The Story of My Life: Virtue, Character and Narrative

Lisa Grover

Contents

Abstract

The primary aim of this research is to develop a new philosophical analysis of the concept of character that reflects the complexity of people and meets the demands of moral explanation. It places the agent's particular perspective and the wider context at the centre of moral judgement. The reason for undertaking this project is to establish an account of morality that is not in conflict with discoveries in empirical psychology. It responds to the challenge that the situation usually has the explanatory role and that character traits rarely function as explanations for action. It argues that the best interpretation of the situationist position is that reasons would have to be features of the situation to separate the situationist argument from behaviourism. However, it argues that this would then commit the situationist to a controversial theory of action where what explains an action need not obtain. It argues that to evaluate a person or his action properly we need to tell an explanatory story and that this narrative construction is what best reflects the richness and complexity involved. It further argues that an adequate attribution of character to an individual will also take narrative form. Hence character traits can explain action because a narrative explanation of why an individual acted in a certain way can also be an explanation of why a certain character can be attributed to that person. It argues that narrative has central importance in the attribution of character traits because the narrative structure gives us an understanding of character that cannot be gained from a non-narrative presentation of the actions and events. This additional knowledge connects with the emotions important to moral evaluation of persons and actions.

Introduction

People are different. Sometimes philosophers forget this. The primary aim of my research is to develop a type of virtue ethics that both respects these differences and reflects the complexity of moral life. This will be achieved in two ways. First, by critical examination of an objection to virtue ethics and secondly through positive argument, developing a new theory of character that reflects the complexity of people and meets the demands of moral explanation. Often the agent's particular perspective and the wider context are ignored in moral theory, but I place these factors back at the centre of moral judgement. The focus will be upon people leading ordinary lives and their complexity rather than the agents that feature in the stripped back, somewhat artificial examples often used in the literature.

This research relates to a growing trend in moral philosophy against defining sets of moral rules. For example, John McDowell argues that morality is uncodifiable (McDowell 1998), Bernard Williams considers talk of moral *theory* to be flawed (Williams 1985) and Jonathan Dancy argues for particularism, which states that what is a reason in one situation may not be a reason, or may be a reason against, in another situation (Dancy 2004). However, this is not to say that there are no moral facts. Although my position will not ultimately define a set of moral rules, this is not to say that there is no fact of the matter as to what is right or that we cannot be wrong in our moral judgements. The point is merely that the complexity of moral life cannot be summarised by a set of simple (or even more complicated) rules. I will argue that to evaluate a person or his action properly we need to tell a story and that this narrative construction is what best reflects the richness and complexity involved.

My reason for undertaking this project is to establish an account of morality that is not in conflict with discoveries in empirical psychology. Much of virtue ethics is based upon an ancient Aristotelian moral psychology and I think it important to consider how this can be developed in light of more recent knowledge. It is important because an empirically discredited view of character would undermine the foundations of virtue ethics. My research focuses on explanations of moral action. The overall aim of my research is to argue that character can explain moral action, a foundational requirement of a traditional character-based virtue ethics. I support this claim by arguing that moral evaluation requires a certain sort of explanation, narrative explanation. I further argue that an adequate

attribution of character to an individual will also take narrative form. Hence character can explain action because a narrative explanation of why an individual acted in a certain way can also be an explanation of why a certain character can be attributed to that person.

Many people are familiar with recent work by, amongst others, Gilbert Harman and John Doris, on moral psychology (Harman 1999b; Doris 1998, 2002). Their challenge is that the situation usually has the explanatory role and that character traits rarely function as explanations for action, thus threatening the foundations of virtue ethics. In my work I critically examine this position, arguing that we are making an error in explaining an action simply in terms of a character trait or a situation, rather than a full narrative explanation. Mine is a novel attack because I analyse the arguments in the wider context of what we require from an explanation, rather than questioning whether the social psychology experiments provide evidence against a conception of character traits found in virtue ethics.

I argue that Harman and Doris are wrong to say that it is the situation that primarily explains human action. The situationist argument has to undermine the notion of intentional actions, those actions done for a reason, because virtues are not just dispositions to behaviour, but are intelligent dispositions to behaviour involving the motives of the individual. In light of this, I argue that the best interpretation of the situationist position is that reasons would have to be features of the situation to separate the situationist argument from behaviourism. However, I argue that this would then commit the situationist to a controversial theory of action where what explains an action need not obtain. I argue that the situationist cannot accept such a theory because he needs the situation to be a causal influence on action, but it seems that the situation cannot be both causal and a reason. This means that the agent's reason for action must be something else and that the situation is not the primary cause of intentional action. I argue that the social psychologist is merely looking at behaviour and attributing its cause to the situation without considering the agent's reasons for action.

I develop what may be meant by a full explanation. In our everyday lives we tend to explain what has happened in our past in the form of a narrative. I investigate whether belief-desire explanations alone are sufficient to explain our actions, or whether we need these fuller narrative explanations. I develop the idea that there are many causal factors that can feature in an explanation and that these factors are often expressed in the form of a

narrative. I also explore the idea that an agent's reasons provide a link between virtuous dispositions and particular actions; it is by evaluating his reasons that we decide whether his action was virtuous. I investigate questions such as the following: What sort of reasons do virtuous people act upon? What is the connection between virtuous persons, virtuous dispositions, virtuous acts and virtuous reasons? Does a virtuous action have to be caused by a virtuous reason that is grounded in a virtuous disposition?

My positive argument is that character traits are best attributed in narrative form rather than as a conditional statement. I argue that narrative has central importance in the attribution of character traits because the narrative structure gives us an understanding of character that cannot be gained from a non-narrative presentation of the actions and events. This additional knowledge connects with the moral emotions important to evaluation of persons and actions. I further argue that my account of character trait attributions as historical narratives provides a genuine alternative to the conditional account and that my account better coheres with our normal moral practice. Such narratives about past episodes are important for understanding character traits because they not only explain how past actions and events have a causal effect on how we are now, but also give us an understanding of how to feel about those past actions and events, giving us the resources for emotional understanding and evaluation of character traits.

Outline of Chapters

In outline, in *Chapter One* I introduce common themes within traditional character-based virtue ethics and establish why one may wish to defend such a view. This explains the basic appeal of the view, justifying the need for a defence. In *Chapter Two* I consider an alternative to the standard character-based virtue ethical view that was introduced in the first chapter. It is important to reject this alternative to establish the appeal of the traditional character-based view. This act-based alternative to virtue ethics is proposed by Thomas Hurka (Hurka 2001, 2006). This chapter explains and analyses this view. The two central arguments are that evaluation of acts is prior to evaluation of persons and that in everyday moral evaluation of acts we do not take into account a person's character. I argue that this view is an important alternative account to consider because it evades recent objections to the persistence of reliable character traits through time. I present five objections to this account. First, as it is dependent upon a consequentialist account of value, it faces some of the same issues faced by consequentialism. Secondly, it is not clear

that the virtuousness or viciousness of acts has conceptual priority. Thirdly, the value of dispositions is unclear. Fourthly, it is not apparent that the value of occurrent states is always greater than the value of dispositions. Finally, in everyday moral discourse we *do* regard both actions and character as having moral value. I conclude that an act-based virtue ethics does not provide an alternative to the character-based account. I propose that instead individual actions 'add up' to general dispositions because character trait attributions are of narrative form.

This conclusion means that the recent objections to virtue ethics, based upon social psychological evidence, that character does not have the level of influence over action that is required by virtue ethics will have to be addressed. In *Chapter Three* I consider and reject Harman's claim that virtue ethics is based upon an error theory about character traits. This view attacks the notion of a stable character trait. It questions whether it is a person's character traits that determines his actions and argues that instead it is the particular situation. Harman argues for the extreme view that there are no such things as character traits. He thinks that in ordinary moral thought we are making the 'fundamental attribution error' (1999b, 316). Harman explains this as meaning that we are making the error of ignoring situational factors and assuming that actions are the result of someone's character traits. In making this argument, Harman denies that character traits can function as explanations for action. A rejection of this view is important because, if Harman is correct, doubt is cast upon the assumption behind traditional character-based virtue ethical positions that people have character traits that explain their actions. I argue that his argument fails upon two grounds. First, that the experimental evidence is open to interpretation and that the most sensible interpretation does not support his conclusions. Secondly, that there is some ambiguity around the notion of a character trait that needs to be settled to establish whether the social psychologist has in mind the same phenomenon as the virtue ethicist. Two questions emerge from this discussion: (i) 'Might a more modest argument based upon this evidence still cause a problem for a character-based virtue ethics?' and (ii) 'What is a character trait?'

In *Chapter Four* I address Doris's more moderate objection that although character traits may exist, these character traits are fragmented and localised to specific situations, e.g. sailing-in-rough-weather-courageousness. He claims that behaviour is very sensitive to the particular situation and that the features of the situation are better predictors of future behaviour than any considerations about character (Doris 2002). He argues for the

existence of 'local' traits that are stable over time and that are situation-particular. He argues that these traits are too narrow to cause the differences in behaviour across situations required by virtue ethics. In this chapter I review some of the literature that responds to this account and suggest why these responses do not adequately deal with the objection. These attack the appropriateness of the experiments, deny that they cause a problem for the notion of character traditionally found in virtue ethics and accept his conclusions about character traits and construct a virtue ethics around this conception of character.[1] In the next two chapters I develop two of my own arguments against the position that Doris sets out.

In *Chapter Five* I put forward my objection to a situationist ethics, challenging Doris's claim that features of situations primarily explain our actions. I interpret the situationist position as claiming that explanatory reasons are features of situations. The explanations for action used by Doris all refer to features of the situation to explain the agent's action. This suggests that the explanatory reason is the feature of the situation and not any desire or belief of the agent. I argue that this commits him to a controversial theory of action, such as that proposed by Dancy (Dancy 2000b). He argues that reasons for action are features of situations. His argument has two unwelcome implications; (i) that what explains an action need not obtain and (ii) that action explanations in terms of reasons are not causal explanations. I challenge Dancy's argument by questioning his notion of the 'normative constraint', a central assumption of his argument. I argue that this rests upon a mistaken interpretation of Williams (Williams 1981). I offer a different interpretation of Williams that allows the claim that normative reasons are facts to be compatible with the claim that explanatory reasons are psychological states. However, this causes problems for the situationist argument because the rejection of Dancy's position involves commitment to the claim that a motivating reason is constituted by a psychological state, so it is psychological states of the agent that are of central importance in explaining action and not primarily features of the situation.

In *Chapter Six* I question Doris's notion of a local character trait by considering the problems raised by analysing local character traits as conditional statements. This is important because the local character trait is central to Doris's explanation of human actions. I consider some traditional problems with analysing dispositions as simple conditional statements, before considering the alternative specific and probabilistic

---

[1] These include responses from Webber (2006a, 2006b), Sreenivasan (2002), Miller (2003), Annas (2003), Merritt (2000), Goldie (2004b), and Vranas (2005).

analysis provided by Doris. I evaluate the metaphysical problems of such an analysis, before drawing on the work of Stuart Hampshire to question whether character trait attributions are conditional statements at all (Hampshire 1953). I argue that even these more localised traits seem open to generalisation and we appear to have no better reason to use one generalisation over another. The problem that Doris faces is that by making his notion of character more in line with the empirical evidence, he decreases the unification of an individual's character. By identifying character traits with specific situations Doris seems to be denying us the ability to make any evaluative connection between the fragments. Without a story to be told about how this may be done, the notion of a local character trait does little more than reiterate the point that we need to take care over attributing general character traits. I conclude with a proposal that character traits are best defined as historical narratives. This connects with my proposal in Chapter Two that individual actions add up to general dispositions because character trait attributions are of narrative form.

In Chapter Five I rejected Dancy's argument that features of the situation provide explanatory reasons for our actions. In *Chapter Seven* I set out a positive account of explanations of action, drawing upon Peter Goldie's argument that we need a fuller explanation of action such as that provided by historical narrative explanations (Goldie 2007). I agree with Goldie that narrative explanations are important because they give us a more complete understanding of why a certain action was done, rather than a basic belief-desire explanation that states only why an action made sense for the individual. Narrative is emerging as of central importance to understanding both character trait attributions and explanations of particular actions. Following on from the conclusions of Chapters Two, Six and Seven, I set out my account of narratives to be defended over the subsequent three chapters.

In *Chapter Eight* I argue that we cannot define the necessary and sufficient conditions for something's being a narrative. My methodology is to consider the definitions put forward by Aristotle, E.M. Forster, David Velleman and Noël Carroll (Aristotle *Poetics*; Forster 1927; Velleman 2003; Carroll 2001). I object to the claims that a complete narrative must have a story with an end and that it must be told in time sequence are necessary conditions for a narrative. I continue to object to Velleman's argument that narratives are necessarily emotive, claiming that the narrative can give rise to emotional understanding of the events without necessarily evoking a particular emotion in the audience. Overall, I reject the idea

that we can specify the necessary and sufficient conditions of narrativity and argue that a narrative is distinctive on the grounds of its general features.

Chapter Nine focuses on what type of knowledge we can obtain from historical narratives. I agree with Velleman's argument that the narrative structure itself generally gives rise to an understanding of the events narrated over and above the causal relationships between the events. I argue that a narrative will normally reveal something about how the narrator feels, how the protagonists feel or how the audience should feel that is lacking from a chronology of events and this information is in addition to the casual connections between the events. I respond to the general sceptical claim that narratives do not provide us with knowledge because they simplify things too much and find connections where perhaps there are none, concluding that narratives are truth-apt and that they can provide the audience with knowledge.

Chapter Ten develops my narrative account of character traits. I argue that character traits are not constituted by a narrative, but that this does not lead to the conclusion that narrative is inessential to understanding character traits. My argument depends upon a development of an account of narrative whereby narrative is not constitutive of character traits i.e. the narrative is about something, namely character traits, which have independent existence. My account argues that narrative is important for understanding character traits because of the knowledge and understanding that is derived from a narrative that could not be derived by any other means. This account provides an alternative to the conception of character trait attributions as conditional statements. It also provides an account of how a narrative attribution of a character trait can give a full explanation of an action.

I conclude that a character trait attribution is best expressed as a historical narrative rather than as a statement of a conditional disposition and that this definition provides a sound foundation for a character-based ethics and for explanation and evaluation of actions.

Introduction

In this chapter I first introduce common themes within traditional character-based virtue ethics. I then continue to establish why one may wish to defend such a view. This outlines the basic appeal of the view, justifying the need for a defence. I briefly explain virtue ethics, defining three central concepts that will be drawn upon in later chapters; virtue, practical wisdom and *eudaimonia*. I continue to put forward five arguments as to why one should be a virtue ethicist thus providing a reason to defend and further develop the view.

1. What is virtue ethics?

There are three main modern normative approaches to ethics: deontology, consequentialism and virtue ethics. Broadly speaking, deontological theories concentrate upon moral rules that guide our actions, claiming that features of the actions themselves have moral significance. Consequentialist theories, as the name suggests, concentrate on the consequences of actions, arguing that it is the consequences that have moral significance. Virtue ethics concentrates on moral character. Virtue ethics has ancient roots, particularly in the work of Aristotle, but has been a growing theory in moral philosophy since the late 1950's. Elizabeth Anscombe's paper 'Modern Moral Philosophy', published in 1958, led this revival of interest. Here she argues that:

> In present-day philosophy an explanation is required how an unjust man is a bad man, or an unjust action a bad one; to give such an explanation belongs to ethics; but it cannot even be begun until we are equipped with a sound philosophy of psychology. For the proof that an unjust man is a bad man would require a positive account of justice as a 'virtue'. This part of the subject matter of ethics is, however, completely closed to us until we have an account of what *type of characteristic* a virtue is – a problem, not of ethics, but of conceptual analysis – and how it relates to the actions in which it is instanced: a matter which I think Aristotle did not succeed in really making clear (1958, 4-5).

What is the appeal of virtue ethics? In general, virtue ethics can be characterised as being centred upon the agent rather than upon acts. It focuses upon being the right sort of person instead of the right actions to perform, hence addresses the question, 'What sort of person should I be?' as opposed to the question 'What types of action should I do?' Of central importance is the idea that ethics is complex rather than codified in rules and virtue ethics

is best placed to accommodate this complexity. So, what exactly is a virtue and how does it relate to moral deliberation and action?

i) Virtue

Virtue is a state of character. Aristotle lists the virtues as courage, temperance, generosity, magnificence, magnanimity, pride, patience, truthfulness, wittiness and friendliness (Aristotle *Nicomachean Ethics*). Each virtue has a corresponding vice, which can be either an excess or lack of the characteristic. Others have subsequently created other lists, but this thesis will not concentrate upon identifying the distinction between states of character and other personality traits.[2] Nor will it concentrate upon listing the set of virtues. The focus will be upon the concept of virtue and its relation to action and explanation. A fully virtuous person, according to Aristotle, is disposed to think, feel, choose, and act 'at the right times, about the right things, towards the right people, for the right end, and in the right way' (*Nicomachean Ethics* 1106b21). On this account the fully virtuous person possesses all of the virtues.

ii) Practical wisdom

Of central importance to virtue ethics is practical wisdom. Practical wisdom, or *phronesis*, is essential for getting things right. The virtuous person must have the wisdom to perform the right action, for the right reason, in the right sort of circumstances. To explain what is meant by practical wisdom, Rosalind Hursthouse gives the example of generosity (1999, 12-13). She says that to be generous means 'giving the right amount of the right sort of thing, for the right reasons, to the right people, on the right occasions' (1999, 12). So, she says, the 'right amount' will vary depending upon the circumstances; if I am poor, I am not mean if I don't give my family lavish Christmas presents and I am not ungenerous if I do not support someone who is idle. On her account every virtue involves practical wisdom so that the agent reasons correctly, given the circumstances, to the correct action. She argues that we cannot obtain moral wisdom simply by being taught (1999, 59). This wisdom is not easy to develop, involving practice over time.

If a person fails to act virtuously, this not only reveals that he lacks virtue but that his reasoning is also faulty. This will be of central importance later when considering

---

[2] For a discussion of the differences between character traits, personality traits, habits, etc. see Goldie 2004b, 7-13.

objections to a character-based account of ethics. It is relatively simple to identify circumstances and actions, enabling experimentation and conclusions to be drawn. However, it is not so easy to access the reasoning processes that link the circumstances and the action. Yet the reasoning processes are an essential part of the moral evaluation of people and behaviour under the character-based account, so the agent's reasons for action must be considered along with features of situations and actions.

iii) Eudaimonia

Eudaimonia is usually translated as 'happiness', 'flourishing' or 'well-being', but has no clear modern meaning. Hursthouse identifies problems with each translation (1999, 9-10). She thinks 'flourishing' is problematic because it can be applied to plants and animals as well as rational beings, but rational beings flourish in a different way to plants and animals. She argues that 'happiness' tends to have a subjective element as, generally, I identify whether I am happy or that my life is happy and am not, in normal circumstances, mistaken in this. She argues that 'well-being' is a problematic translation because it is not an everyday term and does not have a corresponding adjective. However, I think that over the intervening ten years, 'well-being' has become a more commonly used term in everyday language, so I think that this may be the most appropriate translation as it avoids the subjective connotations of 'happiness' and the non-rational attributions of 'flourishing'.

