are liable to change their ideal self-conceptions often. One day they will be praised for their bravery and see themselves in the mould of a fire fighter, the next day they will be praised for their voice and see themselves as a future pop star. Adults should have a more settled – though hardly static – self-conception. We, too, have our heroes who inform our ideal self-conceptions, and we are still liable to adopt and mimic far more than archetypal values when attempting to imitate them. There need not be anything wrong with this. Just as the child's cheap plastic fire-fighting helmet may give her the courage she needs to walk down the dark corridor, so my Madiba shirt may give me the courage I need to deliver a stirring political speech on campus.

This short discussion has already raised some of the features of ideal self-conceptions I wish to look at. These features are closely related, but for the sake of clarity I shall list them now and then

discuss them in consecutive order: articulation, coherence, authority, stability, specification and instantiation. I hope that a discussion of these features of ideal self-conceptions will give more content to the notion of an ideal self-conception. Then I shall describe how an ideal self-conception with the features I have described can function as a standard of assessment in self-assessment.

I expect that most people would have difficulty fully articulating their ideal self-conception. By this I mean, to put it too simply, that most people could not describe their ideal self-conception clearly and comprehensively. For most of us, there are large grey areas in our ideal self-conceptions. However, as we grow up we are expected to develop a better idea of who we want to be. Our struggle for greater coherence in our ideal self-conceptions is closely related to articulation. It is often through the process of articulation that we uncover conflicting elements of our ideal self-conceptions and attempt to resolve these conflicts.

Our ideal self-conception functions as a guide for our own personal and moral development. As Aristotle says, "if, like archers, we have a target, are we not more likely to hit the right mark?"[1] A more clearly articulated and coherent ideal self-conception will function better as a guide, partly because it will have more authority. By authority I refer to the degree to which my ideal self-conception can affect my motivations and so my action. For example, imagine that while reflecting on whether or not to lie on a particular occasion, it occurs to me that honesty forms a central part of my ideal self-conception, and that telling the truth on this occasion is the honest thing to do. If my ideal self-conception has no authority, it is unlikely that this realisation will motivate me to tell the truth.

In the example above, we can see how authority relies on coherence and articulation. If my ideal self-conception is incoherent because I hold both honesty and, say, deception as important values and I have no strategy for reconciling the conflicting courses of action towards which these values move me, then my ideal self-conception cannot motivate me to act.[2] The realisation that my ideal self-conception contains values pertinent to the situation is clearly predicated on an act of articulation.

This is a further reason for seeing the moral life as a quest for coherence, as I suggested in chapter two. Authority is important for personal and moral development, because it is often hard to be guided by our ideal self-conception in the face of immediate temptations and challenges. Even in the process of critical reflective self-assessment, my ideal must be able to stand up to some complaints. When I feel that my ideal self-conception is too demanding because it requires me to have courage and I do not feel very courageous, it is important that I do not simply give up and drag my ideal self-conception down to within easy reach.

---

[1] Aristotle, *Nicomachean Ethics*, Book I, Chapter 2, p. 4.

[2] Explaining what makes certain values authoritative is beyond the scope of this thesis.

The authority of my ideal self-conception is also dependent on stability. If my ideal self-conception changes too easily and frequently, it functions as an annoying moving goalpost rather than a stable goal with which I can identify and towards which I can strive. However, my ideal self-conception must be open to revision in at least two ways. Firstly, if we take internal coherence seriously, conflicting elements of our ideal self-conceptions will often need to be resolved.[3] Secondly, legitimate challenges to my ideal self-conception may come from facts about myself. I may want to save the world by becoming a doctor in a rural part of Africa, but have to acknowledge that my ineptitude in science and new languages makes this simply impracticable for me. Or I may realise that being a moderate social drinker is not a live option for me because I am an alcoholic. So, while it is important for my ideal self-conception to be stable, too much stability is problematic if it is a consequence of not being open to change.

Specification and instantiation are modes of articulation. By specification I mean an explicit description of how (an aspect of) one's ideal self-conception applies to a particular situation. I mentioned a simple example of this earlier, where a specification of an aspect of my ideal self-conception – my honesty – yielded a recommendation of a particular course of action in that particular situation – tell the truth. By instantiation I mean the representation of (aspects of) an ideal self-conception by an instance. Earlier I spoke about how actual and fictional heroes inform our ideal self-conceptions. This is often because they instantiate aspects of our ideal self-conception: the fire-fighter instantiates bravery, Mohandas Karamchand Gandhi instantiates pacifism (and bravery, etc).

It may be helpful to apply both specification and instantiation to the same example. Imagine that I would like, two years from now, to be someone who has a steady job, exercises regularly, speaks at least two languages fluently, and is not awkward when meeting strangers. This intermediate goal self-conception is a specification. My ideal self-conception might include values like security, health, diversity and community. I have distilled these various values into a fairly explicit description of (at least some aspects of) the person I would like to be in two years from now. If I achieve my goal and fulfil this specification, then I will have instantiated these values in these ways. Both the specification and the instantiation are modes of articulating my ideal self-conception.

I want to argue that we use our ideal self-conceptions as a standard of assessment in self-assessment, and that it makes sense to talk about falling (too) far short of our ideals. Indeed, I suggested in chapter one that this is what is involved in shame: the feeling that we have fallen short, perhaps hopelessly short, of our ideal self-conception. However, it might be argued that an ideal self-conception cannot act as a standard of self-assessment because it is a set of unattainable ideals that cannot, and should not, be applied in the real world. The objection breaks into two questions:

[3] See my discussion of coherence and integrity in chapter two.

Are we not simply setting ourselves up for inevitable failure by using an impossible measure as a standard of self-assessment? And, how can we understand what it means to have fallen short of an abstract ideal like "pacifism"? I shall address the second question first.

Specification and instantiation provide perhaps the clearest and simplest examples of how we can understand what it means to fall short of an ideal. We can compare ourselves and our actions to a specification or instantiation of (an aspect of) our ideal self-conception. For example, imagine that I come to see my violent reaction to police aggression at yesterday's political march as a failure to follow Ghandi's example precisely when it was most important to do so (specification and instantiation). Or perhaps I simply failed to fulfil a promise I had made to myself to engage peacefully in the march despite anticipated police aggression (specification).

I hope that these examples make it clear that my ideal self-conception can act as a standard of self-assessment through specification and instantiation. However, my ideal self-conception need not be articulated so clearly to fulfil this function. Remembering my dishonest actions in the past, I may reflect on myself and feel painfully aware of how far short I have fallen of my ideal self-conception as an honest person. I need not have a clearly articulated expectation of myself, or of which occasions of dishonesty would lead to shame and which would not. One factor will no doubt be a matter of degree – in everyday conversation we distinguish between actions that are only slightly dishonest and actions that are terribly dishonest. However, with regards to shame, a far more important factor seems to be the link between action and character. I shall discuss this in more depth in the following section, but let me explain very briefly what I mean here.

Some actions result largely from the circumstances in which the action is performed, while others flow more directly from our long-term dispositions. Shame is particularly likely to attach to the latter, to those actions that appear to reveal deep and serious flaws in our characters. The depth and seriousness of these flaws is determined by their relation to our ideal self-conception. Some aspects of our ideal are more important than others. Suppose that I would like to be honest and have good table manners, but being honest is far more important to me than having good table manners. Then my dishonest action is more likely to occasion shame than my *faux pas* at dinner. Furthermore, let us suppose that I suspect that I have a strong disposition to be deeply dishonest, while I feel that I generally have quite good manners. Again, my dishonest action is more likely to occasion shame on account of what I see as a strong connection between that action and an aspect of my character that falls far short of my ideal self-conception, while I might write my *faux pas* off as an anomaly.

Turning now to the first question raised above: are we simply setting ourselves up for inevitable failure by using an impossible measure (an ideal self-conception) as a standard of self-

assessment? There is obviously a sense in which we always risk failure whenever we challenge ourselves, but this question is not about the possibility of failure, but rather about the possibility of success: even if I could articulate exactly what the ideally honest person is like, could I possibly be that person, or am I doomed to fail?

Again, one tactic is to draw on specification and instantiation to bridge the gap between our everyday expectations of ourselves and our ideals. In the example above, responding in a nonviolent manner towards police aggression at a political march does not require a superhuman adherence to pacifism. Perhaps the true pacifist would respond by giving the police flowers, but the most I feel I can expect of myself is that I limit my response to verbal abuse. In some situations we may articulate the requirements of our ideal self-conception in a way that is too demanding, or not demanding enough, or we may not be able to articulate them very clearly at all. There is no easy rule for getting it right. And, again, my ideal self-conception need not be clearly articulated in order to function as an ideal self-conception. I can feel that I have fallen too far short of my ideal of pacifism without having a concrete idea of what pacifism requires in that particular situation. This can make it even more difficult to know whether your expectations of yourself are too demanding or not demanding enough.