Eudaimonia will not directly be a central theme of this thesis, so I shall briefly say something about it here. *Eudaimonia* is important because the virtues are identified as those character traits that a human being needs to live well. Michael Slote objects to Hursthouse's original characterisation of flourishing on the basis that Aristotle says that 'human flourishing largely *consists in* acting virtuously from a virtuous character over a sufficiently long life' rather than it providing an independent ground for virtue (Hursthouse 1991 reference by Slote 1997, 207). He identifies the following problem with Hursthouse's position: a trait may be necessary for flourishing, and therefore a virtue, yet acting in accordance with that trait on a particular occasion may in fact be inconsistent with his flourishing. He gives the example of benevolence still being a virtue even though exercise of this trait involves self-sacrifice. However, he argues that she does not provide an account of why a trait that is necessary for flourishing is virtuous, yet flourishing is not the basis for the virtuousness of particular actions. Why does flourishing ground traits as virtuous, not individual acts? He argues that to be a distinctive theory, the virtue ethicist

13

needs to argue that virtuous traits are 'ethically fundamental' and that judgements about actions should be derived from these traits (1997, 209). This relationship between the virtuousness of traits and the virtuousness of acts will be further explored in the next chapter. This distinction and the relationship between traits and action will be a key distinction in what follows.

In a later work, Hursthouse argues that virtue does not provide a necessary and sufficient condition for *eudaimonia*, but that virtues are *in general* the best way of living well (1999, 172-4). She draws an analogy with health to explain this point. She gives the example of a doctor recommending that one would benefit from giving up smoking, exercising more and drinking moderately. She says that this is not a sufficient condition for living a long, healthy life, as even if the agent follows this advice, he could still get ill. She argues that similarly being virtuous is no guarantee of *eudaimonia*, but is still the best, most reliable advice available. Neither is virtue a necessary condition of *eudaimonia*, according to Hursthouse. She again draws an analogy with health, observing that there are many counter-examples of people who do not follow their doctor's advice yet still live long, healthy lives. Similarly, she thinks virtue not necessary for well-being, giving the example of a Nazi who escaped to South America and flourished. Hence *eudaimonia* does not provide an independent justification or ground of virtue, but she does identify that, in general, the virtuous life leads to well-being.

On this ground, Hursthouse thinks that virtue ethics has the potential to avoid the problems of cultural relativism that can plague other theories. She thinks this because the virtues are generally the best way of achieving *eudaimonia*, or the well-being, of all human beings. It may be argued that happiness or well-being or flourishing is something that is relative to the individual, but this is to misunderstand what is meant. It can instead be argued that *eudaimonia* is a general concept applicable to all human beings, denoting what it is for a human being, independent of circumstances, to flourish or be well. Martha Nussbaum agrees with this approach, as she thinks that virtue ethics does not reduce to a concern with localised norms because the concept of human flourishing applies to all humans, regardless of circumstances (1999, 177).

2. *Why be a virtue ethicist?*

This section suggests some reasons why one should be a virtue ethicist. First, I argue that virtue ethics can best accommodate the complexity of everyday moral life. Secondly, I argue that it is an advantage of virtue theory that it focuses on more than how people act. Thirdly, I argue that it removes the gap between reasoning and motivation that one often finds in other theories. Fourthly, virtue theory can accommodate degrees of moral understanding. Finally, I argue that it provides an appealing account of moral development, not usually considered by other moral theories.

First, an advantage is that virtue ethics best accommodates the complexity of everyday moral life. Starting with Anscombe's paper in 1958 there has been a move to reject the law-based ethics found in deontology and consequentialism. Anscombe argues that:

> it would be a great improvement if, instead of 'morally wrong', one always named a genus such as 'untruthful', 'unchaste', 'unjust'. We should no longer ask whether doing something was 'wrong', passing directly from some description of an action to this notion; we should ask whether, e.g., it was unjust; and the answer would sometimes be clear at once (Anscombe 1958, 8-9).

Edmund Pincoffs agrees with this approach, picturing Aristotle's primary ethical questions as 'what is the best kind of individual life?' and 'what qualities of character are possessed by a man who leads such a life?' (Pincoffs 1971, 553). He pictures Aristotle as studying types of men as 'possible exemplars of the sort of life to be pursued or avoided' (1971, 553). He is not concerned with how we should act in particular difficult situations but with how we should live. Pincoffs' main claim is that '...reference to my standards and ideals is an essential, not an accidental feature of my moral deliberation' (1971, 564). He does not think that we are morally evaluated in terms of how well we abide by certain rules setting minimal limits on conduct, even though following those rules may be necessary for moral worth. He does not think that moral evaluation should be reduced to consideration of how conscientious a person is at following the rules; he thinks that other qualities are morally important too. He claims that a law-based ethics reduces morality to conscientiousness and that ethics is too focused on finding a rational ground for deciding what to do in situations where it is difficult to know what to do. He argues that instead ethics should take into account character and moral ideals before discussing decision-making.

Pincoffs does not think that morality is reducible to one essential characteristic. It may be important that we are conscientious; following rules, creating new rules where none exist,

and dealing with conflicts between rules. But, he asks, how do we make moral decisions without already being a person of good moral character? He thinks that unless a person is loyal, just, honest, sensitive to suffering, etc. moral dilemmas will not arise for him (1971, 567). He says 'to grant that rule-responsibility is socially essential is not to grant that it is the essence of morality, in that all other moral character traits can be reduced to or derived from some form of this one' (1971, 567). He also criticizes theories which concentrate on the question of 'usefulness' when deciding what to do. He thinks that these theories also do not take into account the moral character of the individual involved. He argues that we may use rules to discover what is permissible or mandatory in a situation, but that this is not all we need to consider when making a decision. He thinks that this is too narrow because we also need to take into account the moral conception we have of ourselves: 'It reduces the topic of moral character to the topic of conscientiousness or rule-responsibility. But it gives no account of the role of the character as a whole in moral deliberation; and it excludes questions of character which are not directly concerned with the resolution of problems' (1971, 571).

M. F. Burnyeat agrees with this type of approach, arguing that:

> ...the noble and the just do not, in Aristotle's view, admit of neat formulation in rules or precepts (cf. 1.3. 1094b14-16; 2.2. 1104a3-10; 5.10. 1137b13-32; 9.2. 1155a12-14). It takes an educated person, a capacity going beyond the application of general rules, to tell what is required for the practice of the virtues in specific circumstances (2.9. 1109b23; 4.5. 1126b2-4) (1980, 72).

He thinks that 'What Aristotle is pointing to is our ability to internalize from a scattered range of particular cases a general evaluative attitude which is not reducible to rules or precepts' (1980, 72). He thinks that we should focus upon being '...the sort of person who does virtuous things in full knowledge of what he is doing, choosing to do them for their own sake, and acting out of a settled state of character (1105a28-33)' (1980, 73).

The appeal of virtue ethics is that it focuses on these questions of character and its role in moral deliberation. Virtue ethics is thus better placed to deal with complexity of morality than deontology or consequentialism, recognising the difficulty of codifying moral behaviour. Williams observed such complexity in moral life and argued that therefore trying to develop a moral theory was inappropriate or misguided (Williams 1985). A key question here is whether virtue ethics is really a *theory*. Some people argue that it isn't a rival theory, or at least not in the same way that consequentialism and deontology are thought to be rivals, because it is interested in answering different questions. For example,

Nussbaum argues that virtue ethics is not a distinctive approach to be contrasted with consequentialism and deontology on two grounds. First, she thinks that it is not a distinctive approach because both utilitarian and Kantian accounts provide a treatment of virtue (Nussbaum 1999). She gives as examples of deontological treatments of virtue Kant's 'Doctrine of Virtue' in his *The Metaphysics of Morals* and John Rawls's *A Theory of Justice*, which place emphasis on virtue (1999, 165). She defends her claim that utilitarians are also concerned with virtue on the grounds that Henry Sidgwick's primary concern in *The Methods of Ethics* is to argue that the virtues have a utilitarian basis and that John Stuart Mill blends the concept of *eudaimonia* with his utilitarianism (1999, 165-7).[3] She questions whether 'there is such a thing as "virtue ethics", that this thing has a definite describable character and a certain degree of unity, and that it is a major alternative to both the Utilitarian and the Kantian traditions' (1999, 164). Her argument is that because virtue can be part of the utilitarian and deontological positions, it cannot be set up as a position in its own right (1999, 167). She calls this a 'category' mistake because 'lots of people are, and have long been, writing and thinking about virtue within the Kantian and Utilitarian traditions. Virtue ethics cannot, then, be an alternative to those traditions' (1999, 200).

Secondly, Nussbaum argues that virtue ethics is not a distinctive approach because there is limited unity between those who identify themselves as virtue ethicists and that even they themselves deny that virtue ethics is a theory (1999, 168). She identifies three claims that are common between virtue ethicists: first, that moral philosophy should be concerned with the agent, as well as with choice and action; secondly, that moral philosophy should therefore concern itself with character and with settled patterns of motive, emotion, and reasoning; thirdly, that moral philosophy should focus not only on individual actions, but also on overall patterns (1999, 170). She does not think these concerns are solely those of the virtue ethicists, relying on the examples above to illustrate the concern of some utilitarians and deontologists with these issues. She argues that virtue ethics has largely emerged as an apparent third type of position because most virtue ethicists are opposed to either utilitarian or Kantian positions (1999, 168-9). She thinks that these philosophers are reacting to the emphasis on choice in much moral philosophy: 'the competing normative theories competed to give the best account of how one ought to choose in a complex situation, and the competing metaethical theories vied to give the best account of what

---

[3] She specifically refers to *On Liberty* Chapter 3, where Mill discusses Greek ideals of self-development (Mill 1978).

ethical discourse and reasoning aimed at choice really were' (1999, 171). However, despite this 'common ground' she argues that there is no unity to this group of thinkers because 'they have different targets and different positive views' (1999, 200). For Nussbaum, although this common ground is significant, its concerns can equally be pursued by Kantians and utilitarians.

It is possible to reject Nussbaum's arguments. First, that some Kantian and utilitarian views can be demonstrated to have a concern with virtue does not undermine virtue ethics as a separate approach. Where such views consider virtue, virtue has a supplementary or complementary role to the central theory. The development of virtuous character will tend to have instrumental value in that it is the best way of ensuring compliance with the moral rules or demands of utility. Virtue ethics is a clearly different approach because it is virtue that has intrinsic value and the value of any rules or concepts of happiness will be derived from virtue. Secondly, this disagreement as to the role of virtue in moral theory supports the claim that virtue ethics is a distinct theory in opposition to the other two positions. Again, that some Kantian and utilitarian views can be demonstrated to have a concern with virtue does not undermine virtue ethics as a separate approach because it is essential that the theories are discussing the same issues for them to be in opposition at all. Finally, that there is disagreement in approach between different virtue ethicists does not undermine this as a category; Nussbaum herself even identifies some common ground. There seems to be no more difficulty in categorising and summarising the views of virtue ethicists than views of Kantians or utilitarians.

Hence, virtue ethics is in a position to take on board some of Williams's concerns regarding the complexity of moral life, yet provide a theoretical framework that can stand in opposition to consequentialism and deontology. Slote agrees, identifying the distinctive features of virtue ethics as being the agent-focus of virtue ethics and the use of aretaic concepts (good, virtuous, etc.) rather than deontic concepts (right, duty, etc.) (Slote 1997). He argues that virtue ethics does not provide exceptionless universal principles and does not claim that there is a single way of resolving all moral issues (1997, 180). He thinks that under a wider concept of 'theory', virtue ethics is still a theory, yet accommodates Williams's concern that moral issues cannot be resolved with a single, simple methodology. It is an advantage of virtue ethics that it can accommodate the complexity of moral life better than consequentialist and deontological approaches.

Secondly, it is an advantage of virtue theory that it focuses on more than how people act. Hursthouse argues that there is more to the concept of virtue than a tendency to act in certain ways (1999, 11). She argues that the virtues have to be states of character because one can give the appearance of being, say, an honest person without actually being an honest person, so there must be something more to the concept of a virtue than merely acting in a certain way. She gives as examples the expectation that a virtuous person will act from certain reasons, he will act in a certain manner (e.g. unhesitatingly), we expect him to have attitudes consistent with the trait (e.g. praise people for their honesty) and we expect him to have corresponding emotions (e.g. distress at dishonesty). She further argues that these characteristics are strongly entrenched, so any change has to happen slowly. Barring unusual circumstances, such as brain damage, we cannot change overnight.

McDowell also approaches the question of how one should live through the concept of the virtuous person (McDowell 1998). He argues for virtue being a type of knowledge, stating that virtues are 'states of character whose possessor arrives at right answers to a certain range of questions about how to behave' (1998, 51). This implies that there is some sort of situation-relatedness built into the concept of character that is missing in the concept of a rule. He says that a kind person has a 'reliable sensitivity to a certain sort of requirement that situations impose on behaviour' (1998, 51). This places emphasis on both the situation and 'sensitivity', a perceptual capacity. This sensitivity results in knowledge and he argues that this knowledge from reliable sensitivity is necessary for possession of the virtue (1998, 52). The concept of the virtue does not have to enter his reasons for the actions, but could be part of an explanation of his action.

L.A. Kosman argues that the virtues that feature in Aristotle's view are not dispositions solely towards certain type of action (1980, 104). He thinks that: 'Throughout his discussion of the moral virtues...Aristotle makes it clear that the activities for which virtues are dispositions are of two sorts, actions and feelings...He rarely mentions virtue...with respect to action alone, but rather in terms of this dual phrase' (1980, 104). On this ground he argues that '...Aristotle's moral theory must be seen as a theory not only of how to *act* well but also of how to *feel* well; for the moral virtues are states of character that enable a person to exhibit the right kinds of emotions as well as the right kinds of actions' (1980, 105). He uses the example of the courageous person to illustrate his point. He says: 'The courageous person is one who is frightened by the right things, in the right way, in the right circumstances, and so on, and who is not frightened when it is appropriate not to be' (1980,

108). He argues that such examples show that virtue is not merely a disposition towards a certain appropriate action, e.g. courageous, in response to a feeling of fear, but that virtue is a disposition towards the appropriate feeling itself (1980, 108-09). He concludes that virtue is complex because it involves a characteristic set of feelings, as well as a characteristic set of actions (1980, 109).

Williams argues that a virtue is a disposition of character to choose to act in certain ways because 'they are of a certain ethically relevant kind' (1985, 8-9). However, he agrees with Kosman that 'virtues are always more than mere skills, since they involve characteristic patterns of desire and motivation. One can be a good pianist and have no desire to play, but if one is generous or fair-minded, those qualities themselves help to determine, in the right contexts, what one will want to do' (1985, 9). He continues to argue that '...if an agent has a particular virtue, then certain ranges of fact become ethical considerations for that agent because he or she has that virtue' (1985, 10). He thinks that '...for Aristotle a virtue was an internalized disposition of action, desire and feeling. It is an intelligent disposition. It involves the agent's exercise of judgement, that same quality of practical reason, and so it is not simply a habit' (1985, 35-6).

It is an advantage of virtue ethics that it focuses on more than action, but does one need all the virtues to truly act from any individual virtue? Aristotle claimed that to be fully virtuous one must possess all the virtues. Do the virtues have to be unified? Williams observes that:

> ...we accept, indeed regard as a platitude, an idea that Aristotle rejected, that someone can
> have one virtue while lacking others. For Aristotle...practical reason required the dispositions
> of action and feeling to be harmonized; if any disposition was properly to count as a virtue, it
> had to be part of a rational structure that included all the virtues (1985, 36).

The fully virtuous person, for Aristotle, possesses all of the virtues and hence knows which ends to pursue. If virtues are dispositions to act according to right reason, then someone who only occasionally acts virtuously is indicating that his moral reasoning is poor. On this account one cannot be said to have one virtue without also possessing all the others, but this is disputed by modern philosophers.

Neera Badhwar considers why Aristotle's claim that the virtues are unified is often dismissed by modern philosophers (Badhwar 1996). She agrees with Williams that modern philosophers tend to reject the Aristotelian view on the grounds that they think that certain virtues are independent of each other. For example, Owen Flanagan argues that

20

benevolence has no relation to courage, so it is possible to be benevolent but cowardly and that similarly one could be courageous but intemperate (1991, 271). Flanagan further argues that certain virtues are incompatible, so the idea that an individual could possess all the virtues is incoherent.

Badhwar argues that although the virtues may be 'disunited across different domains (areas of practical concern)' they are 'united within domains', so have limited unity (1996, 307). She gives as an example an individual being kind towards her friends but not towards strangers, but within the domain of friends she must also be generous, just, temperate, courageous, etc. towards them (1996, 308). She argues this because one individual will not have experience of all spheres of life, so will not be able to exercise his general virtue and wisdom in all areas (1996, 315). She gives the example of a statesman who may 'lack the kind of practical understanding of children that is required of a wise caretaker because of inexperience' (1996, 315). She continues to argue that: 'Conversely, someone who is wise with children and household management in general may lack the kind of experience that is necessary for wise statesmanship' (1996, 315). So, even though both are in general virtuous and wise, each lacks the experience of certain areas of life, hence Badhwar thinks we must reject Aristotle's claim that practical wisdom is a unity. She thinks this because 'no one has experience of all areas of life', so we would have to 'conclude that no one has any practical wisdom' (1996, 315). However, Richard Sorabji argues that '...the virtues are not separate, for courage is not a matter of facing any danger for any reason but of facing the right danger for the right reason (e.g. 3.7 1115b15-20). And what is right here depends partly in the claims of other virtues...' (1980, 207). So according to his interpretation one cannot be virtuous in one sphere of life but not in others because the virtues span all situations.

In general, virtues are states of character that dispose a person to think, feel, choose, and act in the right way. It is clear that the virtues that feature in Aristotle's view are not dispositions only towards action, but also towards thinking, feeling and choosing. That a virtue is more than a disposition to act will be important for my account of character to be developed later in this thesis. What is unclear is whether to be, for example courageous, one also needs to possess all the other virtues. The interplay between different virtuous character traits will be a recurrent theme throughout my thesis.

Thirdly, a key appeal of virtue theory is that it potentially appears to resolve some of the tensions between identifying what is right or good and the motivation to pursue such ends. On McDowell's view 'genuine deliverances of the sensitivity involved in virtue would necessitate action' (1998, 56). Under such an understanding a virtuous person has the required sensitivity to the salient features of the situation, providing a reason to act in a certain way. And because this is a virtuous person this is *just how they act*. They do not deliberate as to whether the action falls under a certain rule or whether it maximises consequences to identify whether it is good or right and then consider whether they are motivated to act in accord with this belief. The fully virtuous agent knows what the appropriate action would be in a particular situation and is motivated to act in accord with this belief because they are the type of person who acts in such ways. This removes the gap between reason and related motivational problems.

It may be challenged that virtue ethics cannot provide an explanation of moral dilemmas; there is not always a most appropriate or 'best' course of action. First, it is open to debate as to whether there are irresolvable dilemmas. If such dilemmas exist, then of course virtue ethics will not be able to provide a resolution for every case, even for the fully virtuous agent. If in principle all moral dilemmas are resolvable, then virtue ethics can present a similar response to that of the deontologist. In cases where different virtues conflict, this is only an apparent conflict, in the same way that the deontologist can argue that where different rules conflict, this is merely an apparent conflict. Where the agent perceives a conflict between two or more virtues he will have to exercise practical wisdom to ascertain whether one outranks the other in this particular situation.

Fourthly, virtue theory allows for degrees of moral understanding. As outlined above, under Aristotelian views, the fully virtuous person must possess all the virtues (unity of the virtues). One virtue cannot operate in isolation; e.g. acting kindly may not be the right act because acting fairly is, and this cannot be identified without the possession of both virtues. We can, of course, talk about the virtues of the non-fully virtuous person in isolation, as the virtues can be possessed to varying degrees of completeness, even if we need sensitivity to all the virtues to properly assess a given situation. Hence, a person can be more or less virtuous, depending upon how developed the person is; virtuousness is a matter of degree.

There are some parallels here with the particularist argument against generalisations

(Dancy 2004). This states that if you have a full set of moral rules which apply to all situations, then these rules will issue conflicting reasons. This argument could be applied to the unity of the virtues; if you have a full set of virtues which apply across all situations, then these virtues will issue conflicting reasons, so surely virtues issue conflicting reasons just as do rules? However, under McDowell's characterisation of virtue ethics, such an objection does not apply. For example, he says 'this reason is apprehended, not as outweighing or overriding any reasons for acting in other ways, which would otherwise be constituted by other aspects of the situation (the present danger, say), but as silencing them' (1998, 56). Virtue ethics admits the importance of moral wisdom; i.e. the fully virtuous person does not weigh different reasons from different virtues to decide action, but has a virtuous person's view of the situation 'in which considerations that would otherwise appeal to one's will are silenced, but nevertheless allow those considerations to make themselves heard by one's will' (1998, 56). The majority of this thesis will concentrate upon these issues. What is the relationship between virtue and the situation? What are the reasons a virtuous person acts upon? How does virtue explain action?

A final advantage of virtue ethics, following on from the previous point, is that *it is additionally appealing because it provides an account of moral development*, through habituation and practice. It takes seriously the idea that we develop from children into moral adults and allows for such progression. An individual can be virtuous to a greater or lesser degree, depending upon their stage of development. Burnyeat thinks that: 'What calls for explanation is how some people acquire continence or, even better, full virtue, rather than why most of us are liable to be led astray by our bodily appetites or unreasoned evaluative responses. It is no accident that Aristotle gives as much space to the akratic as a type of person as to isolated akratic actions' (1980, 85). He continues to argue that the less than fully virtuous person has a 'conflict in terms of stages in the development of his character which he has not yet completely left behind' (1980, 85). This is appealing in a moral theory as often such theories are directed at the fully developed 'moral agent' with little attention paid as to how this development occurred. Instead, here the focus is on how one develops the character of a fully virtuous person.

Others have developed accounts of moral development based upon deontological principles. For example, Lawrence Kohlberg argues that we learn moral obligation through stages (Kohlberg 1981). He argues that there are three levels, each comprising two stages. The lowest or 'preconventional' level is the level at which the meaning of 'right' and

'wrong' is defined in subjective, egoistic and prudential terms. The middle or 'conventional stage' is the level at which what is 'right' and 'wrong' is whatever an authority figure says. The highest or 'post conventional' level is the level at which 'right' and 'wrong' are objective universalizable standards, similar to those of Kantian morality. This is a sequential process and Kohlberg argues that moral obligation has its basis in conventional obligation, so learning social rules would precede learning morality. However, it is not clear that there are such stages of moral development. I do not think that we have to pass through all these stages before we can use concepts such as the Golden Rule. Parents often reason with children by saying such things as 'how would you like it if someone pulled *your* hair/ ripped the leg off *your* doll/...?' We also learn concepts such as fairness at a very early age. Two children when trying to divide something fairly between them may adopt the 'I divide, you choose' policy. Virtue ethics defines no rigid stages, allowing for simultaneous development of social and moral wisdom.

Conclusion

I have argued that a traditional character-based virtue ethics is worth defending for several reasons. Virtue ethics is a genuine alternative to deontology and consequentialism and it best accommodates the complexity of everyday moral life. There is no gap between reasoning and motivation and it provides an appealing account of moral development. For these reasons, it is a theory worthy of defence and development. It is a distinctive view because it takes the person as the fundamental unit of moral evaluation, whereas deontological views prioritise the evaluation of individual acts and consequentialism prioritises evaluation of the world, or aggregations of acts.

Both practical wisdom and virtuous traits can only be gained through experience and practice. The virtuous person not only possesses the virtuous traits, but has the wisdom to perform the right action, for the right reason, in the right sort of circumstances. Actions, reasons for action, circumstances and the character of the individual are all of equal importance to the virtue ethical account. The relationship between these four factors and their role in explanation and evaluation of action will be central themes that recur throughout this thesis. This acknowledgement that it is not only individual acts or only circumstances that have priority in moral evaluation is what marks virtue ethics as distinctive and gives it the resources to reflect the complexity of real-life moral evaluation and explanation.

The next chapter will consider and reject an alternative to this view, an act-based virtue ethics. Rejection of this alternative is important to establish character as of central importance to virtue ethics, thus clearly delineating it from other ethical theories. Subsequent chapters will consider the situationist objection to character-based virtue ethics. McDowell argues that 'a kind person can be relied on to behave kindly when that is what the situation requires' (1998, 51). The situationist objection disagrees with this claim, so must be rejected to maintain the link between character and action required by virtue ethics. Later chapters will develop my positive account of how this link is to be maintained.

Introduction

The argumentative aim of this chapter is to consider whether acts rather than dispositions are the foundation of virtue theory. I analyse Hurka's alternative account that states that virtues are appropriate attitudes to intrinsic goods and evils rather than dispositions. I argue that this account does not provide an adequate alternative to a traditional disposition based virtue ethics because it does not adequately distinguish itself from a sophisticated form of consequentialism. Further, it does not tackle the question of whether these attitudes are best cultivated by developing the appropriate dispositions or whether they are best cultivated by some other means, such as detailed attention to the situation. I consider W.D. Ross's argument that although the virtuousness of acts is primary, virtuous dispositions also have some intrinsic value as dispositions. However, I argue that if we fit character into the occurrent-state view it is unclear what role such dispositions are playing. I conclude that there is some appeal to the idea that our everyday moral assessments of particular acts do not depend upon persistent character traits, yet I wish to find a method of consistently retaining the intrinsic value of virtuous character.

1. Why object to the traditional disposition based virtue ethics?

Hurka defines two uses of the concepts of virtue and vice (2006, 69-70). He thinks that the concepts are applied at a global level when virtue or vice is attributed to people or their stable character traits. He thinks that we also apply the concepts at a local level to the specific acts or mental states of individuals, so a particular act, motive or desire can also be virtuous or vicious. He argues that of course these two uses are connected, because we expect a virtuous person to have virtuous character traits and desires and for them to perform virtuous actions. He identifies two different methods of making this connection. The first he calls the dispositional view; according to which the global use of the concepts of virtue and vice is primary and what are considered virtuous acts and desires derive from these dispositions. He attributes the dispositional view to Aristotle because of his requirement that a virtuous act has to be done 'from a firm and unchanging state' rather than from changeable, temporary motives (Aristotle *Nicomachean Ethics* 1105a27-b1). He takes this view to need an independent definition of a virtuous character trait from which to derive an account of virtuous acts and desires. The second he calls the 'occurrent-state

'view' according to which the local use of the concepts of virtue and vice is primary and virtuous dispositions are identified as those that give rise to such occurrent virtuous acts and desires.[4]

Hurka thinks that the distinction between the two views is important because they disagree about particular cases. He gives the example of a person promoting the pleasure of another 'from an occurrent desire for that pleasure for its own sake' even though the person 'does not normally have such desires and therefore now acts out of character' (2006, 70). Under the occurrent-state view, this action would be considered virtuous because it derives from a virtuous motive. However, the dispositional view would not consider this action virtuous because it does not derive from a stable character trait.

Hurka thinks that in our everyday use of moral concepts we judge a particular act based upon the motive behind it. He thinks that in applying such judgements we refer only to the occurrent states of the perpetrator: 'that attribution concerns only her current motives, apart from any connection to longer-lasting traits' (2006, 71). He thinks that when making an everyday judgement we do not take into account the behaviour of the person at other times when making a judgement about the virtuousness of this particular act. He identifies the core disagreement between the two views as being over whether it is dispositions that are primarily good or whether it is the occurrent states. For Aristotle, occurrent states are virtuous when they are grounded in a virtuous disposition, so dispositions are identified as primarily good.

Given that our everyday evaluation of actions does not appear to take into account an individual's behaviour on other occasions, Hurka questions whether an act performed from an occurrent motive is worse than one that issued from a stable trait. He says 'we must imagine two acts with the same occurrent motive, say, the same desire for another's pleasure for its own sake, with the same motivational force, but where one desire issues from a stable trait of character and the other does not' (2006, 73). He sees no reason for the former act to be more valuable than the latter.

---

[4] Judith Jarvis Thomson also puts forward an alternative act-based view (Harman and Thomson 1996; Thomson 1997). She uses the example of being just and asks what is it that all entities that are just have in common and whether there are entities that are just only by derivation (1997, 280). She argues that acts are prior; for example, being just is what all just acts have in common and that people are just derivatively because just people are those who perform just acts. She thinks that the act being just has metaphysical priority. For example, the character trait of being generous consists in proneness to perform generous acts (1997, 281).

Hurka considers whether an advocate of the dispositional view can object on the grounds that an in-character motive cannot be the same as an out-of-character motive. He thinks that an advocate of the dispositional view could try to defend his position by arguing that an out-of-character motive could occur on a whim, whereas the in-character motive derives from a stable and deep motivational force that has value. He rejects this objection on the grounds that stability and depth of motivation are not necessarily connected. He says 'someone can be stably disposed to act from a motive that is quite weak, so long as the motives that ever oppose it are weaker' (2006, 74). He also rejects this objection on the grounds that a connection between stability and depth does not provide evidence for dispositions. To support this claim he says 'a soldier who has previously been timorous can now want to save his comrades...and can care deeply enough about doing so that he sacrifices his life' (2006, 74).[5] He thinks that the dispositional view is treating dispositions as evidence for depth and strength in occurrent states, but rather the primary intrinsic value is being found in those occurrent states.

Hurka thinks that when we make everyday global judgements about virtue, such as saying that 'a given person is brave or has the standing trait of generosity' we derive those judgements from 'local judgements about the virtuousness of particular acts, desires, and feelings, and takes those states' virtuousness to be independent of any tie to dispositions' (2006, 74). He does not think that our everyday use of the concepts of virtue and vice places dispositions at the centre of our evaluation of certain acts (2006, 75). A virtue ethicist would dispute that when we make a judgement about virtue we are deriving this judgement from evaluations of particular acts and desires. They would argue that when we make such a judgment in everyday morality we are meaning to attribute a certain disposition to an individual. There is some disagreement here as to what our 'everyday' use of virtue and vice concepts entails, which will not be easily resolved.

2. What is Hurka's alternative to a traditional virtue ethicist theory?

Hurka argues for a position that gives virtue intrinsic value, but argues that this value is consequentialist. He defines consequentialism as the claim that right actions are identified by the quantity of good and evil in their outcome (2001, 4). Under such a theory, he says,

---

[5] Harman agrees with the general position that 'we use the terminology of particular virtues and vices not only to specify character traits but also to describe particular acts' (1999a, 4). He thinks 'that a person who is not generally honest or dishonest may yet act honestly or dishonestly on a particular occasion' (1999a, 4). He agrees with this approach, starting with the virtue or vice of particular actions rather than with the virtue or vice of the character traits that the person possesses.

right action is defined in terms of central properties of good and evil, and good is defined as what people have reason to desire (2001, 4-5). The aim of his argument is to show that virtue is intrinsically good and that, therefore, people have reason to desire it.

He develops what he calls a recursive characterisation of good and evil, which has seven main clauses:

1. Certain states of affairs other than virtue are intrinsically good; he includes pleasure, knowledge, and achievement in this category (2001, 11-12).

2. Recursive attitude: 'if x is intrinsically good, loving x (desiring, pursuing, or taking pleasure in x) for itself is also intrinsically good' (2001, 13).

He characterises loving x as having a 'positive orientation' towards it in one's attitudes: 'one can love x by desiring or wishing for it when it does not obtain, by actively pursuing it to make it obtain, or by taking pleasure in it when it does obtain' (2001, 13). By loving x for itself, he means that we love x 'for its own sake', regardless of what consequences x may have (2001, 14).

3. Certain states of affairs other than vice are intrinsically evil; he includes pain, false belief, and failure in the pursuit of achievement in this category (2001, 15).

4. Recursive attitude: 'if x is intrinsically evil, loving x for itself is also intrinsically evil' (2001, 16).

5. Recursive attitude: 'if x is intrinsically good, hating x (desiring, pursuing x's not obtaining or being pained by x's obtaining) for itself is intrinsically evil' (2001, 16).

6. Recursive attitude: 'if x is intrinsically evil, hating x for itself is intrinsically good' (2001, 16).

7. Instrumental clause: 'if x is intrinsically good because it promotes intrinsic good y, loving x because it promotes y is intrinsically good' (2001, 17). This clause allows us to love x as a means and that loving such a means is itself intrinsically good.

He continues to use the recursive characterisation of good and evil to define virtue and vice. He says, 'the moral virtues are those attitudes to goods and evils that are intrinsically good, and the moral vices are those attitudes to goods and evils that are intrinsically evil' (2001, 20). So, under his definitions, virtues are appropriate attitudes to intrinsic goods and evils rather than dispositions. Such a definition of virtue and vice in terms of attitudes places those attitudes in the explanatory role when explaining our actions, rather than dispositions. He argues that this account is similar to the Aristotelian account in that the value of the attitude depends upon the value of its object (2001, 23-4). Aristotle says

'activities differ in respect of goodness and badness, and some are worthy to be chosen, others to be avoided, and others neutral, so, too, are the pleasures, for to each activity there is a proper pleasure. The pleasure proper to a worthy activity is good and that proper to an unworthy activity bad' (*Nicomachean Ethics* 1175b24-30 cited by Hurka 2001, 23). Hurka interprets this to mean that, for Aristotle, pleasure itself has no intrinsic value, but that pleasure in good activity is intrinsically good. However, Hurka identifies two differences between his own account and that of Aristotle. First, Aristotle does not argue that being pleased by another's good action is also good because his account centres upon an individual's own action. And secondly, Aristotle's account only claims that a good action can be the object of a good pleasure, whereas Hurka believes that further pleasure in that good pleasure can be good, i.e. good pleasure can itself be an object of pleasure, not just good action.

Hurka argues that the recursive characterization of good and evil and the definition of virtue and vice 'contain attractive general principles, make attractive particular claims, and use the former to explain and illuminate the latter' (2001, 29). He thinks that the recursive characterisation of good and evil has merit because the idea that it is intrinsically good to desire what is intrinsically good has intuitive appeal (2001, 30). He thinks that our common view of morality does consider that a good desire can be an object of desire. He thinks that his definition of vice and virtue also has merit because it fits with our understanding of virtues as desirable states of persons (2001, 40). He says 'it gives virtue an active form, involving a person's intentional behaviour' (2001, 41). He also thinks that it captures what makes an action virtuous: 'because of their connection to inner states such as motives and desires' (2001, 41).

He continues to consider whether it is a problem that his account of virtue and vice does not involve stable dispositions (2001, 42). He characterises his definition of virtue as treating it 'atomistically', 'finding it in occurrent desires, actions and feelings regardless of their connection to more permanent traits of character' (2001, 42). He thinks that it would be possible to amend his account to include dispositions: 'alongside the values in occurrent attitudes to goods and evils, there are further intrinsic values in dispositions to have these attitudes' (2001, 42). He thinks that making such an amendment would allow his account to state that 'there is a separate and greater value in occurrent good attitudes...but it can...find some value and virtue in appropriate dispositions' (2001, 43). So, it could to

some extent agree with the Aristotelian picture in that it can recognise the intrinsic value of virtuous dispositions.

However, Hurka does not agree with Aristotle that virtuous action has to issue from a stable disposition. He thinks that if an individual acts kindly from a stable disposition to be kind and another acts kindly once, he thinks that each action is equally kind and therefore virtuous. He thinks it possible that the first individual who has a disposition to be kind will, based upon this disposition, act kindly on many other occasions, so on this ground he may be considered more virtuous than the latter individual. He argues that:

> ...on any view, A's disposition is at least instrumentally good. And we have just allowed that his disposition may be to a degree intrinsically good, that is, good just as a disposition...A's kind action may be accompanied by more valuable actions at other times in his life and by a more valuable disposition now, but I do not see that it is any better in itself (2001, 43).

This may be true, but the debate has shifted. What is of central importance is the question above of what is the best account of the value of these things. Hurka believes that virtue is found in occurrent attitudes, as well as in dispositions. Dispositions have mainly instrumental value in causing a person to act well on more occasions, but the value of the individual acts comes from the individual occurrent states, not the disposition. This is the central claim and I shall consider this idea in more detail later.

Hurka argues that another advantage of his recursive account is that it can be amended to accommodate degrees of virtue and vice. He thinks that the intensity of a desire can differ and that the intensity of an attitude can affect its value and that the objects of attitudes can have different values, hence some attitudes are more virtuous than others (2001, 58). He argues that 'if what is good is responding appropriately to values, one should respond more intensely to what has greater value' (2001, 60). He considers a linear view according to which more intense loves of goods are always better (2001, 60-2). He identifies a problem with a simple linear account of the intensity of values in that it would give indifference a zero value, being neither good nor bad (2001, 62-3). He does not think that this is correct because indifference can be evil; he gives the example of imagining 'that B knows A is suffering intense pain but feels no compassion whatsoever for A...His indifference to another's evil seems to involve not just the absence of a good response, but the presence of a bad one' (2001, 63). To avoid this problem, he adds two clauses to his recursive account:

8. 'If x is intrinsically good, being indifferent to x (neither loving nor hating x when, given one's cognitive states, one could do so) for itself is intrinsically evil' (2001, 63).

9. 'If x is intrinsically evil, being indifferent to x for itself is intrinsically evil' (2001, 63).

He continues to argue that to maintain a scale of value, if we believe indifference to be intrinsically evil, we must maintain that 'very mild loves of goods and hatreds of evils, even if appropriately oriented, can be intrinsically evil' (2001, 64). He suggests that there is a threshold at which the intensity of an attitude becomes appropriate, so the two new clauses need to be rewritten to make not just indifference intrinsically evil, but also inadequate love and hatred intrinsically evil. Given such a view, he considers how one should divide love between different goods. He argues that the proportionality view of division give the best account: 'the best division of love between two goods is proportioned to their degrees of goodness, with as much more for the greater good as its value exceeds that of the lesser' (2001, 68).

He analyses whether the linear view, according to which more intense loves of goods are always better, is correct. He argues that the non-linear asymptotic view gives a better account of how a more intense love of a good is always better than a less intense love (2001, 71). This account claims that the value of an increase in intensity gets smaller as the love's intensity increases. The value of the increase in intensity approaches zero, so there is an upper limit on the value love of that good can have. He argues that this view is compatible with the proportionality view of division, outlined above. He thinks that an advantage of the asymptotic view is that 'appropriately oriented attitudes can be in some way excessive' (2001, 74). He argues that this is accommodated because a person who loves a trivial good very intensely is not directing her love to greater goods that would be intrinsically better.

He thinks that another advantage of the asymptotic view is that it can accommodate the idea that attitudes to the neutral are themselves neutral, yet can be evil in combination with other attitudes. He gives the example of 'her desire for fame is in itself neutral, but its being more intense than her desire for knowledge makes for evil in her attitudes as a combination' (2001, 89). On these grounds, he also adds to his account a further clause, the proportionality principle, which states that 'if x is n times as intrinsically good as y, loving x for itself any more or less than n times as intensely as y is intrinsically evil as a combination' (2001, 84). He argues that the combination of the recursive clauses, the asymptotic view about the degrees of value in individual attitudes and the proportionality principle about the degrees of value in combinations of attitudes give us a series of attractive claims about attitudes. He thinks that the recursive clauses are important so that 'certain attitudes, such as loving goods and hating evils, are intrinsically good and others

intrinsically evil' (2001, 91). He thinks that the asymptotic view is attractive because it allows that 'an intrinsically good attitude can be instrumentally evil if, by being disproportionally intense, it prevents a person from having another, better attitude' (2001, 91). And that, given the proportionality principle, 'an intrinsically good but disproportionate attitude can be instrumentally evil in the stronger sense of making for intrinsic evil in a person's combination of attitudes as a combination' (2001, 91).