So, to answer the first question above: yes, we do set ourselves up for almost inevitable failure by using an impossible measure (an ideal self-conception) as a standard of self-assessment. Even when we come quite near to instantiating our ideal self-conception, there is usually room for improvement. However, we have a sense of how much we can expect from ourselves. This is the nature of personal goals in general. Imagine that I plan to finish this chapter by the end of the month. At the end of the day, how do I decide whether I have done enough? While my goal is to finish, I obviously have not failed by not finishing today. Do I divide the chapter by thirty? Do I take into account other commitments, mood swings, weekends, etc?

I am arguing that we use our ideal self-conception as a standard of assessment. I am not arguing that this is straightforward. I raised two difficulties with this, which might clumsily be referred to as “demandingness” and “articulatability”. I have acknowledged that these are real difficulties that we encounter when we engage in critical reflective self-assessment, but I have also argued that we can overcome these difficulties. Specification and instantiation are useful tools here, but ultimately there is no easy way to get it right.

On the view that I am sketching of critical reflective self-assessment, we may see three facets of the self interacting, corresponding to the three elements of self-assessment: the assessor, the object of assessment and the standard of assessment. I have been discussing the ideal self-

conception that acts as the standard of assessment. Next, I shall turn to character, which acts as the object of assessment.

Character

In chapter one, I argued that the object of self-assessment in shame is always the self. In this chapter, I am describing critical reflective self-assessment. I have called it *self*-assessment, so the object of assessment is clearly the self, but in this section I shall be more specific about which aspects of the self form the central focus of this kind of self-assessment. Some people complete a "self-assessment" in the course of filing their income tax return, during which they assesses their own income and expenditure. Income and expenditure are not relevant to the kind of self-assessment I am describing. I shall identify the object of assessment in critical reflective self-assessment as one's character, and I shall discuss which aspects of oneself I consider to be part of one's character.

Self-assessment would be unnecessary if one already knew everything about oneself. However, self-knowledge is not something one can take for granted – indeed, perfect self-knowledge may well be impossible.[4] In one sense, it is usually safe to say that I know who I am: I know my name, appearance, hometown, etc. Nevertheless, it is also perfectly common to speak about self-discovery, about coming to know something about oneself that one did not know before. Self-discovery often occurs by observing how one reacts to a situation. This is one of the attractions of placing oneself in new and challenging situations: they present opportunities for self-discovery because one can observe how one responds to them. Not all self-discovery is pleasant – sometimes one discovers things about oneself that one does not like. In chapter one, I suggested that this kind of self-discovery is often at the heart of shame.[5]

Peter Goldie's excellent book, *On Personality*, is largely about character. After quoting an excerpt from Augustine's *Confessions*, Goldie explains that

> [Augustine] sees the way he has been, and the way he is, and he feels guilty and ashamed of it; and he sees how it is possible to change, to change into a better person. What Augustine recognised, and was ashamed of, was the long-term dispositional structure of his motivations ... Whilst each particular wrong act would have its own particular motives , these motives were expressions of what he recognised and was ashamed of – his *character*.[6]

Goldie identifies the object of Augustine's shame as his character. He highlights the distinction between the particular motives behind particular wrong acts, and the "long-term dispositional structure" of one's motivations.

[4] Damian Cox *et al* argue that it is not even desirable, *Integrity and the Fragile Self*, p. 5.
[5] Although, as I noted there, the discovery may be a rediscovery.
[6] Peter Goldie, *On Personality*, p. 110. I have retained Goldie's italics.

There is a lot to learn by examining the circumstances surrounding a particular action.[7] I may come to see that watching violent movies until late last night, skipping breakfast this morning and then kicking my toe on the doorframe contributed to my passive aggressiveness during this morning's seminar. Perhaps in future I should avoid watching violent movies and skipping breakfast on school nights. But Augustine's focus, Goldie says, is on "the way he has been, and the way he is"; Augustine is focused on those aspects of himself that display some constancy – his character traits. Having a disposition to respond passive aggressively to criticism is a character trait. One reason to focus on character traits in critical reflective self-assessment is their constancy. Character traits are not only an important part of who I am now, but also who I will be tomorrow. And character traits will affect how I behave in future. So focussing on character traits can give me a greater understanding of what and how I would need to change in order to be a different person and act differently in the future.

Focussing on a character trait is particularly useful if one places one's character trait in the context of one's other character traits, and attempts to see how these traits might constitute a coherent whole, which can then be held up against one's ideal self-conception. Applying this to the above example, I could see my passive aggressiveness in the context of building a unified and coherent character, and therefore in the context of my ideal self-conception. I dislike passive aggressive behaviour and see it as an obstacle to traits that I value: affability, kindness, sincerity, etc. These are important elements of my ideal self-conception. So, identifying with my ideal self-conception, I can see that I do not want to be passive aggressive.

I have been discussing character traits. Goldie's book is *On Personality*, so I need to be clear about the difference between personality traits and character traits. I see problems with Goldie's criteria for drawing a neat distinction. Quoting the Oxford English Dictionary (as Goldie does), personality is the "quality or collection of qualities which makes a person a distinctive individual."[8] This is clearly a very broad definition. Goldie presents character traits as a subset of personality traits, and the criteria he suggests for distinguishing character traits are stability, consistency, reason responsiveness and depth. I shall discuss each of these criteria in turn, raising problems with their application along the way. While I shall argue that these criteria fail to generate a neat distinction, I find them useful in other ways and I hope that the value of discussing them at length will become clear. If I am right that a neat distinction cannot be drawn, this calls into doubt my use of "character" rather than "personality" throughout this chapter. After discussing Goldie's criteria, I will address this worry, too.

---

[7] See Peter Goldie's entertaining discussion, *On Personality*, pp. 60-64.

[8] *Oxford English Dictionary Online*, accessed 12 May 2009 <http://www.oed.com>. Goldie uses the 2002 print edition.

Stability is consistency over time, and consistency applies across different situations. So a truthful person tells the truth yesterday, today and tomorrow, and she tells the truth to her friends, her business partners and to the stranger buying her used car. A character trait is also reason responsive, it "involves a disposition reliably to respond to certain kinds of reasons".[9] We see reasons in the situations in which we find ourselves. When I see that an old woman crossing the road is in danger of being hit by oncoming traffic, I see a reason to help her to safety. If you asked me why I ran to help her, I would respond, "Because she was in danger". For me, the fact that someone vulnerable is in danger is a reason (though not always an overriding reason) to help them.

Reliably responding to certain kinds of reasons in certain kinds of situations reveals one's character. I might say, for example, that Celia is overambitious, and mean that in situations where Celia can get ahead, she usually will. She will see getting ahead as an adequate reason for her actions, even at the expense of others. She will abandon her old mother in the middle of the road if her mother is making her late for her job interview.

Character traits are reason responsive in another way, too: they are open to change by reasons. Celia is unlikely to stop being overambitious overnight, but upon reflection (especially the kind of critical reflective self-assessment I am discussing here) she may decide to work towards being less ambitious, or at least putting her ambition within the context of her relationships and other values. This would, in turn, lead her to respond differently to the same situation – to see her responsibility for getting her mother across the road safely as an overriding reason for doing so. Focussing on those aspects of oneself that are open to change puts one in a better position to change those aspects and thereby to become the person one is striving to be.

Goldie admits that some personality traits are reason responsive in both senses. His example of a trait that is not reason responsive is a bad memory: "I might be ashamed of my bad memory, and you might mock me for it ... but guilt and reproach are inappropriate here."[10] (It is interesting to note in passing that Goldie finds shame appropriate towards an object that he is describing as outside the realm of direct responsibility.) It seems to me that his example fails.

A bad memory is reason responsive in both ways. This is clear from an example Goldie gives in another context.[11] Jane asks her husband, Rob, to buy four items at the supermarket and correctly predicts that he will forget the toothpaste. Goldie says that we should not be surprised if Rob has no trouble remembering the football scores. His point is that we do not expect personality traits to be consistent across situations in the same way that we expect character traits to be. My point is that Rob's memory is reason responsive in both the senses discussed above. Rob can

[9] Peter Goldie, *On Personality*, p. 13.
[10] Peter Goldie, *On Personality*, p. 89.
[11] Peter Goldie, *On Personality*, p. 68.

remember the football scores because they are more important to him. He could remember the items on the list if they were more important to him (he would not forget item number one, "beer"). And he could do memory exercises or use rhymes and mnemonics if he wished to become better at remembering lists.

So stability, consistency and reason responsiveness are not always adequate criteria for distinguishing character traits from the broader set of personality traits. What about depth? Goldie says that "character is deeper, personality more superficial, concerned with the surface".[12] By this he means that "someone's personality traits are only good conditionally upon that person also having good character traits", but not vice versa.[13] For example, Jane's exceptional memory is only good because she is kind, and so she uses it to remember people's names and birthdays and so on and in this way makes people feel special. If Jane only used her memory to be nasty, to remember each of her friend's deepest insecurities and frequently pick on them, then her memory would not be a good trait.