3. Problems with Hurka's account

This section introduces five problems with Hurka's account. First, as it is dependent upon a consequentialist account of value, it faces some of the same issues faced by consequentialism. Secondly, it is not clear that the virtuousness or viciousness of acts has conceptual priority. Thirdly, the value of dispositions is unclear. Fourthly, it is not apparent that the value of occurrent states is always greater than the value of dispositions. Finally, in everyday moral discourse we *do* regard both actions and character as having moral value.

First, this is not a distinct virtue theory, but a sophisticated form of consequentialism. Therefore, this theory faces some of the same problems faced by consequentialism. Hurka argues against the view that virtue is a more important good than any other. He summarises his claim in the comparative principle: 'the degree of intrinsic goodness or evil of an attitude to x is always less than the degree of goodness or evil of x' (2001, 133). An implication of this principle is that the sum of hundreds of compassionate attitudes can outweigh an evil. Hurka accepts the conclusion about numbers, but I am not convinced that this is an implication that we should accept. The value of attitudes does not seem to aggregate in this way; each individual compassionate attitude is a lesser good than the pain, or at least it can be, in many different scenarios that we can easily imagine. This remains so even if there are hundreds of people with this attitude. He argues that compassionate pain can sometimes be on balance evil and that this fits our practice of sometimes not revealing our hurts to friends (2001, 145 and 149). He argues that compassionate pain is good as an attitude to its object, but bad as a pain, so if intense, can be on balance evil.

I don't think that this is true; the intense compassionate pain will be good if it is the appropriate attitude to the object i.e. the amount of good and amount of pain will intensify in tandem. If the pain is more intense than the good, the attitude is inappropriate, so is itself

bad. Yes, we sometimes decide not to reveal our hurts to friends because it will cause them pain, but if we did reveal our hurts, their painful response would be appropriate and therefore good, even though painful. Life is more complicated than Hurka allows for, as is friendship. The compassionate pain itself is never on balance evil, but we can decide not to put people in a situation where it would be appropriate to feel this. A situation with no compassionate pain is better than one that contains compassionate pain, but we are withholding the truth from our friends about a serious hurt and, if it ultimately comes to light, they will feel more pain for not having been told. So although we do withhold hurts from our friends, perhaps this in itself is, or could be, bad, because by trying to protect them from pain we cause more pain. We are withholding knowledge, which is in itself intrinsically good, according to Hurka. It is the false belief caused by withholding the knowledge being a good that causes the problem here, not the belief that the compassion will be more painful than good. If the hurt was not serious, it is not clear that our friends' compassion would be painful at all.

Hurka suggests that we can avoid the conclusion that pleasure in great evil, such as the Holocaust, can be on balance good by making pleasure asymptotic: 'the value of an extra unit of intensity in a pleasure gets smaller as the pleasure's intensity increases, diminishing asymptotically toward zero' (2001, 151). He argues that

> if the recursive account retains a linear view of pain and holds that the maximum value of an
> attitude is always the same fraction of its object's value, it holds that the moral evil of loving a
> pain that is considerably more intense is always considerably greater...some pleasures in pain
> can never be on balance good (2001, 151).

This may work for the good of pleasure, but what about knowledge and achievement? Does the value of an extra unit of these diminish as the level of knowledge or achievement increases?

Hurka does not think the value of an attitude of pleasure is affected if it is based upon a false belief (2001, 162). I find this idea a little confusing, for if knowledge and true belief is an intrinsic good, it would seem that an attitude towards a false belief must also itself be, to a certain extent, bad, hence its value is detrimentally affected. Yes, the pleasure may still be good, but overall the situation is bad because of the false belief. It is unclear whether one intrinsic good, such as pleasure, can be valued independently of the others. I would argue that the truth or falsity of the belief seems very important to the pleasure. Pleasure has most value if it is based upon a true belief; the good value of the pleasure diminishes if based upon a false belief. If the pleasure in question is evil, this does not diminish if based

upon a false belief, as here we have a situation containing two evils, false belief and inappropriate pleasure. I think that Hurka ultimately agrees with these concerns when he says that '...in many cases of moral error...there is at least a lack of serious attention to questions of value that involve some vice of moral indifference' (2001, 178-80). In this case, there may be some small good in an attitude, but this is outweighed by the content of the belief. So the value of an attitude of pleasure is not affected if the belief is false, but the overall value of the attitude will be evil because of the content of the belief or moral indifference. However, my question remains as to whether we can value the pleasure independently of other attitudes. Later Hurka claims that 'if an attitude is one of loving, it is good because its object is good' (2001, 189). Why doesn't this apply to false beliefs: if one has a loving attitude towards a bad object (i.e. a false belief), then this loving attitude is itself bad because the object (the false belief) is bad. The love of the object cannot be good because the object is bad.

My second main criticism concerns the conceptual priority of the moral status of acts. It is not clear that the virtuousness or viciousness of acts has conceptual priority. Hurka is attacking a particular Aristotelian version of virtue ethics. However, his account does not fully challenge this position because it does not tackle the question of whether occurrent attitudes are best cultivated by developing the appropriate dispositions or whether they are best cultivated by some other means, such as detailed attention to the situation. He argues that virtue *may* have practical value, as being of good character may be the best way of ensuring that people have the correct occurrent attitudes (2001, 3). Under this interpretation, the character trait of, for example, being kind has a causal relationship with behaviour because it causes me to think and feel in certain ways, which leads me to act in certain ways. This position retains an explanatory role for character traits. Yet, if virtue does have practical significance under this view, then it does not appear to escape the situationist objection that character does not always have the desired effect on our actions. Even if his position is correct, it is no better off against situationist attacks and that is my main concern.

Thirdly, Hurka argues that an individual who has a disposition to be kind will act kindly on many occasions. On this ground, this person may be considered more virtuous than an individual who lacks this disposition; hence virtuous traits have instrumental value. The occurrent states are good, but the best way of ensuring that you have the appropriate desires and act in the appropriate ways is to have a stable disposition to do so. He suggests

it is of no intrinsic importance if a virtuous disposition is a means to produce good, characterising it as having only 'derivative and instrumental significance' (2001, 3). He thinks that 'virtue may be crucial practically, if inculcating it is the best means of ensuring that people fulfil their moral responsibilities. But theoretically it has no intrinsic importance' (2001, 3).

Hurka further argues that a 'disposition may be to a degree intrinsically good, that is, good just as a disposition' (2001, 43). So a person with the disposition to act kindly is more virtuous than an individual who does not have this disposition even if neither of them is currently acting. He argues that the occurrent states view is consistent with finding value in virtuous dispositions (2006, 73). However, if we fit dispositions into the occurrent-state view, as Hurka suggests we might, it is unclear what role such dispositions are playing. If both virtuous dispositions and acts have intrinsic value, it is unclear as to why evaluations of acts are primary. Are dispositions primarily good or the occurrent states? Could it be that the occurrent states are good, but the best way of ensuring that you have the appropriate desires and act in the appropriate ways is to have a stable disposition to do so? The stable disposition has derivative and instrumental value, but might we also intrinsically value virtuous individuals?

Hurka bases his account of the occurrent states view being compatible with the view that virtuous dispositions have intrinsic value on an interpretation of W.D. Ross, who said that 'the state of mind of a habitually unselfish person is intrinsically better than that of a habitually selfish one even when neither is exercising his disposition' (Ross 1939, 291-92 cited in Hurka 2006, 73). Ross argues that virtuous dispositions and actions are intrinsically good. He thinks that actions (or dispositions to act) from morally good motives have intrinsic goodness. He lists these morally good motives as being 'the desire to do one's duty, the desire to bring into being something that is good, and the desire to give pleasure or save pain to others' (1930, 134). He thinks that such actions and dispositions have value apart from their consequences. He argues this to be the case as, if there were two worlds that contained exactly the same amount of pleasure, but in one the people were virtuous and in the other vicious, we would think the former a better world.

Hurka interprets Ross as arguing that the local uses of traits as applied to particular acts or mental states are primary. By taking this use to be primary, Hurka thinks that Ross defines virtuous dispositions as 'dispositions to perform virtuous acts and to have virtuous desires

and feelings' (2006, 70). Ross identifies morally good actions as those that arise from a desire to do your duty, from a desire to bring into being something good and from a desire to produce some pleasure or to prevent some pain for another (1930, 160). This second category, he says, includes actions such as those aiming at improving our own character and at improving our own intellectual capacity. These desires that bring about morally good acts are occurrent states, so Ross identifies occurrent virtuous desires as the motives that bring about virtuous action and then defines virtuous dispositions in terms of these acts and desires; for example, generosity is a disposition to act generously or to have generous desires (Ross 1930, 161-2). Hurka takes this view to need an independent definition of a virtuous occurrent state from which a definition of a virtuous disposition can be derived (2006, 70). He takes Ross to provide this with his claim that it is a virtuous motive that means an act is also itself virtuous, rather than the consequences of the action. Under this position, if an action is done from an occurrent desire to promote the pleasure of another, the action would be virtuous, regardless of whether the individual regularly has such desires, i.e. actions done 'out of character' still have virtue.

Ross later argues that although the virtuousness of acts is primary, virtuous dispositions also have some intrinsic value as dispositions (1939, 291-2). Consider again Ross's example of selfishness; even if two individuals are not exercising their dispositions, the mind of the habitually selfish person is intrinsically worse than that of the unselfish person. Ross says that it is not just 'acts of will, desire, and emotion' that are morally good, but also 'relatively permanent modifications of character even when these are not being exercised' (1939, 291-2). Hurka interprets Ross as thinking that the value of the occurrent acts and desires that arise from these dispositions will always be greater than the value of the dispositions themselves, hence his claim that the occurrent states are primary (2006, 73). Hurka also thinks Ross says that the value of these individual occurrent states is independent of any consistent trait. That Ross argues that dispositions have some intrinsic value gives them value over and above the instrumental value whereby having such a disposition makes the individual more likely to act virtuously or to have virtuous desires.

This leads to the fourth problem with this view. Hurka claims that the value of the occurrent acts and desires that arise from virtuous dispositions will always be greater than the value of the dispositions themselves, hence his claim that the occurrent states are primary. Although the occurrent states may well be temporally primary in the sense that it is necessary that an individual has acted in a certain way for the disposition to be attributed

to them, i.e. the action must come first, it is not clear that we value particular virtuous actions more than virtuous individuals; the actions are not primary in an evaluative sense. Do we value the kindness found in an individual action, such as helping an old lady with her shopping, over and above the kindness of the individual doing the helping? I would suggest not; if we value kindness, we value the kindness of individual actions and the kindness of individual agents equally.

Of course acts *can* be evaluated without reference to the character of the individual as Hurka suggests. He argues that there is no evaluative difference between two acts with the same occurrent motive of desiring another's pleasure for its own sake, where one motive is grounded in a stable character trait and the other is not (2006, 73). What is at issue is the role that stable character traits are playing if it is the evaluation of occurrent motives and acts that is primary. There is a sense in which the occurrent states come first as it is necessary that an individual has acted in a certain way for the disposition to be attributed to them. Although the individual acts are essential for the disposition to exist in the individual, the disposition is essential to link the individual acts together into a structured whole and the whole is often relevant to evaluation of particular acts.

Finally, in everyday moral discourse we do regard both actions and character as having moral value. We often do want to attribute general traits to people and do not tend to see these general traits as simply a function of the previous actions that they have performed, but as having value themselves. We often do attribute general traits to people and take these attributions to be good explanations of why an individual acts in a certain way. If using a character trait to explain George's action, we would say 'George did that because he is courageous when sailing in rough weather with his friends'. It is common to take this particular action as an example of his general trait. In everyday discourse we often *do* assume that people have these general traits and that they explain particular actions, challenging Hurka's intuition that in everyday discourse we derive these general judgements from the particular actions.

What is thus far unclear is how the particular actions add up to the overall dispositions. I agree with Hurka's claim that when we make everyday global judgements about virtue, such as saying that 'a given person is brave or has the standing trait of generosity' we often derive those judgements from 'local judgements about the virtuousness of particular acts, desires, and feelings' but wish to deny his claim that those states' virtuousness is

'independent of any tie to dispositions' (2006, 74). I agree that sometimes our everyday evaluations of actions as virtuous do not consider that the virtuous action has to issue from a stable disposition. But, and this is the crucial issue, how might the individual actions add up to general dispositions? That dispositions only have derivative value, merely comprising the relevant individual acts, does not follow our everyday moral language.

We pick out patterns of behaviour in individuals that are relevant to our evaluation of particular actions. For example, consider the action of walking past someone who needs help because the agent is in a hurry. First, is this hurrying a feature of that particular situation or a characteristic of the person? It is possible for a situation to be engineered such that the agent is made to be in a hurry.[6] However, it is also possible for being in a hurry to be characteristic of the individual. If they are the type of person to be disorganised or to take on too many commitments, then it may be characteristic of the person that they be in a hurry. So, in answer to the first question, the fact that someone is hurrying can be a feature of the particular situation, or his hurrying can be characteristic of the person, or both, where the distinction between person and situation is blurred.

Secondly, does this question matter for our evaluation of that particular action? It is relevant to our evaluation of the action of walking past someone in need of help whether this is part of a regular pattern of behaviour. This becomes apparent if we consider attributions of blame. We may blame an individual less for acting in this way if the agent is rarely in a hurry and this action is out of character, whereas we may attribute more blame to an individual if this action is characteristic behaviour. Although this may on the face of it look like the same action, it is not possible to fully evaluate the action in isolation from the person. The degree of viciousness of the action is affected by both features of the particular situation and general characteristics of the person.

Conclusion

To summarise, I agree that virtuous dispositions also have some intrinsic value as dispositions. This becomes clear in Ross's example of selfishness; even if two individuals are not exercising their dispositions, the mind of the habitually selfish person is intrinsically worse than that of the unselfish person (1939, 291-2). It is clear that in everyday moral discourse, we regard both actions and character as having moral value. We

---

[6] John Darley and Daniel Batson create such a situation in their research based upon the Good Samaritan parable in the Bible (Darley and Batson 1973). This will be discussed in more detail in the next chapter.

often do want to attribute general traits to people and do not tend to see these general traits as simply a summary of the previous actions that they have performed, but as having value themselves. I disagree with the claim that the occurrent states are primary because they always have a greater value than the disposition from which they arise. I shall return to the question of how the individual actions add up to general dispositions in more detail later to explain why we value the dispositions over and above the individual actions, challenging the view that all we really care about are the actions.

I propose that as an alternative to understanding character traits as dispositions, that statements of character take a narrative form. In chapters seven to ten I will develop the argument that character traits have an explanatory, and therefore causal, role because they have the same form as the narrative story that explains why I act in a certain way. For example, my helpful action may traditionally be explained by my desire to help coupled with an appropriate belief. However, there is a narrative-historical explanation as to why I have this particular belief and desire in this specific situation. Under this interpretation, the character trait of being kind has a causal relationship with behaviour because its description is the story that explains what causes me to think and feel in certain ways, which leads me to act in certain ways. This position allows character traits to have an explanatory role in action whilst allowing a central role for individual acts. Individual acts will constitute the description of an agent's character and will constitute the narrative-historical explanation of particular acts; the distinction between evaluation and explanation of actions and character becomes blurred, as each is dependent on the other.

Ross defines virtuous dispositions in terms of occurrent acts and desires; for example, generosity is a disposition to act generously or to have generous desires (Ross 1930, 161-2). If generosity is a disposition to act generously or to have generous desires, it could be argued that the concept of a dispositional trait just *is* a narrative; the narrative describing the individual's past behaviour and desires is constitutive of the character trait and becomes more complex over time. The character trait would not exist over and above the narrative that may be told about it and the only way to attribute a character trait to an individual is to tell a story about it. If attributions of a character trait are taken to be summaries of past behaviour are character traits a narrative construction and not a disposition? What implications might this have for virtue ethics? Before developing this idea, I will first consider some objections to the traditional conception of character traits, or virtues, as dispositions that persist through time.

Introduction

The argumentative aim of this chapter is to summarise challenges to the argument of Harman that there are no such things as character traits and that in ordinary moral thought we are making the 'fundamental attribution error' in explaining actions as the result of an individual's character traits. This is important because if he is correct and there are in fact no character traits, a character-based ethics is seriously undermined. My methodology is to consider whether the social psychology experiments provide evidence to support his argument. I argue that his argument fails for two reasons. First, I argue that the experimental evidence Harman uses is open to interpretation and that the most sensible interpretation does not support his conclusions. Secondly, I argue that there is some ambiguity around the notion of a character trait that needs to be settled to establish whether the social psychologist has in mind the same phenomenon as the virtue ethicist. I conclude that Harman's argument fails because the best interpretation of the situationist position is that although the experiments may cast doubt on the efficacy of character traits in determining action, they do not support the claim that there are no character traits at all. However, two questions emerge from these discussions that remain to be answered: might a more modest argument based upon this evidence still cause problems for character-based virtue ethics and, more generally, what is a character trait? This chapter does not aim to improve on others' criticisms of Harman. However, after encountering Doris's argument in the next chapter, I will add some new criticisms of my own in the subsequent two chapters.

1. The Fundamental Attribution Error

Harman raises a problem as to whether there are stable character traits. He questions whether it is a person's character traits that determine his action or whether it is features of the particular situation. He argues for the extreme view that there are no such things as character traits and that in ordinary moral thought we are making the 'fundamental attribution error' (1999b, 316). This means that we make the error of ignoring situational factors and instead assume that actions are the result of someone's character traits. In making this argument, he denies that character traits can function as explanations for action. The argument runs as follows: there is no evidence that people have character traits, so we are wrong to explain actions on the basis of character traits. In fact, he argues that

attribution of character traits blinds us to what is really important and may lead us to the incorrect action.

No moral theorist will deny the importance of the situation and context in explanations of moral actions. What is at issue here is whether there are character traits or virtues and vices underlying these actions. Harman defines a character trait as a 'long-term, stable disposition to act in distinctive ways' (1999b, 317). He states that 'we ordinarily suppose that a person's character traits help to explain at least some things that the person does. The honest person tries to return the wallet because he or she is honest. The person who pockets the contents of the wallet and throws the rest of the wallet away does so because he or she is dishonest' (1999b, 317). He argues that 'virtue ethics presupposes that there are character traits of the relevant sort, that people differ in what character traits they have, and these traits help explain differences in the way people behave' (1999b, 319). He thinks instead that consistent behaviour can be explained without reference to character traits: 'individuals may behave in consistent ways that distinguish them from their peers not because of their enduring predispositions to be friendly, dependent, aggressive or the like, but rather because they are pursuing consistent goals using consistent strategies in the light of consistent ways of interpreting their social world' (Harman 1999b, 220-221, quoting Ross and Nisbett 1991, 20).