When character traits affect personality traits in this way, Goldie talks about one trait "polluting" another. But is he right to say that character traits can pollute personality traits, *but not vice versa*? There is a sense in which character traits rely on personality traits. Jane's kindness is good whether she has a good or bad memory, but a really bad memory would interfere with her kindness, leaving her unable to remember people's names, birthdays, insecurities, illnesses, etc. Gloominess is another of Goldie's examples of a personality trait, and it seems obvious that deep and persistent gloominess will interfere with many good character traits. So, while it interesting to explore the ways in which traits – those of personality and character – affect one another, this does not provide a definitive criterion for distinguishing character traits from the broader set of personality traits. It is unclear to me how thinking about "depth" would help me to decide, for example, if stubbornness is a personality trait or character trait.

I have been critical of all the criteria Goldie has suggested – stability, consistency, reason responsiveness and depth – but all can be useful if we understand them as properties that traits can display to varying degrees. This allows us to see traits on a sliding scale: where traits display all these properties to a greater degree, we more properly call them character traits, and where traits lack these properties, we call them personality traits. So while some traits fall clearly at one or the other end of the spectrum, there will be many traits that lie somewhere in between. There is no obvious rule for making a clear distinction, for drawing a boundary line around the cluster of character traits in the centre that separates them from the broader cluster of personality traits.

[12] Peter Goldie, *On Personality*, p. 13.
[13] Peter Goldie, *On Personality*, p. 32.

I do not think that this conclusion threatens my description of critical reflective self-assessment, or even my choice of "character" as the object of this assessment. I agree with Goldie that the focus of this assessment is rightly one's character, but I hope to have shown that the borders of one's character are blurry. Deciding which of one's traits are most central to one's character is part of the process of critical reflective self-assessment. The criteria Goldie suggests are indeed useful, but they are not definitive. They help one to locate the focal point of critical reflective self-assessment as one's most central character traits without defining an outer limit of character that would rule other personality traits out of bounds or beyond the scope of assessment.

One of the reasons I have discussed Goldie's criteria is because they are particularly useful, not as criteria, but as ways of understanding our traits. Elizabeth gets very angry when people use her laptop without asking. It is useful to know if she always gets angry when this happens (stability), whether she also gets angry when people use other things of hers (consistency), and whether her quickness to anger pollutes other traits like being a patient teacher (depth).The answers to these questions will play an important role in guiding Elizabeth's attitude towards this trait and her approach to changing it, if she decides to do so.

Sometimes we move too quickly to infer character traits from actions, at other times we stubbornly refuse to acknowledge character traits. There is an important balance to be achieved here, with dangers on both sides. On the one side, there is the danger of exaggerating the importance of character traits, both in understanding one's actions and in one's self-assessment, to the point where they push everything else out. So one might feel pressed to accept, having acted dishonestly on a number of occasions, that one is dishonest. Here one has moved too quickly from action to character.

On the other side, one may refuse to move from action to character at all, or one may simply move one's assessment to focus, not on one's character, but on one's ideal self-conception. In the first case, we find the familiar excuse, "I know I lied on that occasion, but that was different, I am not a liar, I don't do that sort of thing". I have already acknowledged that combinations of circumstances can be very important in explaining our behaviour in particular situations. However, it is often all too easy to avoid admitting to genuine character flaws in this way. In the second case, I shift my assessment from my character to my ideal self-conception. We focus on who we would like to be rather than who we are and say, for example, "I would really like to be diligent, but it's just not in my nature; I'm lazy." Again, there is some truth behind this response. I mentioned earlier that there would be occasions for revising my ideal self-conception in the light of facts about myself that I really cannot, or should not, change. However, once more I suspect that it is often all too easy to avoid the hard work of changing our bad character traits by simply giving up on them in this way.

I do not wish to get into a long and complicated discussion of responsibility. My point here is a small one: in order to learn from our bad acts, we first need to take ownership of them. I have been discussing one way in which we can do that, by acknowledging that at least some of our bad acts flow, at least in part, from bad elements of our characters. This point need not be in conflict with acknowledging that some of our bad acts flow, at least in part, from other sources. And I am suggesting that the kind of critically reflective self-assessment that I am describing focuses on my character in the light of, on the one hand, what my acts seem to tell me about my character and, on the other hand, what I would like my character to be, my ideal self-conception. I shall now look at the reflective facet of the self, the "I" that acts as the assessor in critical reflective self-assessment.

Reflective self

It is worth repeating what I said at the beginning of this chapter: I have separated out three facets of the self in the hope that this will provide greater clarity, but it is important to keep in mind that I am only talking about one self. This is particularly pertinent when thinking about this facet of the self, the "I" that engages in critical reflective self-assessment.

My model of the self describes self-assessment as an interaction between three facets of the self. Someone with a Freudian background might be tempted to compare elements of my model with elements of Sigmund Freud's model. I shall not assess the merits of such a comparison in any detail, but there are some connotations such a comparison might suggest that I wish to explicitly reject. Specifically, the Freudian I am imagining might compare my ideal self-conception to Freud's Super-Ego or, in his later work, the Ego-Ideal. Gerhart Piers argues convincingly that, in the context of shame, it is the Ego-Ideal that it relevant. [14] The role of my reflective self might be compared to that of Freud's conscious Ego.

On my model of critical reflective self-assessment, my reflective self assesses my character in the light of my ideal self-conception. On the classical Freudian model, the Ego arbitrates between the Super-Ego and the Id. But on Piers' development of the Freudian model, the Ego arbitrates between the Ego-Ideal and the Id. Freud's Super-Ego is a patriarchal internalised authority figure, and the Ego-Ideal is not much better for my purposes.[15] Both connote a set of restrictive rules and boundaries rather than attractive values and ideals – we are driven by the Id, we are merely kept in check by the Super-Ego/Ego-Ideal. In contrast, on my model we may be driven by ideals and attracted to values. The background assumption on Freud's model seems to be that those rules and boundaries are internalised in a passive way. On my model, while I have noted that the elements of

[14] Gerhart Piers and Milton B Singer, 'Shame', p. 147.
[15] Sigmund Freud, 'Origin of the Sense of Guilt'.

one's ideal self-conception come from the world, I have described the ideal self-conception in a way that also acknowledges one's capacity to take control of it, to choose to adopt, revise or reject elements of it.

Most importantly for this section, the comparison between the roles of the reflective self and the Ego suggests moderation between two extremes: an over-controlling Super-Ego/Ego-Ideal and an instinctual, libidinal Id. On my model, my reflective self may sometimes come up with restrictions, with the goal of bringing my character into greater coherence with my ideal self-conception. But this is a collaborative process, because my character aspires to instantiate my ideal self-conception, and my reflective self reflects on how this aspiration can be fulfilled. In chapter four, I shall suggest that the image of a golfer examining her swing more accurately captures this relationship.

I used the word "I" above to describe the reflective self because one usually refers to this facet of one's self in the first person. I have already referred to the temptation to separate oneself from one's acts, and even from the character that one might see as responsible for those acts. This is particularly true of actions from the more distant past. One's character changes over time, and so if enough time has passed since one committed a bad act, it is tempting to attribute it to the character of one's past, and dissociate that past character from one's present character. However, it is more difficult to do the same with one's reflective consciousness, to disown one's past reflections in the same way that one disowns one's past actions. It feels like it is inescapably me that reflects and assesses, that engages in critical reflective self-assessment.

Let me call this "identification" – the feeling that a particular facet of the self is "I" at a given moment. My claim that I identify more strongly with my reflective self can be taken in a weaker and a stronger sense. The weaker claim is merely that when I am engaged in critical reflective self-assessment, I tend to identify more closely with my reflective self because that is the facet of the self that is active, that is engaging in assessment. The stronger claim is that I always tend to identify more closely with my reflective self, even outside of the activity of critical reflective self-assessment. I think that both claims are true, but here I want to discuss the danger of this tendency in the context of critical reflective self-assessment, so the weaker claim is more pertinent.

There is a danger in identifying too closely with my reflective self, and distancing myself from the facet of the self I am assessing, my character. Firstly, it is dangerous because, identifying with my reflective self, I am likely to ignore the fact that my character affects my assessment, including my critical reflective self-assessment. In other words, my reflective self may display the same character flaws as my character. If I am characteristically rash and impulsive, it is likely that my self-assessments will also tend to be rash and impulsive. By identifying simultaneously with both my

reflective self and my character, by bearing in mind that it is the same self assessing and being assessed, I am in a better position to acknowledge and compensate for character flaws that would otherwise interfere with my assessment.