Harman uses two experiments to provide evidence for his view that there are no character traits. In one, John Darley and Daniel Batson conducted research based upon the Good Samaritan parable in the Bible (Darley and Batson 1973). This parable suggests situational and personality differences between those who didn't stop to help and the one who did. The priest and the Levite were probably thinking about religious matters and in a hurry whereas the Samaritan could have been thinking about anything and was less likely to be in a hurry. Darley and Batson decided to test whether it was the case that thinking about a subject of a religious nature, their personal dispositions or being in a hurry had the greater influence on whether an individual stopped to help someone. They did this by asking a group of seminary students to participate in a study on religious education and vocations. They were first given a questionnaire to test their religious personality. They were then asked to move to a different building for the second part of the test in which they were to give a talk. Between these two buildings the subjects passed a 'victim' slumped in an alleyway. It was observed whether and how the subjects helped the victim. The variables were how much the subjects were told to hurry to the next test and the subject of the talk

they were to give. Some were to talk about the Good Samaritan parable and the others about vocations.

The results show that the hurry variable was significantly related to helping behaviour, but what the subjects were thinking about was not. For low hurry 63% offered to help, intermediate hurry 45% and high hurry 10%. The determining factor of whether people stopped and helped was not what subject matter people were thinking about, but whether they were in a hurry. Harman uses this evidence to defend the position that a disposition to be helpful would not be enough to predict what a person would do in such a situation; in fact the particular situation determines how you act.

Harman also cites Stanley Milgram's experiments on obedience to provide evidence to support his claim that features of the particular situation are what are more relevant in determining actions, not character traits (Milgram 1974). There were many variations of this experiment each raising many issues, but here I shall describe just the variation Harman refers to. In this version of the experiment a volunteer comes to a laboratory at Yale University to participate in a study of memory and learning. At the laboratory is an actor pretending to be another volunteer. They draw lots to determine who is the teacher and who is the learner, but unbeknown to the real volunteer the draw is rigged so that he is the teacher. The learner is seated in another room with his arms strapped to a chair and an electrode attached to his wrist. The experimenter tells them both that this is a study into the effects of punishment on learning. The learner is given a list of pair words to learn and recall; whenever he makes an error the teacher is told he must give him electric shocks of increasing intensity. The teacher is seated in a different room with a generator that has switches from 15 Volts to 450 Volts, in 15 Volt increments, labelled from slight shock to severe shock to XXX. Starting at 15 Volts the teacher is told to increase the level each time the learner makes an error. The learner responds to the shock at 75 Volts with a grunt, at 120 Volts he complains verbally, at 150 Volts he demands to be released and at 285 Volts he only screams agonisingly, then falls silent from this point onwards. If reluctant to continue, the teacher is told by the experimenter that he must continue.

In this experiment almost two-thirds of teachers were prompted to give the learner a 450 Volt shock, essentially because they were instructed to do so by an authority figure. Harman states that the intuitive response to these experiments is to conclude that these people are of bad character, but this has the disturbing result that the majority of people are

of bad character. He makes this comment because he believes we regularly use people's character traits as explanations for their actions. He uses the example of explaining that a person will return a lost wallet because he is honest. Analogously, the teacher shocked the learner to 450 Volts because he is bad or evil. He claims that the above experiments show such explanations to be dubious. The teacher did not shock the pupil because he is bad but because of particular features of the situation.[7]

The conclusion Harman wants to draw from these experiments is that character traits do not explain actions, situations do. Lee Ross and Richard Nisbett suggest that it is situational factors that give the reasons for action; in the Milgram case these are the stepped nature of the experiment, the difficulty of stopping and the non-sensical situation (1991, 56-58 quoted by Harman 1999b, 322-23). Harman concludes that we consider these people to have a character defect, that of destructive obedience. This is interesting, as presumably obedience is a trait usually to be praised, so he has to bring in the evaluative term 'destructive' to make his point.[8]

Harman also argues that using character traits to explain action can cloud our judgement as to the real reasons. He takes the example of violence in the former Yugoslavia and says that if we attribute this to ethnic hatred rather than understanding how the violence arose out of the situation, then we will not see all the possible ways of ending the violence (1999b, 329). He reaffirms his claim that it is better not to talk in terms of character traits because of the negative effects it can have on how we understand a situation (2000, 224). However, this is highlighting the difficulties of correctly attributing character traits, rather than providing an argument for their non-existence. These experiments suggest either that having a certain character trait is insufficient for the correct moral appreciation of the particular case, or that character traits do not exist at all. Harman is arguing the latter; his argument is based upon finding examples of people seemingly acting 'out of character' and taking this evidence to suggest there are no character traits at all.

---

[7] There may be an issue here related to a divergence between an agent's explanation of his own action and an observer's explanation of that same action. It may be the case that in explaining our own actions we cite situational reasons, but we tend to explain other's actions in terms of their character traits. There is thus a divergence between first personal and third personal explanations.

[8] Could we borrow some terminology from Ross (1930, Ch. 2) and argue that obedience is 'prima facie' a good reason to act, but in this case is defeated or outweighed? Or is it that obedience is good except when it is destructive i.e. we did not have the full statement of the character trait?

2. Problems with this argument

Jesse Prinz identifies the main responses from virtue ethicists to Harman as either accusing him of misusing the empirical results or misunderstanding virtue ethics (Prinz, forthcoming). I shall consider these approaches in the two sections below. In the first section I will consider three ways in which it has been suggested that the data has been misused. First, that the aim of the experiments was not to prove that character traits did not exist, so methodologically crucial information is missing from the experiments. Secondly, that the experiments show a conflict of character traits rather than their non-existence. Thirdly, that the evidence that people often incorrectly attribute character traits is consistent with the claims of virtue ethics. In the second section I will consider whether the virtue ethicist and the social psychologist are operating with the same concept of 'character'.

i. Misuse of the empirical results

I don't think the aforementioned intuition that the people in the Milgram experiment are of bad character is correct. In fact, the results of the experiment are all the more worrying because the subjects are ordinary people. Similarly, Christian Miller criticises the conclusion that Harman draws from the Darley and Batson case on the grounds that we would use the students' character to predict what they would do only if we had a reason to believe that they had the relevant character traits, but because we have no grounds for this assumption, he thinks we would take situational factors into account (Miller 2003). I think that we consider individuals bad only when they consistently act in such a way across situations. The Milgram example appears to show a failure of moral wisdom, not that obedience is always wrong, or that the trait does not exist. This example doesn't seem to have proved that the trait of obedience does not exist, just that all agents do not necessarily have the wisdom to apply it correctly to all situations. I think this experiment in fact shows that people (across people) have a very stable character trait of obedience, not that they don't have the trait. This experiment shows that, in general people are obedient, highlighting the problems of having this characteristic and the depth of knowledge needed to apply it correctly in a particular situation, not that the characteristic does not exist. We are criticising people for lack of moral wisdom; i.e. 'a person with the relevant character trait has a long term stable disposition to use the relevant skills in the relevant way' (Harman 1999b, 317); what is failing here is not the long term stable disposition, but the

ability to use the relevant skills in the relevant way.

a) First problem with the empirical results: the aim of the experiments

In Nafsika Athanassoulis's response to Harman's paper she concentrates on the experiments and how they are used (Athanassoulis 2000). Her first objection is that the Milgram experiment did not aim to prove that there was no such thing as character traits (2000, 216). She points out that actually the experiment was designed to investigate the extent to which people would follow orders that ran contra to their normal moral inclinations. The experiment did not aim to prove that character traits do not exist. For this reason, it would not be surprising if this conclusion could not be drawn from the evidence. In an earlier work Flanagan also makes a similar point, stating that this experiment is about the conflict between two traits, compassion and obedience (Flanagan 1991, 293-8).

Athanassoulis argues that the conclusion of the experiment is that, in this situation, most people are obedient, not Harman's conclusion that there are no character traits (2000, 217). She thinks that because the experiment was not designed to find differences in character traits, it has little to say about the long term, which she thinks is inherent in the nature of what it is to be a character trait. I agree that this experiment shows that the majority of people in this situation would on one occasion be obedient. We would need to repeat the same experiment many times with the same people to see if they have a stable character trait of being obedient. Showing that x amount of people are likely to be obedient in one situation says nothing about how they may act in another. Harman, however, extrapolates from the fact that people in general were not compassionate in one situation that people as individuals do not have the character trait of compassion, but this experiment does not prove this as it is conducted at the level of people in general and not the individual level.

Gopal Sreenivasan considers this objection in some detail (Sreenivasan 2002). He introduces the notions of temporal stability and cross-situational consistency used by psychologists (a distinction also referred to by Miller 2003, 367). Temporal stability refers to the reliability of a character trait in specific types of situation, for example, the reliability of cheating in situations where cheating is possible. Cross-situational consistency refers to the reliability of a character trait across different situations, for example the reliability of honesty across situations where cheating or stealing is possible. Generally, psychologists think that there is temporal stability of character traits, but

Sreenivasan argues that it is the cross-situational consistency that is important in the situationist's case against character traits, as it is these broader traits that are referred to in virtue ethics (2002, 55).

Sreenivasan argues that the Milgram experiment and the Darley and Batson experiment only provide evidence of the fundamental attribution error (2002, 52). We make the mistake of attributing a character trait to someone on the basis of one example. He says that a character trait can only properly be attributed to someone if you have observed numerous pieces of evidence. Following from this mistake, people will also make the mistake of predicting what a person would do if confronted with this situation again, based on the one example. He identifies a third error as our failure to distinguish between temporal stability and cross-situational consistency, for example we take one example of being honest by not stealing as evidence that in situations where the person could lie or cheat, they would also act honestly. He allows that we may make the fundamental attribution error, but that this does not determine whether anyone has a character trait, just that we often attribute them incorrectly (2002, 53-4).

In Harman's response to Athanassoulis he seems to accept that the Milgram experiment does not itself challenge the notion of a character trait (2000, 223). He shifts the argument slightly, making a distinction between the first and third person. He argues that a third personal observer of the Milgram experiment is likely to attribute the way someone acts to his character rather than to aspects of the situation. I do not think that Athanassoulis would have a problem with accepting this; in fact she argues along similar lines, saying that the experiments show that individuals attribute character traits without sufficient evidence. The error of ignoring situational factors is compatible with virtue ethics and the existence of character traits; in fact most virtue ethicists would agree that it is an error to explain actions solely by reference to character traits.

In a later paper, Harman again admits that these experiments alone do not show that there are no character traits (2003, 91). He admits that what they show is how important aspects of a situation can be to how someone acts, as I and others have been arguing. He says that observers of actions tend to attribute these actions to character traits rather than to features of the situation and that, therefore, these explanations are erroneous. He emphasises that there are a large number of experiments, not just those he quotes, that demonstrate the importance of the situation, the lack of awareness of the situation by observers and the

tendency towards the fundamental attribution error (2000, 223).[9] Again, I do not think this is a problem for virtue ethicists, as the argument is that the experiments miss their target, highlighting the importance of the situation and the complexity of identifying character traits, rather than proving that they do not exist. It is the interpretation of the experiments that is at issue, not the number. This makes it seem possible to argue that the fundamental attribution error is merely an error of explanation i.e. the error of explaining an action simply in terms of a character trait, rather than a 'full' explanation. This would result in what constitutes an explanation being at issue, rather than a debate about whether character traits exist. Later chapters will return to this question of what type of explanation we need for moral evaluation of actions.

Sreenivasan thinks that it is the one cross-situational experiment that Harman briefly refers to that may provide evidence against broad character traits.[10] The Hugh Hartshorne and Mark May experiments into honesty found that there were very low average correlations of .13 between stealing and lying, .13 between stealing and cheating, and .31 between lying and cheating (Hartshorne and May 1928). Overall, they found only a .23 correlation between any pairs of honesty related behaviour measures.[11] Sreenivasan claims that we have to make two assumptions for this experiment to provide evidence against cross-situational character traits. The first assumption is that, for example, the .13 average correlation applies to an individual case i.e. that a particular individual who steals has a .13 chance of also lying. He argues that we need to make this assumption because the results of the experiment are aggregated across the group that the experiment was conducted on, thus an average .13 correlation between stealing and lying does not exclude the possibility that there are some, although obviously not many, for whom the correlation was much higher. He points out that the aggregated results do not provide an obstacle for some virtue theories in which only some people will have full virtue, with most ordinary people varying in degree of virtue. He claims that the second assumption we have to make is that the 'behavioural measures properly operationalize the character trait honesty', by which he means we need to be clear how we understand honesty as a cross-situational character trait (2002, 56).

Sreenivasan continues to discuss three issues, all of which provide objections to the

---

[9] I will consider some of these additional experiments in the next chapter.

[10] Harman refers to this experiment at 1999b, 326

[11] Sreenivasan explains that this type of correlation across situations is measuring the consistency between 'a set of behavioural responses to one relevant situation and another set of behavioural responses to a second relevant situation' (2002, 49).

relevance of the Hartshorne and May experiment to cross-situational character traits. The first of which is whose specification of which behaviours are honest and which situations call for honesty is important? He uses an example of an individual who believes 'finders keepers', so does not regard taking some loose change as stealing, which makes this action consistent with his honesty, whereas the experimenters do not take his beliefs into account and simply regard the behaviour of taking some loose change as stealing and evidence that he is not honest across situations. The second issue he discusses is the relevance of the behaviour in assessing honesty (2002, 58-9). He suggests that pocketing some loose change is not a paradigmatic case of theft and perhaps a clearer cut situation such as shoplifting would perhaps result in higher correlation.

The third issue he raises is that character traits as found in a virtue theory are not merely responses to situations, but are responses to a reason for action in a situation (2002, 59). He uses two examples to make his case. One is again the Hartshorne and May experiment, where in their lying situation, the lie stops another child getting into trouble. He again claims that this is not a paradigm case of lying, as preventing another child getting into trouble could provide a greater reason for action, so action in this case is not a good behavioural measure of honesty, unless we conceive of an honest person as someone who always acts honestly regardless of the reasons for action in the particular situation. The second example he uses is helping someone in distress, which is paradigmatic compassionate behaviour. He argues that one instance of not helping does not count against them being compassionate or show that they are inconsistent, as the reason to help can be defeated: 'Whenever that reason is defeated, failure to help someone in distress is perfectly consistent (in that case) with the trait of compassion' (2002, 60). Thus even a cross-situational experiment such as this does not provide clear evidence against the existence of character traits.

In summary, what concerns are raised here? First, that the experiments only provide evidence of the fundamental attribution error, but that this error is consistent with the existence of character traits. Secondly, issues as to how a character trait or situation is described have been raised. This relates to questions around what sort of explanation of action is needed for moral evaluation. The agent's beliefs or reasons for action will be of central importance to my later claims about what is important for an explanation of action.

b) Second problem with the empirical results: conflicting character traits

Rachana Kamtekar objects that Harman does not consider that the experiments may show conflicting character traits, for example compassion versus obedience, so whatever action the individual undertook he would have been acting against one of his character traits (2004, 473).[12] As she points out, even though there is only one action that can result in any situation, how one individual arrived at this action will differ from the next. The post-experiment reflections cited by Milgram demonstrate this. Harman would accept this point, but does not view it as a challenge to his argument. He claims that all differences in how people act can be explained without reference to character traits (2000, 223). He further argues in a later paper that even if there are character traits such as obedience and promptness, these general traits cannot explain the differences in the behaviour observed in the above experiments (2003, 91). He gives the example of the difference in obedience displayed by someone who immediately shocks someone at the highest level and someone who is incrementally led to do this. I am not convinced by this example, as it is not his trait of obedience that would explain the difference between someone who shocks the pupil at a high level immediately and someone who is led to it incrementally, but the interaction of his trait of obedience with all his other traits, as well as his feelings, reasons and beliefs. Again, this issue with the use of the empirical results raises questions about what type of explanation of action is required.

c) Third problem with the empirical results: over-optimism

Athanassoulis makes a further objection that the experiments show that people are over-optimistic in their attribution of character traits, not that they do not exist. She continues to argue that Harman's paper is useful in that it points out the problem of attributing character traits when you do not have enough evidence (2000, 218). She thinks that these experiments provide evidence of the Aristotelian concepts of virtue, continence, incontinence and vice. In the case of a continent person, although they may act in the same way as a virtuous person, this was a result of a conflict between reason and desire, and so it is possible to confuse the continent with the virtuous. She claims that the virtuous person has character traits, or long term stable dispositions, but the continent person does not. The virtuous person has the right reasoning and desires, but the continent and incontinent persons face the battle between right reasoning and their desires. She also adds that the

---

[12] Athanassoulis (2000) also argues that the Milgram experiment does in fact show that there are differences between people as they react to the pressurised situation in different ways.

continent and incontinent persons are much more common as these stages are part of our moral development. She continues to link this back to the experiments cited by Harman (2000, 219). She argues that the experiments concentrate only on outward behaviour and not the long term stable character traits behind the action, so it is impossible to distinguish between the virtuous, continent, incontinent and vicious. She argues that depending on the level of an individual's moral development, they will find it easier or harder to resist influence and, as the experiments focus on single events, they can tell us nothing about this.

Athanassoulis reinterprets the Darley and Batson experiment as showing how the situation can affect the manifestation of character traits, not that the situation directly affects action (2000, 219). She argues that people who do not act with compassion when in a hurry are not fully virtuous; a fully virtuous person will always act compassionately, despite external factors and she does not think that it matters for the virtue theorist if there are very few, or even no, people who are fully virtuous. In common with Athanassoulis' paper, Miller agrees that becoming fully virtuous involves a struggle in overcoming character defects and situational distractions, which indicates that most of those in the Milgram experiments were not fully virtuous. He claims that the best conclusion that can be drawn is that people fall short of being fully virtuous and that this is not a problem for virtue ethicists because full virtue is hard to obtain (2003, 379). Athanassoulis presents some different explanations of why people acted as they did in the Darley and Batson experiment:

> ...student A did not help because he wrongly judged that being on time was a greater moral requirement than giving assistance, an error in judgement, whereas student B did not help because although he realised that he was morally obliged to help he could not resist the self-centred desire to present his lecture, a conflict between his reason and his desire. It may even be the case that student C who did help did not really exhibit the virtue of kindness, as his motive was to be recognised as a hero by the student newspaper for his act (2000, 219-20).

These examples of possible explanations highlight the limitations of the experiment in discovering the grounds for an individual's action. Her conclusion is that these experiments suggest that we may attribute character traits to people too easily and without enough evidence. She argues that this conclusion is consistent with our ordinary thinking and with the experiments cited by Harman.

Miller also makes the point that the social psychology experiments give evidence that observers tend to underestimate the effect of the situation, but that this does not cause problems for virtue ethics and does not provide evidence that character traits do not exist

(2003, 371). He thinks that the Milgram experiment causes problems for a virtue ethicist only if they believe that if someone has a virtuous character trait, they will display it on every occasion that calls for it. He does not think that a virtue ethicist would have such a strong, implausible account (2003, 378). Kamtekar also highlights the problems inherent in the experiments that focus on one character trait. By taking a character trait in isolation the experiments are failing to take into account how other character traits may affect or be affected by the situation. She thinks that the experiments make the assumption that if someone has a character trait, it will always be displayed in any relevant situation. She argues that this is an unreasonable claim to make, pointing out that we do not commonly take, for example, one instance of not being helpful as evidence that someone is not a helpful person (2004, 475).

In response, Harman points out that many of the people who do not agree with the conclusions that he draws from the Milgram experiments interpret the results as showing that *most* people are of weak character or will not be compassionate under pressure, when in fact every person in the initial experiment applied shocks in excess of what was expected prior to the experiment (2000, 225). Kamtekar draws attention to the fact that Harman dismisses the possibility of everyone having the character trait of destructive disobedience without argument (2004, 468). Athanassoulis also comments that Harman does not explain why 'destructive obedience' cannot be a character trait (2000, 217). The paper appears to be reliant on the fact that we would not want to attribute this character trait to such a large percentage of, or all, people and that it plays on this intuition to cast doubt on the existence of character traits. Harman plays on our horror that so many people can act in a way that seems morally abhorrent and that we don't want to believe that so many people are 'bad'. However, it is not true of all the subsequent Milgram experiments that everyone displayed the trait of obedience; for example, two people in the Bridgeport version of the experiment refused to give even the smallest shock (1974, 61). Miller also criticises the conclusion that Harman draws from the Milgram experiment on these grounds. He thinks that Harman does not make it clear what character trait is supposed to be lacking and does not explain why he does not think that those who do not comply were exercising compassion and those who do comply were being obedient (2003, 369).

Harman concludes from his evidence that virtue ethics rests on an error because there are no character traits. I don't think that these experiments alone will convince a person who thinks that character traits exist to change his mind. For example, Flanagan considers all

the experiments Harman refers to and does not regard them as evidence for non-existence of character traits (1991, 293-298). Flanagan argues that character traits are 'highly context sensitive' i.e. that they are only appropriate in certain situations and even when appropriate may still be affected by other situational factors (1991, 280-1). Athanassoulis concludes that although experiments in social psychology raise interesting questions and highlight the importance of other factors in moral judgement, they do not threaten virtue ethics. She argues that even if there is less consistency in character traits than we commonly think, this does not undermine the principles of virtue ethics (2000, 217). The notion of having a certain character trait means we 'for the most part' act in accord with it. There will of course be exceptions when other factors cause us to act in other ways.

In summary, these concerns that the evidence does not conclusively show that there are no character traits raise some interesting questions. The evidence suggests that we do make the fundamental attribution error. The fact that we make this error will have implications for using character traits to explain action. What do we demand of an explanation of action? And what role do character traits and situations have in explanations of action? These concerns also illuminate confusion as to what exactly a character trait is that merits further investigation, which will be addressed in later chapters. How is a character trait to be distinguished from the situation?

ii. Misunderstanding Virtue Ethics

Other criticisms focus on whether the situationist arguments misunderstand virtue ethics. It can be inferred from the situationist argument that character traits, even if they exist, are not reliable guides as to how people will act in a given situation. Every situation is different from the next and if it is the situation that explains your action, then your action is not a demonstration of any character trait that you may possess. How you act in one situation does not reliably say anything about how you will act in another. However, an assumption seems to be made that for a character trait to be reliable we must be able to predict what someone would do in any particular situation. But this requirement is too strong. Of course there are times when you act 'out of character'. It appears that, in the Milgram experiment and the Darley and Batson experiment, the one occurrence of acting unhelpfully or obediently does not immediately identify you as a morally repugnant person.

Harman clarifies that if his arguments for the non-existence of character traits succeeds, it delivers a fatal blow to a certain conception of virtue ethics in which the agent is striving to obtain the character traits that a virtuous agent would have (2000, 224). If there are no character traits, it makes no sense to strive to obtain something that does not exist. Maria Merritt thinks that the attacks on 'virtue ethics' are misplaced, as there is not a single conception of virtue ethics (2000, 367). She thinks many of the situationist attacks on virtue ethics approach it as the view that we work out the right thing to do by imagining what the virtuous person would do (2000, 369). She agrees that this may be a useful tool in difficult situations, but, as argued in Chapter One, deliberation about the right action is not traditionally the primary concern of virtue ethics; the primary purpose is to answer questions about how one should live (2000, 370). She continues to argue that traditionally virtues are not supposed to help us live by providing paragons about whom we can theorise to decide on right action; they are supposed to help ourselves be 'reliably responsive' to situations (2000, 371). She argues that reflection on paragons has to be a secondary purpose: 'The ideal life is supposed to be the life of having the virtues, not the life of thinking about other people who have the virtues' (2000, 371).

What is meant by a character trait according to the virtue ethicist and the social psychologist? Julia Annas defines virtues as character traits that are 'dispositions to act, think and feel in certain ways, which are taken by the agent to be the way they are, their character' (2003, 21). She contrasts this with the character traits we find in the experiments, where it is a disposition to produce behaviour, enabling prediction of future actions, making no reference to the person's reasons (2003, 23). Miller defines the character traits referred to by Harman as being long term dispositions, involving habits, which are broad and explanatory. He argues that although this list may be *necessary* for a virtue ethicist it will not be *sufficient* for their richer concept of virtue. This is similar to Kamtekar's argument that the type of character trait attacked in these experiments is not the same concept as the one commonly used in virtue ethics. She thinks that the experiments concentrate on individual dispositions, considered independently from how people reason, whereas the concept used in virtue ethics is more holistic, taking into account character as a whole and how people reason, giving rise to a set of non-conflicting motivations. Inconsistency in behaviour is explained by the not fully virtuous character being swayed by conflicting desires or beliefs rather than the non-existence of character traits. Kamtekar provides a different view of what a character trait is by appealing to an interpretation based largely on Aristotle (2004, 477-486). She argues that the traditional

view of virtues is that they are character traits enabling an individual to respond appropriately to any situation, not to display particular behaviour associated with the character trait in any situation. She argues that Aristotle does not think virtue or vice necessitates particular actions, but that they are character traits that 'incline us to act in particular ways' and that these character traits are not by themselves explanatory (2004, 479). She also makes it clear that central to the Aristotelian concept of virtue is practical wisdom i.e. virtues are not merely character traits to do certain acts, but involve reasoning to do these acts appropriately (2004, 480).

Annas agrees with Kamtekar and others, such as Sreenivasan and Miller, who claim that situationists are wrong about character traits; they attack the view of social psychologists that character traits are dispositions to cause unreflective behaviour. She criticises the notion of a character trait found in the experiments, because she thinks that we do not discover our own character traits by predicting what we may or may not do in given situations. She questions why these accounts do not consider factors such as the individual's reasons for action (2003, 23). Her main interest, however, is investigating what effect situationism has on virtue, with virtue defined as a reliable, habitual disposition. She defines situationism as making a distinction between the person and the situation as explanatory factors and that it is the situation that has the explanatory role. She also defines what we mean by a situation, identifying it as 'an event with features to which a person is (or is not) sensitive at a particular time and place' (2003, 22). She makes it clear that virtue is not a habit, as these are mindless; a virtuous disposition is exercised by practical reasoning. She uses an Aristotelian claim that virtue is a disposition that is built up by making choices and not just a summary of past actions.[13] She adds further that this disposition is used in making choices. She uses the example of an honest person not taking something that is not hers to illustrate her point; this action is not a predictable reflex caused by habit, but a decision that is driven by her honesty and that further establishes her honesty (2003, 24-25).

She breaks down the claim that virtue is a 'disposition to do the right thing for the right reason in the appropriate way' into two parts, the affective and the intellectual (2003, 25). She describes the affective part as the sensitivity to the fact that someone may do the right thing but have varying feelings. She takes the example of classical virtue theories that claim the fully virtuous person always acts without feeling contrary inclinations (2003, 25).

---

[13] Later chapters will build on this idea that virtuous traits are more than a summary of past actions.

She describes the intellectual part of virtue as acting for the right reason and that reason is identified because she has a character that enables her to understand the situation correctly (2003, 25). She develops the picture of virtue within the classical virtue ethics tradition by discussing how we learn to be virtuous (2003, 25-6). She says we start by learning from others, then begin to reflect on what we have learnt to try and unify our judgements and actions, or, in other words, virtue 'requires experience and practice' (2003, 25). She continues to argue that having ethical expertise involves an understanding of what we do and is similar to other sorts of practical knowledge, such as building expertise.

The point of this elaboration of virtue is to make it clear that ethical expertise is sensitive to situations:

> It is only the absolute beginner who does what he does because he has been told to do so, or is copying the expert, and who acts in a way which is not sensitive to the specific demands of the situation. As soon as he develops understanding of what he is doing, he brings to each situation an understanding of what he should do which has been built up by practice, but is active and responsive to what needs to be done now, in this situation (2003, 27).

She uses this conception of virtue to show that the notion of character traits as dispositions to reliably produce behaviour regardless of the situation used by situationists is not the same as virtue (2003, 27). She elaborates two reasons why this is not the case. First, virtue is not a disposition that causes behaviour, but a disposition to decide and, secondly, it is a disposition to respond to features of a particular situation, not act blindly. She makes the point that we will not be able to predict what someone will do unless we have built up the same level of understanding, through experience, as the individual concerned. She says that we must have an understanding of the way an individual thinks: 'When we have no background knowledge of a person, and in particular know nothing whatever about her views on honesty in various different kinds of circumstance, how could we expect to foresee accurately what she will do?' (2003, 28).

Annas argues that virtue requires attention to the situation, so is not dissimilar to what the situationist is telling us (2003, 28). She says the situationists underestimate the intellectual part of virtue:

> The virtuous person not only judges what is the right thing to do, he does this from understanding, something which enables him to criticize the judgements he originally started from, and to explain and give reasons for the judgements he makes. He does this in the light of an understanding of his life as a whole and the workings of both the virtue in question and other virtues to which the situation is relevant (2003, 28).

Prinz is not convinced by this argument, as he thinks that subjects in the experiments do deliberate about what to do and exercise their rationality (Prinz, forthcoming). He argues that if character traits or virtues are a skill used in practical reasoning then they should be available in these cases. He uses the example of justice in the Milgram case, defining the disposition of caring about justice as understanding what actions are unjust and understanding the conditions under which it may be necessary to intervene to maintain justice. He argues that if this disposition was an enduring character trait for some people, then we should expect to see some evidence of it in the Milgram experiment. He states that in the initial experiment, no one acted in accordance with this disposition.

He suggests three possible explanations: that no one has that disposition, the disposition is present but rationally overruled by other beliefs or that the disposition is there but is inert, all of which, he claims, do not help save virtue theory. I find the first rather a grand leap, as the first experiment involved a small sample of individuals and the experiments over the period 1960 to 1963 only involved around 1,000 subjects (Milgram 1974, 1-26). Admittedly, the vast majority of subjects complied, but there were the two people I mentioned earlier who refused to give any shocks at all and I think this is enough to cast doubt on the conclusion that no one has that disposition as here is some evidence of it.

The other two explanations amount to the same thing i.e. the disposition is there but not affecting action for whatever reason. Prinz dismisses the second explanation because he finds it troubling. He thinks that in the experiments, subjects who are obedient or punctual are making a moral mistake and he thinks it is incorrect to argue that there is a reason (i.e. being punctual or obedient) for the subject to act as he does. He illustrates his point by appeal to the Milgram experiment; how can there be a reason to continue being obedient when you believe the learner has been hurt or killed? He argues that even if we can rationally explain why someone obeys an authority figure, even when doing so causes harm, this misses the point, as our ability to make decisions is also affected by features of the situation. Prinz favours the latter explanation, as he thinks that in pressurised situations our ability to reason is impaired by the stressful features of the situation. He thinks that the fact that our reasoning can also be affected by situational factors means that virtue ethics cannot be saved by appeal to the fact that traits/virtues involve practical reasoning. I am not sure that this argument causes any new difficulties for the virtue ethicist, as I think they can admit that situational factors can cause failure of practical reasoning in the less than fully virtuous agent. This goes back to Athanassoulis's arguments that the level of an

individual's moral development, in terms of possessing virtues and the capacity for right reasoning, will affect how much they are influenced by external factors.

The debate between social psychologists and virtue ethicists over what a character trait is will not resolve the dispute as to whether or not the evidence is damaging to virtue ethics. The social psychologist can take on board the virtue ethicists concerns about the rationality of virtuous traits but, as we shall see in the next chapter, can argue that action is of primary importance. They argue that the virtue ethicist has to be concerned with the relationship between virtuous character and action, as what is the point of a virtuous character if it does not lead to right action? The social psychologist can make appeal to Aristotle at this point, as he argued that you cannot be virtuous just by possessing the virtues, but only by acting in accord with them (*Nicomachean Ethics* 1098b30-1099a5). This is not to say that there is no interesting debate to be had over the concept of a virtuous character trait, but the focus of my investigation will be whether the experiments highlight a problem for this particular type of virtue ethics that does demand a reliable causal link between character and action, rather than on constructing an account of virtue ethics that makes no such demand.

3. Conclusion

In conclusion, these experiments cannot be used to deny that we have character traits, just that they alone are not enough to explain our actions. We must also take the situational context into consideration. For this reason, we cannot on the basis of character exactly predict what someone may do in the next situation that demands use of that character trait because it may be over-ridden by a situational factor. The factor does not alter your character trait; it just shows that this alone cannot explain action. So we can still see what a good person would do in a case with no over-riding factors and use this as the basis of morality.[14]

However, there remains the problem that even if the experiments do not establish the non-existence of character traits, can they still cast doubt on the efficacy of character traits in producing right action? A second, and related, problem that has arisen from this discussion is a lack of clarity as to what a character trait is; whether a general trait that causes consistent behaviour across all situations, a more specific trait that causes specific behaviour in specific situations or a disposition to reason in a certain way. The next two

---

[14] A question can be raised here about what I mean by a case with no over-riding factors, or normal circumstances. This will be considered in further detail in Chapter Six.

chapters will discuss the first of these problems, before moving on to consider the second.

Introduction

The argumentative aim of this chapter is to consider the implications for virtue ethics of accepting Doris's argument that character is fragmented and that character traits are specific to particular local situations. This is important because if he is correct and character is fragmented, a traditional character-based ethics that relies upon global character traits is seriously undermined. My methodology is to consider first whether the social psychology experiments provide evidence to support his argument and, secondly, whether the evidence is consistent with virtue ethics. I consider and reject an interpretation by Merritt that argues that the evidence is compatible with a virtue ethics that treats virtuous traits as being socially sustained. Secondly, I consider Jonathan Webber's argument that the evidence provides evidence not for fragmentation of character traits but regularity and that this is compatible with virtue ethics. However, I question whether this in fact answers the situationist's challenge. Thirdly, I reject Peter Vranas's claim that the situationist argument results in indeterminacy. I conclude that the most promising response derives from consideration of how individuals may evaluatively connect the different fragments. In the next two chapters I will suggest some problems of my own with Doris's view that improve upon the responses of the commentators considered in this chapter. It is important to respond to Doris's argument in some detail because he is making a stronger, more appealing claim than Harman. Harman makes a radical claim about the non-existence of character traits, which does not seem to be supported by the evidence, but Doris makes a more modest yet ultimately stronger claim that character traits do not regularly explain behaviour in the way that we perhaps think they do.

1. The sensitivity of behaviour to the particular situation

Doris finds truth in the general situationist claim that behaviour is very sensitive to the particular situation and further that the features of the situation are better predictors of future behaviour than any considerations about character (Doris 2002). His general argument is that character does not in any reliable way determine behaviour and he discusses what implications this has for moral philosophy. He thinks it important that any account of morality is not based upon an empirically discredited view of character, so moral philosophy should take empirical psychology into account (2002, 2). He justifies the

relevance of empirical psychology to moral philosophy because talk of character involves the descriptive, explanatory and evaluative. He draws attention to the fact that moral philosophers regularly make descriptive claims about character and psychology, citing Alasdair MacIntyre as an example (MacIntyre 1984). He argues that this descriptive content means that ethical character can properly be the subject of empirical investigation (2002, 5-6).

Doris defines a character trait as a disposition 'to behave in a certain way in certain eliciting conditions', which he argues is the conception generally used in philosophy (2002, 15). He acknowledges that describing virtue as a behavioural disposition is not sufficient, as the concept of virtue involves emotional response, deliberation and decision making, as well as behaviour. He justifies a primary interest in behaviour because, even though it is only part of the psychology of virtue, an account of virtue necessarily involves action (2002, 17). He cites Aristotle's claim that the activity of virtue is necessary; you cannot be virtuous just by possessing the virtues, but also have to act in accord with them (*Nicomachean Ethics* 1098b30-1099a5). He also highlights that the philosophical literature is concerned with these dispositions being stable or reliable, or what he calls 'robust'. He takes a consequence of this robustness to be a consistency in trait-relevant behaviour over trait-relevant situations.

Doris describes the concept of character under discussion as 'globalism', consisting of three theses. The first is that character traits are consistent, reliably producing trait-related behaviour in a wide range of differing trait-relevant situations. The second is stability, or that the trait-related behaviour is reliably produced over numerous similar trait-relevant situations. The third is evaluative integration, whereby an individual with a trait evaluated in one way is likely to have other traits similarly valued (2002, 22). He claims that the theses are detachable, with philosophy and psychology tending to be committed to the first two, but with less commitment to the third thesis. It is this global conception of character traits that Doris thinks is undermined by the lack of empirical evidence. However, this does not commit him to scepticism about all conceptions of character traits. He refers to many experiments in social psychology that point to a lack of empirical evidence for behavioural consistency indicative of global character traits. He summarises the three central claims of situationism: i) behavioural variation is caused by features of the situation rather than differences of character; ii) attribution of robust traits is problematic because people behave inconsistently with the attribution standards; iii) the evaluative status of

dispositions may vary from situation to situation (2002, 23-25). His conclusion is that the empirical evidence gives him grounds to reject the consistency and evaluative integration theses of global character traits, but leaves him room to defend some version of the stability thesis. He argues for the existence of 'local' traits that are stable over time and situation-particular that cause differences in behaviour (2002, 25).

First, Doris argues that character traits are not consistent i.e. behavioural variation is caused by features of the situation. He concludes from Milgram's experiments that situational factors can lead people to be destructively obedient and that attempts to explain the results based upon differences in character have failed (Milgram 1974 cited by Doris 2002, 39-51).[15] He argues that the experiments raise a question about whether people have robust compassionate character traits and that the experiments provide strong evidence for situationism. He agrees with Harman that the Good Samaritan experiment described in the previous chapter also provides strong evidence that it is the situational feature of being in a hurry that affects behaviour and not the character traits of the individuals (Darley and Batson 1974 cited by Doris 2002, 33-4).

Doris introduces some additional experiments to support his view. The first set of experiments highlight the effect of mood on behaviour (2002, 30-2). In one experiment Alice Isen and Paula Levin set up an experiment in a shopping centre in which some people found a dime in a phone box and others did not (Isen and Levin 1972). They observed whether subsequently an individual was more or less likely to help another individual who had dropped some papers. The results found that those who found a dime helped on 14 out of 15 occasions, whereas those who did not find a dime helped on only 2 out of 26 occasions. Doris argues that based upon these percentages 'finding a dime' is a plausible explanation of the behaviour, showing that a trivial situational factor can have a disproportionate effect on behaviour, undermining the existence of reliable, stable traits. The second set of experiments introduces the effect of being in a group on behaviour (Latané and Darley 1970; Latané and Rodin 1969 cited by Doris 2002, 32-3). In one of these experiments Bibb Latané and Darley pumped smoke into a room where undergraduates were filling in forms (1970, 44-54). After a few minutes there was enough smoke to smell, obscure vision and affect breathing. The results found that when the subject was alone 18 of 24 reported the smoke within four minutes, but when the subject was with two others who remained passive only 1 in 10 reported it. Doris takes this to

---

[15] This experiment was described in Chapter Three.

show that the mild social influence of being in a group can affect behaviour. He explains that these experiments found that being in a group influenced how a subject interpreted the situation; when in a situation where others are not helping, the subjects find the situation ambiguous and are not sure if help is needed. He continues to explain that some of these experiments also found that a lack of clear individual responsibility lead to ambiguity as the subject was not sure whether it was up to him to intervene. He concludes that minor social pressures can affect moral behaviour. The third experiment he uses to support his claims is the Stanford Prison experiment (Doris 2002, 51-3). In this experiment 21 male college students were randomly given the role of prisoner or guard and the experiment was scheduled to last for two weeks (Haney et al. 1973).[16] Prisoners were kept in a simulated prison 24 hours a day and the guards were allowed to administer non-physical punishments, such as force-feeding prisoners, making them clean toilets with their bare hands and putting them in solitary confinement (Haney and Zimbardo 1977). Five of the prisoners had to be released early, due to emotional distress, yet most guards appeared to enjoy their role. The experiment had to be terminated early after only six days due to the extreme behaviour of the guards towards the prisoners. Doris takes this experiment to provide evidence for his view because the participants knew they were participants in a short-term experiment, hence their extreme behaviour must be the result of situational factors and not personality traits. He concludes from all this evidence that character traits are not consistent i.e. trait-related behaviour is not reliably produced over numerous differing situations.