Secondly, identifying too closely with my reflective self and distancing myself from my character may encourage me to respond in a restrictive and authoritarian way to the flaws I find in my character. I suggested above that this danger may be exacerbated by some connotations of a Freudian model of the self. I shall return to this in chapter four. For now, I would like to suggest that identifying simultaneously with my reflective self and my character, and bearing in mind that my character flaws are my own, is likely to lead to a more constructive response to these flaws. And here too, an authoritarian response that negatively focuses on rules and restrictions may inspire rebellion next time I am in a situation where I am faced with temptation and challenge. As one switches roles, identifying at one moment with one's reflective self and at another moment with one's character, one simply rejects the assessments and decisions one has made while adopting the previous role. In this way, one struggles to develop a stable approach towards oneself. A more positive approach involves taking ownership of one's character flaws during self-assessment, acknowledging that they are mine, and focusing on the value one places on one's ideal self-conception. This way, one is more likely to realise that, for instance, wanting to be an honest person means wanting to act honestly, or at least wanting to want to act honestly.[16]

I have been discussing the danger of identifying too closely with my reflective self and distancing myself from my character. At the other extreme lies a danger in identifying too closely with my character and not allowing enough distance between my reflective self and my character. Some distance is required to make an assessment at all. Beyond this, while I have suggested that it is important to acknowledge the likelihood of character flaws being displayed in one's reflection, the distance between my reflective self and my character also makes it possible to transcend these flaws to some degree. In my everyday life, I pursue a rich variety of different and sometimes conflicting goals and experience the pull and push of numerous temptations and challenges. When I step back and engage in critical reflective self-assessment, many of these fade to the background, enabling me to focus more clearly on my character and on my ideal self-conception.

In chapter one, I described a situation in which my failure at some display of macho male bonding occasioned irrational shame – irrational, because my negative self-assessment did not make sense in the light of my values or standard of assessment. It should now be clear how my reaction may be connected to a character flaw. We can imagine that I have an excessive tendency to seek the approval of others and that, in this situation, this contributes to my uncritical adoption of the

[16] See my discussion of Frankfurt in chapter two.

audience's negative assessment of me. If we further imagine that the audience is made up of my rugby teammates, and that this situation occurs in the locker-room, then we can easily imagine how distancing myself from my character can help me transcend my flaws. While I am in the locker-room seeking my teammates' approval, I focus exclusively and excessively on what they think of me, and I ignore my other values and their relationships to this one – my character flaw is precisely my tendency to do this. Distancing myself from the situation allows me to distance myself from the pursuit of approval, which enables me to acknowledge my other values and evaluate my pursuit of approval in the light of these other values. And evaluating my tendency to seek approval enables me to distance myself from it, acknowledge it, and even take steps to compensate for it, and thereby possibly transcend it.

I started this discussion about the reflective self by reiterating how important it is to remember that all three facets of the self are merely facets of one self. I have been discussing the importance of identifying with both one's reflective self and one's character, but identification is no less important concerning the other facet of the self, one's ideal self-conception. It might seem odd to talk about identifying with my ideal self-conception given the way I have been talking about identification. My ideal self-conception does not act in the way that my character and my reflective self act. But there is a sense in which I need to "identify" with my ideal self-conception, especially when I am engaged in critical reflective self-assessment, which might be referred to rather as associating oneself with, or feeling a special affinity for, one's ideal self-conception. While I cannot watch the process of self-assessment from the perspective of my ideal self-conception, my reflective self should feel itself strongly allied to my ideal self-conception.[17]

So once again, I am recommending a balance between identifying too strongly with and distancing oneself too much from any facet of the self. In critical reflective self-assessment occasioned by shame, I suspect that one more often identifies too closely with one's reflective self, because one's character is tainted by shame. Perhaps the best way to describe the balance is to say that we need to find a way of feeling both that we are *looking at* and *being looked at*. Most often, we seem to do this by moving between the two perspectives. This process is sometimes described as an inner dialogue, and this image is helpful. However, what I suggest that we should try to adopt both perspectives at once, we should try to identify with both voices in the dialogue at once, and we should never lose sight of the fact that both voices belong to one self.

[17] My thinking about identification has been influenced by reading Marilyn A Friedman, 'Autonomy and the Split-Level Self'.

Interaction

I have described how critical reflective self-assessment can be seen to involve three facets of the self which I called the ideal self-conception, character and the reflective self. I stressed throughout the unity of all three facets of the self and drew attention to various ways in which these facets of the self interact. Indeed, the critical reflective self-assessment I am describing in this chapter is itself an interaction between the three facets of the self. The primary interaction here is between the reflective self and character. I suggested that the image of a patriarchal internalised authority figure punishing an aspect of the self for transgression is inappropriate here. Rather, the reflective self should aim to achieve a careful balance between identification and transcendence.

In an experience of shame, I may feel utterly crushed, a failure. On reflection, I will compare my character to my ideal self-conception to make sense of my emotional response. Of course, my shame may not make sense in that it may not be related to my self-conception in a coherent way. Imagine, again, that I failed at some display of macho male bonding, but on reflection I realise that macho male bonding has no place in my ideal self-conception, that I place little or no value on it. On my conception of shame, described in chapter one, my shame is irrational.

There are a number of ways in which I can respond to irrational shame. My emotions can draw attention to values I have unwittingly and so tell me important and interesting things about myself. In this way, even irrational shame can be valuable. I shall discuss this further in chapter four. The pertinent point here is that, following irrational shame, I may wish to revaluate my ideal self-conception. Perhaps I need to acknowledge that macho male bonding is more important to me than I have been prepared to admit, or that it is a value that is inextricable from the constellation of values that I already hold. Here, my reflective self changes my ideal self-conception, sometimes in light of my character.

If my shame does make sense in light of my ideal self-conception, then my character is at odds with my ideal self-conception. This is what I shall call accurate shame: my character has a flaw (for example, I am dishonest) and this conflicts with an element of my ideal self-conception (honesty). In response, my reflective self should attempt to change my character in light of my ideal self-conception.

As I have described them, all three facets of the self are open to change by one another, at least insofar as the reflective self is informed and influenced by both the ideal self-conception and character, and makes changes to the ideal self-conception and character on such information and under such influence. While I am discussing interactions, I want to avoid the thought that these facets of the self only interact with each other. The self is not sealed off from the world. I began my discussion of the ideal self-conception by talking about various sources for children's unstable ideal

self-conceptions. Our initial character traits are learned and habituated as we grow up in a process that is largely beyond our control. I take it for granted that the self in all its facets develops in the world and in response to the world.

Conclusion

In chapter four I shall suggest that shame can lead to the kind of critical reflective self-assessment I have been describing, and that this plays a valuable role in personal and moral development. In this chapter, I hope to have sketched a picture of what I mean by critical reflective self-assessment. I have done this by separating out three facets of the self, in the hope that this provides some clarity without implying an unnecessary multiplication of entities. Self-assessment necessarily involves an assessor, an object of assessment and a standard of assessment. I have tried to show the value (and danger) of attributing these three roles to three facets of the self.

Chapter Four: Shame leads us back to ourselves

Introduction

Shame can make sense and be valuable to us through its role in promoting coherence and thereby contributing to our moral lives. Put simply, the moral life is a project of trying to be or to become a certain kind of person who lives life in a certain way. In chapter two, I showed how this project can be understood as the quest for coherence in various ways. The moral life is a project in which one can engage with more or less success – one can achieve greater or lesser degrees of coherence. I am arguing that shame can help one to engage more successfully in this project by promoting coherence.

This chapter is divided into two sections. In the first, I shall focus on the value of shame as an emotion of self-assessment. I shall argue that shame is valuable because of its relationship to the process of critical reflective self-assessment, and that this process is valuable because it promotes coherence. I shall build on my description of critical reflective self-assessment developed in chapter three, and highlight features of that process that can be particularly conducive to promoting coherence and thereby contributing valuably to the moral life.

In the second section, I shall focus on the connection between emotions and values. I shall develop the intuition that valuing, for example, honesty implies responding emotionally to the presence or absence, promotion or abuse of honesty. I shall argue that shame is a manifestation of the connection and that it plays a valuable role in one's moral life by turning one's attention towards one's values and motivating one to promote one's values. In this way shame promotes coherence between one's character and one's ideal self-conception.

Self-assessment

In this section, I focus on the relationship between shame and critical reflective self-assessment. For clarity, I shall use paradigmatic shame as my model. By paradigmatic shame, I mean an episode of shame consisting of three elements: accurate shame, critical reflective self-assessment, and post-reflective shame. I begin by briefly revisiting what makes shame rational and accurate on my conception, and clarifying my distinction between the self-assessment involved in shame and the process of critical reflective self-assessment. I discuss the way shame leads to critical reflective self-assessment, and I describe the features of what I take to be the most valuable critical reflective self-assessment in the context of paradigmatic shame. Finally, I argue that non-paradigmatic shame can also be valuable in similar ways.

On my conception, shame is rational only if my negative self-assessment is coherently related to my ideal self-conception, and my shame is accurate only if my negative self-assessment correctly identifies a mismatch between my character and my ideal self-conception. Accuracy therefore requires that I have knowledge of my character, my ideal self-conception, and the relationship between the two. As such, accurate shame is always rational. In paradigmatic shame, accurate shame then also leads to critical reflective self-assessment.