Secondly, Doris argues that character traits are not stable. He argues that the lack of intrapersonal consistency found in social psychology experiments implies that individuals do not act consistently across trait-relevant situations (2002, 63). He argues that the Harshorne and May experiments introduced in the previous chapter provide evidence to support this view as they found a low correlation of similar behaviour over trait-relevant situations (Hartshorne and May 1928 cited by Doris 2002, 63-4). He also suggests that Theodore Newcomb's experiments provide evidence for the lack of stability of traits (2002, 24). This experiment investigated the levels of introversion and extraversion in problem boys (Newcomb 1929). It concluded that behaviour relevant to these traits was specific to particular situations and was observed to be inconsistent rather than in reliable patterns. Doris takes this study to provide further evidence for his argument that character

---

[16] All the subjects were normal or average according to personality measures and there were no significant differences between prisoners and guards (Haney et al. 1973, 81-90).

traits are not stable i.e. that character traits do not reliably produce trait-related behaviour in a wide range of trait-relevant situations.

Doris questions whether what cross-situational consistency we find depends upon how we have defined the situation (2002, 76). The experiments discussed focus on aspects of the situation independent of the subject; for example, group size, finding a dime, hurry. Doris cites social and personality psychologists who think that focusing on 'objective' aspects limits the findings of the experiments, as the behaviour may seem inconsistent by these 'objective' experimental criteria, but seem perfectly consistent to the individual. He is not convinced by their claims as he thinks this is leading to a type of relativity of consistency. He thinks that a situation will always be consistent with some outlook, but that with something as important as morality, the inconsistency found in one outlook is not removed by pointing to consistency under another outlook. He thinks that inconsistency in any context is important to moral traits and moral psychology. This seems to be changing the problem somewhat as it is now focusing on what is the right definition of the situation. Doris claims that we need to know which inconsistencies can be ignored in making moral judgement and argues that any inconsistencies related to moral traits such as honesty and compassion cannot be ignored because our moral standards are socially shared standards about how we should interact with one another (2002, 80). He sums his position up, describing the concerns of morality to depend on shared standards rather than consistency within one individual.

Doris uses evidence for local, situation-related traits found in these experiments to argue that character is fragmented, or 'an evaluatively disintegrated association of situation-specific local traits' (2002, 64). He makes four observations to support his view. First, that the low consistency across situations does not provide evidence for robust traits. Second, that the strength of the influence of situational factors undermines the use of robust traits. Third, that the empirical evidence does not support a link between character and behaviour. Fourth, that biographical stories about individuals reveal fragmentation. He cites the example of Oskar Schindler, who saved over a thousand Jews in Poland, but who was a manipulative womaniser (Keneally 1982 cited by Doris 2002, 59). He continues to argue that local traits would satisfy his standard for attribution; that 'there is a markedly above chance probability that the trait-relevant behaviour will be displayed in the trait-relevant eliciting conditions' (2002, 66). However, the conditions will be very narrow, so these local traits are not robust, i.e. they are not reliable across all situations, just narrow

situations relevant to that behaviour. The problem that Doris faces is that by making his notion of character more in line with the empirical evidence, he decreases the unification of an individual's character (2002, 71).

Doris argues that reflecting on a situationist moral psychology helps us to judge and act better in particular situations than if we were dependent on a moral theory (2002, 109). He says that empirical psychology is not going to tell us what is 'good' but reflection upon an empirically supported moral psychology can help us do right actions and help us to lead morally good lives. He considers whether ethical reflection is best achieved through assuming broad character traits exist, whether this is empirically supported or not. Under such a view character traits may describe an ideal and reflection on this ideal may lead us to moral improvement even if few or no people actually achieve virtuous status (2002, 110-111). He argues that this view cannot be justified on the grounds that moral outlooks that do not make room for character are 'impoverished' because there are many outlooks within both Eastern and Western cultures that do not place emphasis on character, so this claim would need further argument as to why these outlooks were 'impoverished'. Another argument he considers is that ideals of character are stronger than more abstract ideals such as justice and therefore have stronger motivational force. However, he thinks this raises a further question of whether a moral psychology based upon general character traits or a moral psychology based upon the influence of situations is better at producing morally correct action.

Doris argues that accepting a situationist moral psychology should lead us to be cautious about speaking in terms of general evaluative traits such as 'she's a good person' because we would, in that case, think it unlikely that anyone had such general traits (2002, 114-115). He continues to argue that we could still evaluate people as more 'good' than 'bad', as even if an individual's behaviour is not consistent, it can tend more in one direction than the other. In this case, we may use 'good person' as shorthand for 'generally more often behaves well than badly', without committing ourselves to the existence of such general traits. He also thinks that such summaries of past behaviour may serve a role in assessing responsibility, as it may highlight unusual behaviour that can be excused on some ground (2002, 131).

To summarise, Doris's argument takes the following form:

Premise One: Any account of morality should not be based upon an empirically discredited view of character, so moral philosophy should take empirical psychology into account (2002, 2).

Premise Two: Character traits are stable, consistent and evaluatively integrated, reliably producing trait-related behaviour in a wide range of trait-relevant situations and in numerous trait-relevant situations.

Premise Three: Possession of a character trait results in certain behaviour: 'if a person possesses a trait, that person will engage in trait-relevant behaviours in trait-relevant eliciting conditions with markedly above chance probability p' (2002, 19).

Premise Four: Social psychology experiments show that behaviour is very sensitive to the particular situation, with behavioural variation being caused by features of the situation rather than differences of character. The low consistency across situations does not provide evidence for robust traits.

Premise Five: Social psychology experiments show that character does not in any reliable way determine behaviour, making attribution of character traits problematic. The strength of the influence of situational factors undermines the use of robust traits.

Premise Six: Social psychology experiments show that character traits can have a different evaluative status in different situations.

From these premises, Doris draws three conclusions:

Conclusion 1: features of the situation are better predictors of future behaviour than any considerations about character.

Conclusion 2: people do not often possess global character traits, but may display consistent behaviour over specific local situations.

Conclusion 3: reflecting on a situationist moral psychology helps us to judge and act better in particular situations than if we were dependent on a moral theory (2002, 109). Doris thinks that instead of trying to develop characters that are largely independent of situations, we should instead concentrate more on the features of situations that influence our behaviour. He gives the example of infidelity, whereby you avoid a situation where it would be a possibility not because you:

> ...doubt that you sincerely value fidelity; you simply doubt your ability to act in conformity
> with this value once the candles are lit and the wine begins to flow. Relying on character once
> in the situation is a mistake, you agree; the way to achieve the ethically desirable result is to
> recognize that situational pressures may all too easily overwhelm character and avoid the
> dangerous situation (2002, 147).

66

He thinks we would get things right more often by concentrating on the features of situations.

It is unclear as to the extent that Doris's positive argument is in opposition to virtue ethics. Doris's 'nuanced attention to particular circumstance' sounds compatible with McDowell's version of virtue ethics, described in Chapter One. The sensitivity or perceptual capacity that McDowell suggests is central to virtue ethics does not sound dissimilar to the nuanced attention recommended by Doris, illustrating a lack of clarity in Doris's argument as to how features of the situation influence behaviour, unless through being perceived in a certain way. McDowell thinks that this perceptual capacity is essential for an agent to recognise the salient feature of a situation and that being the sort of person who has this perceptual capacity is at the core of virtue ethics (1998, 73). Both McDowell and Doris see attention to the details of particular situations as fundamental, so, on this aspect of virtue ethics at least, there does not seem to be large disagreement. This is more an argument that our practical rationality needs to be developed to ensure that we are aware of the situational factors that may influence our practical rationality, rather than any argument against virtue ethics. The virtue ethicist would agree that situational factors can affect the reasoning and therefore the behaviour of less than fully virtuous agents, but disagree that the best way to avoid the influence of the situational factors is by avoiding the situations, but to be the type of person (fully virtuous) that will perceive the relevant salient features of the situation and act on them; in the experiments involving normal agents, the less than fully virtuous agents are failing to properly exercise practical rationality.

Doris would deny that such character development and unity of the virtues is necessary for an awareness of the influence of features of the situation. However, I think that there has to be an important role for character development. Doris thinks that we would do better by concentrating on managing the situations in which we find ourselves, but surely we already have to be the type of person (of moral/virtuous character) to care about doing better for this to work? The evidence that Doris puts forward suggests that people do not in general have the global character traits that he thinks necessary for virtue ethics. However, neither does this evidence suggest that people in general have sensitivity to the features of situations that may influence their actions. He does not provide an argument as to why it would be morally more successful to cultivate sensitivity to the features of situations rather than to develop global character traits to combat the effect of features of situations on action. Nor does he adequately demarcate what it would be to have such sensitivity from

what it would be to have a virtuous character. Doris's evidence could in fact provide compelling evidence for the development of virtuous traits. The empirical evidence suggests that people in general have neither sensitivity to features of situations nor the virtuous traits required for morally correct action. There is nothing to suggest that such sensitivity can be developed independently of virtuous traits.

There is also a concern as to whether the experiments provide evidence for situationism. It seems clear that this evidence raises questions about whether people have robust character traits, but it is not so clear that the situationist conclusion follows. A particular action may well be a necessary condition for attribution of a character trait, but it is not sufficient for attribution of a character trait. Doris cites many different versions of the obedience experiments (by Milgram and others) and highlights that, despite the differences, an average of two thirds of subjects are fully obedient and close to all subjects are partially obedient. Surely, if our behaviour is so situation sensitive, the variation in experimental conditions should have produced a variation in results? It is unclear how this is evidence for situationism, as it seems to indicate that there is a robust trait of destructive obedience. The experiments do not prove that it is features of the situation that cause people to be destructively obedient; why is this not evidence of a widespread trait of destructive obedience? How will any experiments ever be interpreted as evidence for character traits if, in what appears to be a set of experiments with robust, consistent results, these results are attributed to the situation rather than to character? And why can't the small variation of disobedient people be attributed to their compassionate character trait? Doris is looking for variances in responses to a fixed situation that can be attributed to character traits; the disobedient people seem to fit the bill. He wants to observe a variety of different behaviours, but in this situation the choice is either shock or don't shock, so it is not surprising that there is not a wide variety of behaviour. In fact, interviews and observation of the subjects suggest a great deal of character differences underlying their actions. However, this evidence is taken seriously and deemed to be evidence for situationism by a wide range of social psychologists and philosophers. I will not here consider again the issues with the evidence, as raised in the previous chapter, but will now concentrate upon the implications of accepting Doris's argument that character is fragmented. Assuming that the experiments do provide some evidence for situationism, what implications would fragmented character have for virtue ethics?

2. What implications does accepting the fragmentation theory have for virtue theory?

This section considers three methods of addressing Doris's concerns with the empirical evidence for robust character traits. First, it evaluates Merritt's argument that the evidence is compatible with a virtue ethics that treats virtuous traits as being socially sustained. Secondly, it considers Webber's argument that the evidence provides evidence not for fragmentation of character traits but regularity. Thirdly, it discusses Vranas's claim that the situationist argument results in indeterminacy.

i) Socially dependent character traits

Merritt thinks that it is the evidence for local character traits that provides the largest worry for virtue ethicists (Merritt 2000). She is worried that such local traits are subject to change because they are closely related to our social setting. She argues that this leads to the conclusion that our moral dispositions are at least partly socially sustained. She thinks that this presents a challenge for traditional virtue ethicists, as they tend to argue that the ideal virtuous motives should be independent of social settings. She argues that it is possible to construct a virtue ethical theory based upon socially dependent character traits. She thinks that it is a more sensible project to be careful of your choice of social settings and relationships, rather than trying to make your character independent of the influence of social settings (2000, 378). Obviously there are social settings and relationships about which you have no choice, such as parent to child, but she argues that in these cases we can decide how much importance to give to the relationship. She regards this as a Humean approach, which characterises virtues as stable beneficial qualities. She thinks that this allows for virtue to be socially sustained rather than individual.

There are three problems with this view. First, Merritt characterises Aristotle's virtue ethics as being independent of social settings. This does not seem quite right, as Aristotle was talking about the virtue that can be achieved by a small proportion of Athenian men of a certain social class, a very specific social setting. Secondly, it is not clear that we have a sufficient level of choice we can operate in our social settings and relationships. We are lucky to be well educated, capable of being financially well off and live in a Western culture so, to a certain extent, can operate a degree of choice. This is not so true of, say, someone born into poverty in Ethiopia, as presumably the need for survival overrides any reflection about whether his social setting and relationships are morally beneficial and,

even if they did reflect, they would not have the resources to implement change. Thirdly, this also seems to involve a 'view from nowhere' from which you can evaluate your social settings and relationships. For example if you are indoctrinated into a political or religious regime it would be an exceptional person who would both be able to question and evaluate this from an external perspective and then choose to do otherwise. Merritt does admit that her demands may not be practically easy to fulfil, but does think that they are ethically modest. However, her minimum requirement of 'belonging to a society where conventions of cooperation allow us to live more peacefully and prosperously than we could in their absence' can itself be ethically problematic, as there is a conflict between the general 'us' and what may benefit myself most as an individual (2000, 379).

There is something to the idea of avoiding morally problematic situations (common sense? practical wisdom?), but this is not the appeal of a character-based ethics. The latter's appeal is what it can say about how we should act in these morally problematic situations should we end up in them through no fault of our own. If we cannot exercise choice in our situations and relationships, are we unable to be virtuous? We get into many situations incrementally and, as the Milgram experiment illustrates, incremental progression is hard to deal with. What makes one get out at one point when it is only slightly worse than the previous point? In developing a 'realistic' account of how our characters are socially supported, Merritt is forgetting to ask whether this is how things should be. Doris seems to agree that a socially sustained virtue is not sufficient for virtue ethics. He says that if the character traits are only reliable under certain social circumstances, they cannot recommend how to act in difficult situations, and it is precisely those situations that are relevant to virtue (2002, 90). The difficult project becomes developing a society supportive of virtue, rather than that of individually developing virtue.

ii) Regularity theory of character traits

Webber does not think that we should accept Doris's thesis of fragmentation of character, as he thinks that the evidence suggests this only if read in a behaviouristic way. Further, he thinks a behaviouristic reading is compatible with the more traditional theory of character found in virtue ethics (Webber 2006a). He considers the experiments used by Doris (and others) to argue for fragmentation of character and argues that three of them fail to provide relevant evidence (2006a, 195-198). He rejects the Isen and Levin experiment that suggests that finding a dime has a large influence on subsequent helping behaviour. He rejects it on

the grounds that it only shows that we should not judge a character based upon one instance of helping to pick up some papers because small changes in the situation can affect behaviour in what Webber regards as trivial situations (Isen and Levin 1972 cited by Doris 2002, 30-2). He also discounts the evidence from the Stanford Prison Experiment (Haney et al. 1973 cited by Doris 2002, 51-3). As the experiment has never been repeated and as there was no control group, Webber thinks that Doris's claim that most people would act in the same way in this situation is unwarranted. He also thinks it would be difficult to infer anything from this extreme situation about how individuals would act in more usual situations. The final experiment he rejects is the Hartshorne and May study of honesty in school children, which found that there were very low correlations across stealing, lying and cheating situations in which people could act honestly (Hartshorne and May 1928 cited by Doris 2002, 63-4). He thinks that this experiment can be used only to illustrate how character traits may develop, but say nothing about how character traits influence behaviour in adults.[17] However, as Webber points out, Doris does not rely heavily on this experiment, partly because the experiment involved children and not adults. He thinks that the main reason Doris does not rely on this experiment is that his argument hinges on behaviour being influenced by situational factors that we would not normally consider morally relevant.

Webber argues that the Milgram experiment is central to Doris's argument because it shows that the subjects do not think that the instructions of the experimenter do not make compassionate behaviour inappropriate, yet they still shock the victim, so the expectation of virtue ethics that people will act appropriately is not met. He also thinks that the Good Samaritan experiment and the experiments into the effect of bystanders on helping behaviour could provide evidence for Doris's view (Darley and Batson 1973 cited by Doris 2002, 33-4; bystander effect cited by Doris 2002, 32-3). However, Webber is not clear how these experiments provide evidence for fragmentation of character. He thinks that confusion arises because there is no clear definition of a character trait as used in virtue ethics. This leads to confusion in interpretations of Doris, although Doris himself makes it clear that virtues are conceptually different from character traits and that he is arguing that character traits do not exist in the form that we normally think of them.

---

[17] Sreenivasan, however, thinks that this is the only cross-situational experiment that Doris and Harman refer to that may provide evidence against broad character traits (Sreenivasan 2002). This is discussed above in Chapter Three.

Webber continues to argue that people sympathetic to Doris's view do not make it clear why they think that the data shows that situational features drive an individual's behaviour, not dispositions.[18] He thinks that there is confusion over what is meant by cross-situational consistency. He thinks that Doris is concerned with finding consistency in situations that have the same trait-relevant feature, but differ in other respects (2006b, 652). Doris himself uses higher than chance probability as evidence for traits. Webber reads Doris's interpretation of the experiments as saying that they do not show a diversity of behaviour, so do not provide evidence for traditional character traits (2006a, 203-204). Doris expects to find inconsistency in action in a situation due to people's character traits. Why do we expect people to have widely differing, individuating character (particularly moral) traits?

Webber wonders whether a character trait might be better described as a 'disposition towards certain behavioural inclinations in response to a particular kind of stimulus' (2006a, 204). His definition of a character trait as a disposition to have an inclination towards behaviour, in response to particular features of a situation, contrasts with that of Doris, who defines a character trait as a disposition to behave in a certain way in certain situations. How does having a disposition to have an inclination towards behaviour provide an answer to the challenge from situationism, where having a disposition to behave cannot? If character traits are dispositions to have inclinations, are they compatible with the experimental evidence for situationism and is this definition of character traits compatible with virtue ethics? Webber thinks that the answer to both of these questions is 'yes'.

Webber thinks that a situation will give rise to a range of competing inclinations, not all of which can be acted upon. He calls this the 'regularity' theory of character and thinks that the experiments above provide evidence for his view. The Good Samaritan experiment shows differing inclinations towards helping and punctuality; the bystander experiment shows differing inclinations towards helping and peer-approval; and the Milgram experiments show differing inclinations towards obedience and inflicting pain. A situationist could accept Webber's claim that the Milgram experiments show differing inclinations towards obedience and inflicting pain, but argue that these inclinations are generated by features of the situation and deny that they are the result of their character traits. The experiments do not prove that there are no such things as character traits (whether 'global' or 'regular'), just that if they do exist, they are uncommon and that the

---

[18] He considers Goldie (2004b) and Merritt (2000) to be sympathetic to Doris's view.

situation is far more likely to drive behaviour.