All shame involves a self-assessment, so I need to be clear about the difference between non-paradigmatic shame that merely involves a self-assessment and paradigmatic shame that also involves a process of *critical reflective* self-assessment. In general, the negative self-assessment that occasions shame may be arrived at in any manner. It may be the conclusion of a process of critical reflective self-assessment, or one may simply adopt it without criticism or reflection from one's peers (as I suggested that Arnica does, in chapter one). For paradigmatic shame, while there are no restrictions on the way in which this self-assessment is adopted, it must be accurate. Furthermore, in the paradigmatic case, shame then leads to a process of critical reflective self-assessment. I described this process in chapter three, and pointed to some parallels between this process and the self-assessment which occasions shame. In chapter one, I mentioned that this self-assessment involves an assessor, an object of assessment, and a standard of assessment and, in chapter three, I attributed these three roles to the three facets of the self involved in critical reflective self-assessment: the reflective self, character and the ideal self-conception.

These structural parallels between the three elements of self-assessment involved in shame and the three facets of the self involved in critical reflective self-assessment are not necessarily any more than parallels. When one's shame involves a reflective assessment of an element of one's character in the light of one's ideal self-conception, each aspect of the self is involved. However, irrational shame may involve an unreflective assessment of an aspect of oneself that is not part of one's character in the light of a value that is not part of one's ideal self-conception. The more nearly the elements involved in an episode of shame mirror those involved in critical reflective self-assessment, the more effectively that episode will facilitate the shift from a self-assessment to the process of critical reflective self-assessment. The most significant parallel here is that between the object of assessment in shame – the self – and the object of critical reflective self-assessment – the character, which is an aspect of the self.

The object of assessment in shame is the self. This is perhaps most clear when we contrast shame with another painful emotion, guilt. The same action may lead to both guilt and shame.[1] Guilt focuses on the action and looks outward from the action to its effects. We may feel guilty simply

[1] Although it is not only actions that occasion shame, as discussed in chapter one.

because we have transgressed a boundary or broken a rule,[2] or because we have harmed somebody. Shame looks inward from the action to the self. In an experience of shame, I am painfully aware that it is I who performed the action, I am painfully aware of the connection between myself and my action. Bernard Williams describes the difference in *Shame and Necessity*:

> We may feel both guilt and shame towards the same action. In a moment of cowardice, we let someone down; we feel guilty because we have let them down, ashamed because we have contemptibly fallen short of what we might have hoped of ourselves. As always, the action stands between the inner world of disposition, feeling and decision and an outer world of harm and wrong. *What I have done* points in one direction towards what has happened to others, in another direction to what I am.[3]

I do not want to suggest that guilt, regret, remorse, and other negative reactions to one's own actions do not play a valuable role in one's moral life, but I do want to argue that shame plays a role that these other reactions cannot. Shame plays a unique and valuable role in our moral lives by leading us back to ourselves, by focusing our attention on who we are and who we can be in a way that no other emotion can. I think this is important because, in contemporary Western culture at least, shame is so often conceptually swallowed up by guilt.[4] When attempting to identify and work through our emotional responses to our own wrongdoing, we more easily identify our guilt and often fail to identify our shame. I think our lack of conceptual familiarity with shame, and our lack of understanding of how to respond to shame, is largely responsible for this. This certainly leads us to miss out on the opportunity provided by shame for personal and moral growth, and may also leave us unable to overcome the shame which still lurks, unacknowledged, in our lives.

I am arguing both that shame *does* frequently lead one to engage in critical reflective self-assessment, and that it generally *should* do so. (I say "generally" to acknowledge that there are some situations in which it should not do so, and I shall discuss these situations later.) The essential focus on the self in shame parallels the essential focus on the character in critical reflective self-assessment, and this parallel is the primary force behind the move from shame to critical reflective self-assessment. This move often begins with a simple question prompted by the painful experience of shame, like "Why am I feeling shame?" or "What am I ashamed of?". Because the object of shame is the self, as soon as one reflects on such questions, one has already started to engage in some form of reflective self-assessment. Furthermore, shame involves a painful conflict between two different elements: one's self-conception and one's ideal self-conception. One is likely to have a critical attitude when one reflects on one's shame, because neither of these elements is easily abandoned,

[2] Sigmund Freud, 'Origin of the Sense of Guilt'.
[3] Bernard Williams, *Shame and Necessity*, p. 92.
[4] For a psychological explanation of this, see Erik H Erikson, 'Autonomy v Shame and Doubt'. J David Velleman connects our shamelessness to our shrinking sense of privacy in 'The Genesis of Shame'.

nor is the conflict between them easily ignored. So one's critical attitude stems from an attempt to simultaneously defend each of these elements that are in conflict. Yet, even when one reflects critically on one's shame, it will not always exemplify the process of critical reflective self-assessment that I described in chapter three. There are more or less valuable ways to engage in such a process. In this section, I am focusing on the valuable role that shame and critical reflective self-assessment in paradigmatic shame can play in our moral lives. The value of this role will be most obvious in episodes of shame that involve the most valuable kind of critical reflective self-assessment, so I shall now discuss some features of what I take that to be, contrasting them with some of the ways in which self-assessment can go wrong.

What I mean when I call a certain kind of critical reflective self-assessment the "most valuable" is that this is the kind of self-assessment which can contribute the most to our moral lives by promoting coherence most effectively. In chapter two, I discussed how coherence is not just *conducive to* becoming the person I want to become, it *means* becoming the person I want to become. The main role of critical reflective self-assessment is that it promotes coherence, and the better it promotes coherence the more valuable a role it will play in our moral lives. I take it to be intuitively obvious that reflecting critically on who one is and who one would like to be is generally likely to help one to become that person. However, reflecting in different ways will have different results, and I shall now look at some features of more valuable critical reflective self-assessment.

Perhaps the most obvious feature of valuable critical reflective self-assessment is accuracy. Critical reflective self-assessment is a conscious, intentional process, through which I identify my flaws and decide how to correct my personal and moral development. I have stressed throughout this thesis the importance of seeing critical reflective self-assessment in the context of a moral life: so I do not just think about who I am and who I would like to be, but go further and think about how to change from who I am into who I would like to be. Clearly, the more accurately I assess who I am and who I would like to be, the more likely I am to think accurately about how to become that person.

Critical reflective self-assessment is a good opportunity to reassess the assessment involved in shame and reject it if it is found to be inaccurate, but the value of this opportunity depends on how effectively we use it. In chapter three, I suggested a number of reasons for thinking that one's critical reflective self-assessment will be more accurate than one's less critical, less reflective self-assessments. In particular, I highlighted the opportunity presented by critical reflective self-assessment to achieve a critical distance from one's character and one's character flaws while taking ownership both of one's flaws and one's values insofar as they constitute one's ideal self-conception. However, there is no guarantee that critical reflective self-assessment will always work

as a corrective process. One can go wrong, and the self-assessment at which one arrives can be inaccurate in a variety of ways. In short, one can simply be wrong about oneself.

Self-deception is especially a concern here. Acknowledging one's flaws is often painful and difficult. Reflecting on oneself can be an opportunity for spinning a web of self-deception around one's flaws as much as it is an opportunity for clearing away such webs. There is a real danger that one will reject an accurate but painfully negative self-assessment when it is reassessed during critical reflective self-assessment, and there is no easy way to reliably avoid this possibility. The best one can do is try to engage in critical reflective self-assessment sincerely and bravely. On the other hand, one can also fail to reject inaccurate assessments when reassessing them through critical reflective self-assessment. In chapter three, I spoke about the common tendency to infer character traits on insufficient evidence, like deciding that one is dishonest after lying on only one or two occasions. Again, there is no easy way to ensure that one's reassessment is accurate. One can only do one's best to consider carefully what all the available evidence tells one about one's character.

Importance is key to accuracy. One must correctly assess the importance of flaws in one's character. In chapter three, I argued that there is no easy rule for deciding which flaws are central to one's character – one feature of valuable critical reflective self-assessment is the accurate assessment of which flaws are central to one's character and which are not. There is a danger that one can fixate on less important flaws, sometimes to avoid focusing on one's more central and problematic flaws. The best strategy here is to try to see all one's traits against the background of the whole character that they constitute, and each corresponding value against the background of the whole ideal self-conception.

I have stressed that critical reflective self-assessment should always be placed in the context of the moral life. So one must not only accurately identify one's character flaws and accurately assesses their importance, one must also respond appropriately to these flaws. One's post-reflective response to one's flaws is influenced, to a significant extent, by one's attitude during critical reflective self-assessment. In chapter three, I suggested that a Freudian model of the self may have unwanted connotations about conflicting extreme elements of the self (the Id) being moderated by another element (the Ego). In contrast, in developing my model of the self, I emphasised the self's unity, and suggested that all three aspects of the self share in the whole self's goal of closer unity in the form of greater coherence. The model on which one understands the self affects the attitude with which one engages in critical reflective self-assessment. If one sees the self as consisting of competing extremes, one may feel that the only choice left to one is to pick sides, or one may feel that the only fair, balanced response is to consistently compromise. Neither of these responses

involves taking ownership in the way that I recommended in chapter three, and I return to this below.

It is difficult to think of an analogy for critical reflective self-assessment based on my model of the self that captures the attitude towards the self that I would recommend, but perhaps that of a golfer examining strobe photographs of her stroke and thinking about how to improve will be helpful. The golfer is carrying out a self-assessment based on her own goals. A manual on how to improve one's stroke will only be of use to her insofar as it helps her to more regularly hit the golf ball with greater accuracy (assuming that is her goal). She cannot necessarily make direct changes to each aspect of her stroke, but this does not mean that she cannot change those aspects at all. In order to change them, she may need to acknowledge facts about herself (she is overweight) and change that or something else (her clubs, perhaps) to improve that aspect of her stroke. Given the effort involved in losing weight and the cost involved in buying new clubs, she may need to think carefully about how important her stroke accuracy is to her – perhaps she values the time on the golf course with friends rather than hitting the ball accurately. If she chooses to lose weight or buy new clubs, she should do so because she wants to hit the ball accurately. In the same way, if critical reflective self-assessment leads one to choose to make changes to one's character, one should make those changes autonomously with the corresponding values in mind, not because one succumbs to some inner instinctual drive or enforces passively internalised rules and boundaries.

The correct attitude to take, then, is one of ownership. In critical reflective self-assessment, *I* assess *my* character in the light of *my* ideal self-conception. Put another way, *I* think about who *I* am and who *I* would like to be. While it is important to acknowledge that it is not a solipsistic project, my moral life is inescapably mine, and critical reflective self-assessment is one way in which I take ownership of that project. The idea of ownership can be misunderstood, and this misunderstanding tracks a problem with the way some people respond to their flaws. Some people speak as if their flaws were an important part of their individual identity, as if they would be losing an important part of who they are if they were to overcome these flaws. While it is true that people are sometimes identified by their flaws ("Who's Tom? He is the duplicitous one."), it is another thing to say that people would be losing something important if they overcame their flaws. This threatens to turn ownership into something which negatively restricts one's capacity to become the person one wants to be. But ownership, correctly understood, enhances autonomy because ownership of flaws *as flaws* only makes sense alongside ownership of values. Ownership of one's own duplicity *as a flaw* only makes sense alongside ownership of the value of candour (valuing candour is what makes its antithesis, duplicity, a flaw). The attitude that I am recommending in critical reflective self-assessment entails joint ownership of one's character and one's ideal self-conception. Indeed, it is

this joint ownership that is the cause of accurate shame. As I highlighted above, shame occurs because one simultaneously holds a conflicting self-conception and ideal self-conception. So what I am suggesting is not that one only develop joint ownership while engaging in critical reflective self-assessment, but rather that one retain it and pay attention to it. As I discussed in chapter three, taking ownership of one's flaws should also lead one to recognise and compensate for the ways these flaws may affect how one engages in critical reflective self-assessment.

In chapter two, I developed a notion of coherence and a model of value on which coherence between values can be understood, and I described how the moral life can be seen as a quest for coherence between values and between values and action. In chapter three, I showed how the quest for coherence in these two spheres is mutually supportive and interdependent, and together can be seen as a quest for coherence between one's ideal self-conception and one's character. These are all different ways of looking at the simple and familiar idea of trying to be or become a certain kind of person who lives life in a certain way. Put simply, I have been arguing that shame encourages one, often forcefully, to think about who one is and who one would like to be, and this helps one to be or to become such a person.

I have been discussing the way in which shame can lead to critical reflective self-assessment but, as I suggested earlier, critical reflective self-assessment can also lead to shame. This happens in two ways. On the one hand, critical reflective self-assessment can lead us to make a new negative self-assessment that can occasion a new instance of shame. On the other hand, critical reflective self-assessment can merely lead to shame occasioned by a post-reflective assessment that concurs with the earlier assessment which occasioned shame in the first place. I shall call the shame in both these situations post-reflective shame. So pre-reflective shame is occasioned by a pre-reflective self-assessment, and post-reflective shame is based on critical reflective self-assessment.

One reason to think that post-reflective shame will be more valuable than pre-reflective shame is that it is likely to be based on a more accurate self-assessment, because critical reflective self-assessment will generally act as a corrective process. Another related feature of post-reflective shame is that we are aware of this – we *know* that post-reflective shame is likely to be based on a more accurate self-assessment. After thinking sincerely, bravely, and carefully about how the self-assessment that grounds an episode of shame is related to our character and our ideal self-conception, we have good reason to endorse that shame.

In this section, I have been focusing on the relationship between the self-assessment that occasions shame and the process of critical reflective self-assessment. I have argued that the self-assessment prompts and informs the process of critical reflective self-assessment, and that this process informs and corrects the self-assessment. In paradigmatic shame, where the self-assessment

is accurate and the process is most valuable, each therefore supports the other. The self-assessment will prompt the process, which will then reaffirm the original self-assessment. However, the relationship between the two is also valuable in non-paradigmatic shame.

Shame in which one's negative self-assessment does not make sense in light of one's standard of assessment is irrational. Irrational shame usually occurs when one adopts an audience's negative assessment of oneself without adopting that audience's standard of assessment. I discussed two cases of irrational shame in chapter one, involving Arnica and Jennifer. In both cases I suggested that they would be better off if they did not feel shame, and suggested ways in which they could overcome their shame. However, if we consider Arnica and Jennifer carefully, we can see some value in their experiences. In each case, insofar as their shame prompts a process of critical reflective self-assessment , it prompts them to attend to their moral lives, to consider their characters and their ideal self-conceptions. This means that Arnica and Jennifer may not literally have been better off if they had not felt shame in these situations. They may be better off if they do feel shame, engage in critical reflective self-assessment, and then reject their shame.[5]

A similar point applies to inaccurate shame. Consider the example in which my shame is occasioned by my negative self-assessment that I am dishonest. My shame may be rational but inaccurate if I value honesty, but I am not really dishonest. (Perhaps I moved too quickly to infer dishonesty from only a few actions.) Here again, I may correct this inaccurate self-assessment through a process of critical reflective self-assessment, and this process may be valuable for reasons other than simply this correction (as I discussed in chapter three). So even irrational or inaccurate shame can be valuable when it leads to a process of critical reflective self-assessment.

I began this section by clarifying my distinctions between shame that is rational or irrational, accurate or inaccurate and paradigmatic or non-paradigmatic. I also clarified my distinction between the self-assessment that occasions shame and the process of critical reflective self-assessment, but went on to show the close parallels between them. I argued that these parallels explain their interdependent relationship. I then discussed some features of the most valuable critical reflective self-assessment as a way of highlighting the valuable role shame and such critical reflective self-assessment can play in one's moral life.

I have been focusing on the self-assessments involved in shame. I only hinted, earlier, at a role that might be played by the painfulness of shame when I suggested that it might make one more critical both of oneself and of the negative self-assessment involved in any particular episode of shame. I turn now to consider this phenomenological feature of shame more carefully.

[5] Arnica is less mature and less well equipped to deal with her irrational shame, and as such it is less likely to be valuable. However, if over time she were to successfully reject her shame through critical reflective self-assessment, this could be a particularly important moment in her development.

Emotions and values

In the previous section, I suggested some reasons that one has to endorse the self-assessment involved in one's post-reflective shame. However, shame is more than just a self-assessment, it is also a painful experience. So endorsing the self-assessment that occasions one's shame does not yet mean that one should endorse one's shame experience. The only reason I have given to think that the painfulness of shame is valuable is because of its role in prompting critical reflective self-assessment. In this section I want to argue that the painfulness of shame is valuable for other reasons. I aim to show that shame is one way in which one's emotional connection to one's values manifests, and that it presents an opportunity to reaffirm this connection.

My intuition is that valuing implies more than simply having a belief that something is valuable. To value fairness, for example, implies something other than simply recognising that honesty has extrinsic or even intrinsic value. As Elizabeth Anderson says, "a mode of valuation includes distinctive emotional responses to the apprehension, achievement, and loss of things related to what is valued."[6] Here is a situation that will act as an example to tease out this intuition. It involves Anne, who values fairness:

1. Anne reads in the newspaper that a very poor old lady has been waiting to be allocated a government-sponsored house for fifteen years because, when her name finally reached the top of the long waiting list ten years ago, a government official swapped it with his girlfriend's name which was near the bottom.
2. The following week, Anne reads that the old lady has been given a home.

My intuition is that in situations 1 and 2, if we take them at face value, we expect more from Anne than simply an inward intellectual recognition that the value of fairness has or has not been upheld. We expect an emotional response from Anne. It may be difficult to say exactly what emotional response we expect in each case, and I do not think that we should be too specific. It would be misleading to claim that there is only one appropriate response to a value being promoted or abused, like indignation or anger. My intuition is that the emotions we expect correlate with whether the value in question is being promoted or abused; when our values are promoted we expect a positive emotion, and we expect a negative emotion when our values are abused.[7]

What makes the example interesting is that Anne does not stand to lose or gain anything through the unfairness that is manifested. She reacts, if she does, to the unfairness in the situation, rather than any unfairness to her. Her emotional reaction to the situation may be complex and multifaceted: she may react with empathy for the old woman and with anger at the official. What I am pointing to, though, is clearest if we consider emotions she may have which are not directed at

[6] Elizabeth Anderson, *Value in Ethics and Economics*, p. 11.
[7] The strength of one's emotional reaction depends on variables like, for example, physical distance (although this example makes it clear that it is debateable which variables should matter).

anyone in particular, but more generally at the unfairness in the situation. For example, Anne may simply feel angry or sad that this sort of thing happens, that vulnerable people are treated unfairly.

It is also possible to embroider 1 and 2 with details that would lead us not to expect an emotional response, and this is why I added the caveat "if we take them at face value". For example, we can imagine that Anne works for the Department of Human Settlements as an anti-corruption investigator. Anne does value fairness, but it would be too emotionally draining for her to respond to unfairness on every occasion in the way that we would otherwise expect. She sees gross unfairness every day. So, on this occasion, Anne calmly sends her colleague in that district a note to follow up on the case. That in the absence of the expected emotional reaction we would expect such a story only reinforces my point.

My example is designed to trigger the intuition that there is something odd about an absence of an emotional reaction when our values are promoted or abused. If Anne claims to value fairness but does not respond emotionally when people promote or abuse that value, this leads us to question whether or not she really does value fairness, even if her actions support her claims. I am not arguing that emotional reactions speak louder than actions, but rather that emotional reactions are *also* important. Against the background of this general connection between values and emotions, I want to say something more specific about the connection between values and emotions within the context of shame and the moral life. Ultimately, I want to show that the pain we feel in experiences of shame is largely due to the emotional connection that we have with our values, and that this emotional connection can also provide a positive motivation for us to promote our values.

I have discussed the way in which we expect an emotional response from someone when they become aware of an occasion on which their values have been promoted or abused. Here I want to narrow my focus to the reaction I have when I am the one who promotes or abuses my values. Again, an example may be helpful. Imagine another situation involving Anne, who values fairness:

> 3. During an annual review with her supervisor, Anne is shown persuasive evidence that she has systematically investigated female colleagues more rigorously than male colleagues, and recommended harsher disciplinary actions against them.

Following from my discussion above, my intuition here is that we expect Anne to do more than simply acknowledge that she has been unfair and make an effort to be fairer in future. We expect her to have a negative emotional reaction.

Anne's reaction in 3 should be different to her reactions in 1 and 2. Here, she is the one who has abused her value, she is the one who has been unfair. I am not suggesting that this negative

reaction must be shame. I have already mentioned the care with which we need to avoid moving too quickly from action to character, and I have argued that the object of one's shame should be one's character. However, if Anne takes her actions to flow from a flaw in her character (and thus assuming that Anne values fairness, in order to see unfairness as a flaw), she should feel shame. The pain involved in Anne's shame is, at least in part, the pain of turning away from her value, or of realising that she has done so. And, insofar as values come together in one's ideal self-conception, shame is experienced as the pain of turning away from one's ideal self-conception.

Shame is a painful manifestation of the connection between valuing and emotions: one feels the pain of turning away from one's ideal self-conception and the values that constitute it. In this way, shame can inform one about one's values. For example, my shame occasioned by my dishonesty can lead me to become aware that I value honesty (if my shame is accurate). Here I am arguing that shame can do more than this. It can also make one aware of one's emotional connection to one's values. For example, my shame can make me aware of my broader emotional connection to honesty, because when I experience shame occasioned by my dishonesty, I can recognise my shame as a painful manifestation of this connection. This new or reinforced awareness of the unpleasant side of one's emotional connection to one's values can motivate one to avoid abusing or betraying those values, and it will almost certainly make it harder to simply ignore them.

However, one's emotional connection to one's values can also be felt pleasantly. As I said above, we expect Anne to react to 2 in a positive emotional way. Similarly, if Anne identifies a flaw in her character and takes successful steps towards correcting it, then we expect Anne to react in a positive emotional way as she moves towards her ideal self-conception. Insofar as Anne's values come together in her ideal self-conception, she has a strong emotional connection to her ideal self-conception. Her positive emotional response when she comes closer to embodying her ideal self-conception, when she achieves greater coherence between her character and her ideal self-conception, is not just a reaction to her success. It flows naturally from her emotional connection to her values.

So the connection between values and emotions can act as both carrot and stick. It can motivate one to promote one's values and not to abuse one's values, and it always returns one's attention to one's values and to their relationship to one's character. This points to the positive role that the emotional connection plays in one's moral life. As a manifestation of the connection, shame itself plays this valuable role. However, I said above that shame not only plays this role as a manifestation of the connection, but also makes us aware of the connection. Now I shall argue that shame also plays another important role by providing one with the opportunity to reaffirm and reinforce one's emotional connection to one's values. This may seem counterintuitive because

shame itself is a manifestation of the unpleasant side of the connection. Nevertheless, developing the connection is an important part of the moral life because of the role it plays in promoting coherence, and through an experience of shame one may come to realise this role and endorse it.

It may be helpful to illustrate this with another example. Imagine that I value honesty. On a number of occasions I have lied to get out of trouble. Shame has led me to engage in critical reflective self-assessment, which has helped me to understand that being honest in such situations is an important part of realising my ideal self-conception. Faced with similar situations in the future, it is likely that I shall want to be honest, but I shall also want to get out of trouble. What I am now suggesting is that having a more developed emotional connection with honesty will motivate me to be honest. I will feel drawn towards the honest action and repelled from the dishonest action. Placing this back into the context of the moral life, it becomes clear that my emotional connection to my values can promote coherence, and coherence can promote my emotional connection to my values. It will be harder to develop simultaneous emotional connections with two inherently conflicting values, like honesty and duplicity. In the case of two values which form a mutually supportive pair within a coherent set of values, my emotional connection with each of the two values will also be mutually supportive. For example, my emotional connection to honesty and my emotional connection to candour will, for the most part, reinforce each other. This will promote coherence in both ways: firstly, it will promote a more coherent set of values and, secondly, it will promote coherence between those values and action. Shame can motivate me to avoid dishonesty and duplicity, make me aware of my emotional connection to honesty and candour, and provide an opportunity to reinforce that connection so that I will be more likely to act with honesty and candour in future.

For clarity, I have been discussing the connection between emotions and values within the context of paradigmatic shame. The valuable roles that I have attributed to the connection are particularly clear in that context. For example, it is clearer that shame provides an opportunity for me to become aware of my emotional connection to my values when my shame is accurate and leads to a process of critical reflective self-assessment. However, I suggested earlier that shame can also play a positive role in one's life when it is not followed by that process. In the absence of critical reflective self-assessment, shame is still a manifestation of one's emotional connection to one's values and can still play the valuable roles that I have been discussing. I introduced the notion of paradigmatic shame both to improve the clarity of my discussion, and to make a normative point – I said that *generally* our shame should aspire to the paradigmatic case, because in most situations paradigmatic shame will be more valuable than non-paradigmatic shame.

Earlier in this section, however, I said that shame neither does nor should always lead to critical reflective self-assessment. So shame should *not always* aspire to the paradigmatic case. One reason to think that shame should not always lead to critical reflective self-assessment is a worry that this will recommend too much thinking and not enough acting. This may be expanded to state a more general worry that this thesis recommends an overly abstract or intellectual approach to one's emotional reactions or one's moral life in general. The moral life, as I have presented it, is a quest for coherence. I have recommended that shame should most often be followed by a process of critical reflective self-assessment, and I have recommended that one engage in this process with the moral life in mind. So my recommendations apply specifically to shame and more generally to the moral life.

However, I am not recommending any approach at the expense of others. Unlike *Seppuku* (ritual suicide practiced by Samurai warriors to rid themselves of shame), the approach to shame that I recommend is compatible with many others. I have left room for complementary ways of promoting coherence in response to one's shame, such as unreflective spiritual exercises or dedicating one's life to a worthy cause. And while I have focused on a reflective approach to the moral life, I have also argued that bringing one's character into a more coherent relationship with one's ideal self-conception requires change. Critical reflective self-assessment provides an important way of thinking about change, but it will be meaningless if that change is not enacted. This is one of the reasons that I have stressed the important role that shame plays, not merely in making us aware of how we need to change, but also in motivating us to change.

Furthermore, we should note that on some occasions thinking and acting will compete for our time and attention, and we can identify some situations in which acting should be given priority. Many episodes of shame are similar enough that we need not subject each to a process of critical reflective self-assessment. I suggested in chapter two that Jennifer should think carefully about her mother's shaming criticism regarding her abortion. I also suggested that, having thought carefully about it, she may come to reject it. While she may have reason to reconsider this rejection at some point in the future, it would be counterproductive for Jennifer to engage in critical reflective self-assessment every time she is faced with her mother's shaming criticism, even if she cannot help but feel shame in her mother's presence. In that situation, ignoring her mother and distracting herself might be exactly the right response. A similar point applies to accurate shame. If I am frequently ashamed at my dishonesty, and critical reflective self-assessment has led me to accept my shame as accurate, then it may be a waste of time to critically reflect on my dishonesty every time I feel ashamed of it, particularly if it is at the expense of other potentially more effective transformative responses.

Conclusion

Not every episode of shame is valuable. Sometimes one's shame is inaccurate and does not promote coherence in any way. It may even create conflict, crushing one's faith in oneself and confusing one about one's values. So a particular episode of shame can play a negative role in one's moral life. However, shame generally plays a valuable role by promoting coherence and thereby contributing to one's moral life.

In this final chapter, I have divided my discussion of how shame can play this valuable role into two sections. In the first, I concentrated on the connection between shame and self-assessment. Building on my discussions of coherence and the process of critical reflective self-assessment in earlier chapters, I argued that this process promotes coherence. I argued further that, due to the relationship between them, shame and critical reflective self-assessment in combination are particularly conducive to the promotion of coherence. In the second section, I concentrated on the connection between emotions and values. I argued that this connection promotes coherence and that, insofar as shame manifests this connection, it promotes coherence. I also argued that shame promotes coherence by prompting a new awareness of this connection, and by providing an opportunity to reaffirm and reinforce this connection.

I hope that it is now clear that coherence and the moral life are not simply technical notions, but rather capture the process of becoming. There is a sense in which one becomes more truly oneself as one achieves greater coherence and thereby makes progress in the moral life. It is not simply that one's true self is one's ideal self-conception, so that becoming oneself can be straightforwardly understood as one's character coming to mirror one's ideal self-conception. Rather, the self includes the ideal self-conception and the character. As the elements of each of these come together to form coherent wholes, and as each of these coherent wholes come together to form a coherent self, so one becomes oneself. Shame is intimately connected to this process of becoming. It acts as a "coherence canary" in one's moral life, sounding the alarm when one has stepped off the path towards greater coherence. Shame can be a false alarm, but if one checks the accuracy of the alarm and then responds appropriately, shame can play an invaluable guiding role.

Conclusion

In the Introduction, I suggested that this thesis can be approached from two different perspectives. These perspectives generate two sets of questions that might be described as theoretical and pre-theoretical. I hope to have answered the theoretical questions clearly and convincingly. I provided accounts of shame and integrity, I explored the notions of coherence and the moral life, and I developed models of values and the self.

The pre-theoretical questions centred around your personal experience of shame. I tried to show how your shame can make sense and be valuable to you. However, shame is a painful experience. I did not claim that shame is or should be pleasant, nor that any particular experience of shame should be welcomed. As Laurence Thomas rightly notes, "[a] moral person recognizes the importance of being made aware of her shortcomings. But being made aware of them is not a source of delight for her."[1] Shame implies the identification of a shortcoming, and this is not in itself something that we should welcome or enjoy.

However, painful experiences take on a different character if one also recognises the positive role that they play in one's life. While one may jokingly refer to a dental procedure as "torture", there is something significantly different about getting a filling as opposed to having a tooth removed by a sadistic madman. At the very least, I have given an account of shame that encourages one to see one's shame as more like the former than the latter, as a necessary evil in order to live well, rather than a pointless ordeal. But I hope to have done far more than this. While I would not suggest that one should welcome any particular episode of shame, I have given reasons for one to welcome shame in general as a feature of one's life; I have provided an argument in support of the intuition that one would be worse off if one were shameless. If one understands the role shame can play in one's moral life, then one should come to see shame, not as something to be tolerated, but as a valuable feature of one's moral life that one should strive to recognise, understand and learn from.

So, to conclude, shame is valuable. My shame is valuable to me because it plays a positive role in my moral life by promoting coherence. I have values, and I want to be a certain kind of person who lives life in a certain way, which means both embodying and promoting my values. Coherence is crucial to each of these deep wants. Shame can make me painfully aware of incoherence in each case, and instruct and motivate me to seek and achieve greater coherence. In every sense, shame leads me back to myself.

[1] Laurence Thomas, *Living Morally*, p. 141

References

Anderson, Elizabeth. *Value in Ethics and Economics.* Cambridge, Mass.: Harvard University Press, 1998.

Aristotle. *Nicomachean Ethics*, trans. Roger Crisp. Cambridge: Cambridge University Press, 2000.

Calhoun, Cheshire. 'An Apology for Moral Shame', *The Journal of Political Philosophy*, 12 (2), 2004. pp. 127–146

Cox, Damian, Marguerite La Caze, and Michael Levine. 'Integrity'. *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition). Edward N Zalta (ed), <http://plato.stanford.edu/archives/fall2008/entries/integrity/>.

Cox, Damian, Marguerite La Caze, and Michael Levine. *Integrity and the Fragile Self*. Aldershot: Ashgate, 2003.

de Sousa, Ronald. 'Emotion'. *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N Zalta (ed), <http://plato.stanford.edu/archives/fall2008/entries/emotion/>.

John Deigh. 'Shame and Self-Esteem: A Critique', *Ethics*, 93 (2), 1983. pp. 225-245.

Ekstrom, Laura Waddell. 'A Coherence Theory of Autonomy', *Philosophy and Phenomenological Research*, 53 (3), 1993. pp. 599-616.

Erikson, Erik H. 'Autonomy v Shame and Doubt', Herbert Morris (ed), *Guilt and Shame*. Belmont, Calif.: Wadsworth, 1971.

Frankfurt, Harrry G. 'Freedom of the Will and the Concept of a Person', *Journal of Philosophy*, 68 (1), 1971. pp. 5-20.

Freud, Sigmund. 'Origin of the Sense of Guilt' in Herbert Morris (ed), *Guilt and Shame*. Belmont, Calif.: Wadsworth, 1971.

Friedman, Marilyn A. 'Autonomy and the Split-level Self', *The Southern Journal of Philosophy*, XXIV (1), 1986. pp. 19-35.

Goldie, Peter. *On Personality*. London: Routledge, 2004.

Hugo, Victor. *Les Misérables*. Herfordshire: Wordsworth Classics, 1994.

Hutchinson, Phil. *Shame and Philosophy: An Investigation in the Philosophy of Emotions and Ethics.* Houndmills, Basingstoke:: Palgrave Macmillan, 2008.

Lynd, Helen Merrell. *On Shame and the Search for Identity.* New York: Harcourt Brace & Co, 1958.

Mayo, Bernard. 'Moral Integrity', *Human Values*, 11, 1976. pp. 27-43.

McFall, Lynne. 'Integrity', *Ethics*, 98 (1), 1987. pp. 5-20.

Piers, Gerhart and Milton B Singer. 'Shame', Herbert Morris (ed), *Guilt and Shame*. Belmont, Calif.: Wadsworth, 1971.

Ramsay, Hayden. *Beyond Virtue: Integrity and Morality*. London: Macmillan, 1997.

Rawls, John. *A Theory of Justice.* Oxford: Clarendon Press, 1972.

Sartre, Jean Paul. *Being and Nothingnes,* trans. H E Barnes. London: Routledge, 1969.

Schauber, Nancy. 'Integrity, Commitment and the Concept of a Person', *American Philosophical Quarterly*, 33 (1), 1996. pp. 119-129.

Simpson, Joe. *This Game of Ghosts*. London: Vintage, 1994.

Smedes, Lewis. *Shame and Grace: Healing the Shame We Don't Deserve.* San Francisco: Zondervan Publishing House, 1993.

Solomon, Robert C. *Thinking about Feeling : Contemporary Philosophers on Emotions*. Oxford: Oxford University Press, 2004.

Sorabji, Richard. *Self: Ancient and Modern Insights about Individuality, Life, and Death.* Chicago: University of Chicago Press, 2006.

Taylor, Gabriele. *Pride, Shame, and Guilt: Emotions of Self-Assessment.* Oxford: Clarendon Press, 1985.

Thomas, Laurence. *Living Morally: A Psychology of Moral Character.* Philadelphia, Pa.: Temple University Press, 1989.

Velleman, J David. *Self to Self: Selected Essays.* Cambridge: Cambridge University Press, 2006.

Velleman , J David. 'The Genesis of Shame' in *Philosophy and Public Affairs*, 30 (1). pp 27-52.

White, Stephen L. *The Unity of the Self.* Cambridge, Mass.: MIT Press, 1991.

Williams, Bernard. *Shame and Necessity.* Sather Classical Lectures, vol 57. Berkeley, Calif.: University of California Press, 1993.