Webber can be interpreted as arguing that the character traits as defined by the regularity thesis are compatible with our normal view of character traits being reliable and stable. This means that we can ask the question of whether the regularity theory satisfies Doris's theses of globalism. It is not clear that divorcing behaviour from the disposition, through the medium of inclination, establishes their globalism. Doris uses the situationist arguments to attack the link between dispositions and reliably produced behaviour found in his first thesis of consistency. Webber criticises the importance of this link between dispositions and behaviour to undermine Doris's argument, but this approach is perhaps unfair. There is a close link between dispositions and behaviour. If the behaviour has never happened, then the person cannot be said to have the disposition. It is true that an instance of helping behaviour can be indicative of being a helpful person but that it could also be indicative of self-interest, for example, wanting the praise for helping. In cases of what looks to an observer like helping behaviour, we need to know about the individual's motives, beliefs, desires, and so on to ascertain whether the helpful behaviour is truly virtuous. Strictly we are incorrect to say whether this person is helpful or not without understanding of his reasons for action, even if the behaviour itself is helpful.

If this instance of helping behaviour does not occur, there is no question of whether the person is helpful; on this evidence, he is not helpful. There is no need to go away and investigate his motives and find that, for example, punctuality 'won' over helpfulness. The fact that he has a defeated helpful reason for action does not count as evidence for him being a helpful person. These motives have to result in behaviour the majority of the time to even count as possible virtuous behaviour, about which we can investigate whether he has the right reasons, motives, etc. The attribution of the disposition has to, in some sense, work backwards from behaviour; if there is no display of the relevant behaviour, then the disposition cannot be attributed. Nor should you attribute the disposition merely on the basis of observed behaviour; the disposition also involves the person having the right thoughts and feelings. Webber seems to go too far in divorcing the intellectual part of virtue from action. He thinks that character traits may create inclinations that are not actually acted upon and that these still count as evidence for possession of the trait. In ethics we have to be interested in action, along with reasons, feelings and beliefs.[19] If every honest inclination I have fails to issue in action, I cannot be said to be an honest person.

---

[19] Doris agrees with this Aristotelian argument (*Nicomachean Ethics* 1098b30–1099a5; Doris 2002, 17).

Doris argues that the fragmentary theory of character traits is compatible with thesis two, stability. Are both the fragmented and regularity theories of character traits consistent with the conclusions of social psychology and resulting local character traits? The character traits under the regularity theory do not meet the criterion of being regularly produced over numerous situations. As mentioned above, the situationist can claim that if there are competing inclinations, they are derived from the situation and not from dispositions. Situationists do not deny the existence of dispositions, just that there is any evidence they regularly cause behaviour. As the regularity theory is operating with situation independent character traits, it cannot answer the situationist challenge. The fragmented theory can answer the situationist challenge because it ties the character trait to a specific situation, so it is the concatenation of the situation and the character trait that has the causal power; the trait alone is inert. Webber would also find it hard to defend any version of the third thesis of globalism, as he describes character traits as competing inclinations, not as evaluatively integrated.

Webber thinks that it is character traits as described by the regularity theory that feature in virtue ethics. He argues that this conception of character traits is compatible with an Aristotelian virtue ethics and that we should prefer his theory to the fragmentation theory (2006a, 205-211). He thinks his conception of character provides a deeper level of explanation and prediction because it explains why a certain situation leads to certain behaviour. This is because it refers to the relative strength of several inclinations of the individual. He thinks that situations will not necessarily link with one trait, but with several in conflict, one of which will 'win'. However, the characterisation of virtue as a disposition to have an inclination towards behaviour is not compatible with an Aristotelian virtue ethics for two reasons. First, this seems to preclude the normative element of virtue ethics. If all we have are conflicting inclinations to differing actions how do we know which is correct? Surely the point of traditional virtue ethics is that the fully virtuous agent acts in the right way, in the right circumstances, for the right reasons and does not engage in a conflict of priorities? Webber's characterisation may be a true picture of the less than fully virtuous agent, but does not seem to capture the fully virtuous agent.

Secondly, Webber's view commits him to the position that having a character trait or disposition means you have an inclination and that inclination causes certain behaviour. This seems to be simplifying a disposition or character trait, reducing it to an inclination that causes behaviour. A character trait, and particularly a virtue, is more complex,

including reasons, beliefs, desires and behaviour, as well as inclinations. Take the example of honesty; it is not merely a disposition to have an inclination to tell the truth that leads to truth-telling behaviour, but it is much more complex, with the individual needing practical wisdom to ascertain that the honest action is the right one in that particular situation and that they are doing it for the right reasons. An example is the case where a lie stops another child getting into trouble mentioned in the previous chapter. Again, Webber's conception of character traits as inclinations seems to describe only the less than fully virtuous agent.

In summary, Webber's regularity theory does not provide a response to the challenge from situationism because it is based upon criticism of a link between dispositions and behaviour. The situationist can claim that if there are competing inclinations, they are derived from the situation and not from dispositions. The regularity theory also does not seem to be more compatible with a traditional view of virtue ethics than the fragmentary theory of character. Inclinations that are not acted upon do not count as evidence for possession of a trait and a fully virtuous agent does not have conflicting inclinations. Further, a traditional virtual ethical view does not argue that a virtuous person is motivated by inclination.

iii) Does accepting the fragmentation theory result in indeterminacy?

If we accept the fragmentation of character what are the implications? Vranas argues that a person who regularly behaves well in some situations and also regularly behaves badly in other situations is 'indeterminate' (Vranas 2005). Doris's view is that we can evaluate people as more 'good' than 'bad' because even if an individual's behaviour is not consistent, it can tend more in one direction than the other. Vranas criticises this view on the grounds that a person of fragmented character is instead indeterminate. He thinks that because most people have a fragmented character, this results in a paradox. He states this indeterminacy paradox as: 1) most people have a fragmented character and 2) fragmentation entails indeterminacy, therefore most people are indeterminate. He defines 'indeterminate' as a person who is not good, bad or intermediate. He further defines indeterminacy as the property of lacking character. His claim that a person who regularly behaves well in some situations and also regularly behaves badly in other situations is indeterminate therefore states that such a person is not good, bad or intermediate (2005, 17).

Vranas uses an analogy with temperature to argue that a person who is neither good nor bad is indeterminate rather than intermediate (2005, 19). He argues that being between hot and cold is to have an intermediate temperature, but that a lake that consists of discrete hot and cold areas cannot be said to have an intermediate temperature; he thinks that is has no overall temperature. He considers another analogy; if a person believes arguments for and arguments against a position, but does not believe that one side outweighs the other, then he does not hold a position between for and against, as he does not have a position to fit on the for/against scale. He argues that if a person was intermediate then they would have to be better than every bad person. He states his argument as follows: a bad person (*b*) acts badly in every situation that a fragmented person (*f*) behaves well and acts neutrally in every other situation. This means that *f* would behave better than *b* in some situations (those where *b* acts badly, but *f* well), but also that *f* behaves worse than *b* in some situations (those where *f* behaves badly and *b* behaves neutrally), so *f* cannot be said to be better than *b* (2005, 20).

It is not clear that Vranas's analogies establish a paradox. Taking the lake analogy, we can still assess whether there are more hot patches than cold and an average temperature would give us an indication of this even if no part of the lake actually had that temperature. The analogy with the person who believes arguments both for and against a position can be evaluated on the for/against scale, depending on how many for or against arguments he agrees with and how strongly. We quite commonly say we are more for a position than against on such a basis. Similarly, in his comparison of the bad person (*b*) with the fragmented person (*f*); it is not that one instance of *f* acting worse than *b* means that *f* cannot be said to be better than *b*, but looking at the overall balance of whether, in general, *f* is better than *b*. We do not have to accept the conclusion that people who are neither good nor bad are indeterminate, or lack character. Doris is correct to maintain that people who are neither good nor bad are intermediate; overall character assessment is a matter of degree. As they stand, Vranas's objections are not enough to refute Doris's position.

3. Could there be an evaluative connection between fragmented character traits?

Doris thinks that an idealistic virtue ethics will be subject to the same alienation worries as more 'theoretical' approaches to morality: 'worries about theoretical mediation may recur, if ethical practice consists in regulating behaviour by reference to an ideal of virtue, instead of simply acting from virtuous dispositions' (1998, 520). He considers an alternative to

copying the behaviour of moral exemplars; we may ask what advice the virtuous agent would give, based upon the situation and how it may affect the less than fully virtuous agent i.e. the advice is not to do what a fully virtuous person would do, but what this particular non-fully virtuous person should do (1998, 518). Doris questions whether this is really a virtue ethical theory, as he thinks it more important to this view that we try to emulate the fully virtuous, not just ask their advice. He continues to suggest that the two methods complement each other; in some cases emulating the virtuous agent will be the best thing to do, in others we may need to do something else, as suggested by the advice model. He continues to argue that both the emulation and advice models depend upon analysis of the situation, thus situationism is normatively useful in deliberation, as well as descriptively correct (1998, 519). The virtue ethicist would not deny the importance of the situation, but he would deny that it is the situation that provides the normative force; what provides the normative force is the value of being a certain sort of person.

Under this model, Doris thinks that virtue ethics would no longer be able to claim an emphasis on character development rather than reflection on an ideal. It could no longer allow our normal use of character in explaining action and it no longer avoids problems with theoretical deliberation. Miller does not think that these conclusions obviously flow from Doris's arguments (Miller 2003). His arguments, in Miller's view, admit of the possibility of more modest robust character traits, so it could still be an aim to develop such traits and use them in explanations, even if only a few people were fully virtuous. He thinks the objection holds only against a theory that claims everyone is fully virtuous, as otherwise the problem of an individual being uncertain what to do in a new situation is one that faces every moral theory and is not a new problem that arises from Doris's arguments. He argues that becoming fully virtuous involves a struggle in overcoming character defects and situational distractions, which indicates that most of those in the Milgram experiments were not fully virtuous. He claims that the best conclusion that can be drawn is that people fall short of being fully virtuous and that this is not a problem for virtue ethicists because full virtue is hard to obtain (2003, 378).

Miller continues to develop a positive account that he thinks allows for global trait acquisition and is compatible with the evidence from social psychology (2003, 382). He thinks we should construct an account based upon local character traits, for which there is evidence of wide possession and is compatible with Doris's arguments. An individual who possesses a local trait will act consistently and his action can be predicted across a narrow

type of situation. He thinks that for prediction of what he may do in some other type of situation, we would need to know whether he had other local character traits that may suggest what he would do in that situation. He thinks that an individual needs a range of complementary local traits to appear consistent across different situations.

Miller argues for the existence of a 'personality network' of cognitive, affective and motivational states that produce local trait-specific action in narrow situations, based upon the Social-Cognitive theory of personality (2003, 383-384).[20] He thinks that this personality network will determine what features an individual sees as relevant in a given situation, hence why two people in the same situation may act differently, or why an individual may act differently in two similar situations. He thinks such an account of local character traits is supported by the social psychology literature. He continues to consider how such local character traits may be linked to the more global concepts used in virtue ethics. He thinks that this hinges upon an account of moral education, explaining how we develop character traits. He holds that an individual that is influenced by something insignificant, such as not finding a dime, has not received adequate moral education. An adequate moral education means we will habitually recognise and respond to the morally salient features of the situation. He argues that there is empirical evidence for this kind of habitual personality training (2003, 385).

However, Miller needs to be careful with the statement of this view, as we do not act morally out of mere habit, but for the right reasons. He is in danger of losing the intellectual element of the virtues. He is trying to argue that we need to train ourselves to habitually respond to features of situations and that this personality response then triggers reasons, goals and actions relevant to the situation. Perhaps it is more the recognition of the morally salient features that needs to be habitual, but what actions are generated will still be subject to reason. He thinks that our local character traits will only be grouped together

---

[20] The Social-Cognitive theory of personality argues that stable variability in behaviour between individuals can be explained by an individual's 'cognitive-affective personality system', which contains his beliefs, feelings, goals, competencies and strategies (Shoda et al 1994, 682-3; Mischel and Shoda 1995, 254; Mischel 1999, 677-80; cited by Doris 2002, 77). They argue that there is something about the individual's character that causes consistently variable behaviour. Doris cites Yuichi Shoda et al's example of a child that displays slightly below average verbal aggression when teased by a peer, displays substantially above average verbal aggression when praised by an adult, greatly above average when warned by an adult and somewhat above average when punished by an adult (Shoda et al 1994, 677-80, cited by Doris 2002, 77). They think that although this child's behaviour is not consistent across different situations, if the child was put in the same sequence of situations, he would demonstrate the same variability of behaviour. Walter Mischel and Shoda argue that this is the type of psychological consistency referred to by personality, not behavioural consistency over nominally similar situations; hence they also disagree with Doris's definition of a character trait (1995, 257).

as more general concepts, such as honesty, after lengthy habituation. He argues that for a minimal possession of a character trait we must act in the appropriate way in paradigm situations. Goldie also argues that we should harmonise these fragments into something akin to traditional character traits (Goldie 2004b). To demonstrate such a view, Webber gives an example of sociability, suggesting that if an individual is sociable in specific circumstances and values sociability, then using circumspection and strength of will, he will manipulate his surroundings so that he is sociable more often, thus making his original situation-specific sociability a wider disposition (2006a, 194). Kamtekar also makes a similar argument, suggesting that what starts off as a narrow disposition that can be explained as a strategy developed to deal with a certain situation could come to inform our actions in other situations and become broader-based (Kamtekar 2004, 469).

Doris rejects the Social-Cognitivist picture because he argues that although people may be individually coherent, this does not mean that they behave consistently with shared ethical standards. He thinks that virtue ethicists need consistency across nominal features of situations, not just psychological features of situations, so any inconsistency over nominal features of situations (i.e. shared ethical standards) will be problematic, regardless of whether the individual is consistent over psychological features of the situations. But what if the individual coherence is consistent with the ethical standard; this is surely the aim of virtue ethics, for the character structure of individuals to cohere with that of the ideally virtuous agent? If a character-based ethics can be based upon psychological features of situations, then it appears consistent with social psychology. The Social-Cognitivists argue that people can individually behave in a consistent and stable way, and this possibility seems to be exactly what the virtue ethicist needs. It is a further question then whether the individuals are acting consistently with shared ethical standards. The individual coherence explains why people act the way they do, but the shared ethical standards determine whether this behaviour is justified. It is not the convergence on the shared ethical standards that explains an agent's behaviour, but his personality structure, including character traits as well as beliefs, emotions and other commitments of the agent. For example, subjects in the Milgram experiments may well consistently obey authority figures in experiments but not obey the authority of the government in compiling their tax return, with this variability in obedient behaviour explained by their personality structure. Whether their behaviour is justified in any particular situation depends upon the shared ethical standards; if their behaviour does not conform to the shared ethical standards, this does not show that they do

not have robust personality structures, including character traits, but that on this instance, although their behaviour can be explained, it is not justified.

This is where the situationist argument goes wrong. Doris admits that the situationist argument does not show that robust character traits are impossible and that it is possible that there are a small percentage of people who are fully virtuous (1998, 511). He thinks that an Aristotelian account can only be committed to the more modest claim that there are only a few individuals who act as exemplars and the rest of us guide our conduct by reference to them. However, this is not the only interpretation. A coherent theory may only need there to be a few individuals who are fully virtuous, but only so that striving for full virtue is a worthwhile enterprise as it may actually work; we do not need these exemplars so that we can decide what to do by reference to them. This is not what virtue ethics is limited to if fragmentation of character is accepted. As Goldie suggests, we learn to value certain local traits and try to use these virtues in other situations, thus acting from our local virtuous dispositions in working towards obtaining a virtuous ideal.[21] He suggests that we discover local traits we have that are valued by ourselves and others and then try and apply these traits to other situations. This is an idealistic view that is based upon our own values, not deliberation about what the ideally virtuous would do, thus avoiding concerns about alienation:

> If Susan identifies with her character trait of kindness, considering it a virtue, she will be idealistic about it: she will consider that she ought to be robustly – stably and consistently – kind, even if she isn't fully virtuous and is occasionally tempted to act selfishly, and not to do what a kind person ought to do...her dispositionism about herself is a form of idealism – being robustly kind is what she thinks she ought to be, where this 'ought' is idealistic and moral as well as predictive (Goldie 2004b,70).

Conclusion

By identifying character traits with specific situations Doris seems to be denying us the ability to make any evaluative connection between the fragments. Without a story to be told about how this may be done, the notion of a local character trait does little more than reiterate the point that we need to take care over attributing general character traits as it is highly likely that we do not have enough evidence to make such an attribution and unlikely that the person in question actually has such a robust character trait. This chapter has

---

[21] Doris also suggests that the virtue ethicist may have an answer along the lines of the fact that an action is virtuous is not part of the deliberative reasons for doing that action (Williams 1985, 10). Goldie also agrees that an action's being virtuous is not necessarily one of your reasons for action (2004b, 15).

introduced a promising suggestion that an individual's character is a network of integrated traits. This idea will be further developed and explored in later chapters.

However, there remains the problem that there is a lack of clarity as to what a character trait is; whether a general trait that causes consistent behaviour across all situations, a more specific trait that causes specific behaviour in specific situations or a disposition to reason in a certain way. There are two problems with the notion of situation specific local character traits that could provide the foundation blocks of this personality network. First, the next chapter will argue that there is an inconsistency between the situationist claims and the reasons an agent has for action. Secondly, Chapter Six will argue that there is a problem with the specification of the conditions of locality.

Introduction

The aim of this chapter is to raise a new problem for Doris's position outlined in the previous chapter. This problem is based upon the relationship between motivating reasons for action and features of the situation. I argue that it is possible to interpret Doris's argument to imply that motivating reasons for action are features of the situation. However, I argue that this would then commit him to a controversial theory of action, such as that proposed by Dancy. He argues that reasons for action are indeed features of situations, but his argument has two unwelcome implications for Doris; that what explains an action need not obtain and that action explanations in terms of reasons are not causal explanations. I argue that Doris cannot accept these implications.

I argue that the situationist can challenge Dancy's argument by questioning his notion of the 'normative constraint', a central assumption of his argument. This states that all motivating reasons should be the 'right sort of thing' to be normative reasons, and hence that motivating and normative reasons cannot be different types of reason. I argue that this conclusion rests upon a mistaken interpretation of Williams's claim that if an agent acts for a reason, then his reason must feature in a correct explanation of his action. I offer a different interpretation of this principle that allows the claim that normative reasons are facts to be compatible with the claim that motivating reasons are psychological states. I reach a different conclusion from that of Dancy and argue that we can consistently hold that motivating and normative reasons can be different types of reason, so his argument that reasons for action are features of the situation does not follow. However, I argue that this does not resolve the problems for the situationist because the rejection of Dancy's position involves commitment to the claim that a motivating reason is constituted by a psychological state, so it is psychological states of the agent that are of central importance in explaining why an agent acted in a certain way and not features of the situation.

1. Situationists and motivating reasons

Doris's position on motivating reasons is not entirely clear. However, he is committed to consideration of the agent's motivating reasons for action because he agrees that virtue as a behavioural disposition is not sufficient, as the concept of virtue involves emotional

response, deliberation and decision making, as well as behaviour (2002, 17).[22] The virtuous person's life and actions are reasoned ones. If virtues are dispositions to act according to reason, then someone who only occasionally acts virtuously is indicating that his moral reasoning is poor, so consideration of the agent's reasons for action has to play a part in Doris's account. Motivating reasons report why an agent did something and as Doris is interested in explanations of behaviour he must have an interest in motivating reasons.[23] The problem with his argument is that it does not appear to draw a distinction between intentional action and mere behaviour. If intentional action is taken to be action for which there is a reason, it is difficult to see how intentional action fits into the situationist position; if it is the feature of the situation that causes behaviour, there seems to be little space for the agent's intentions and motivating reasons for action. Doris would argue that people are regularly wrong about the motivating reasons for their intentional actions (2002, 139). On the one hand, it seems that I cannot be wrong about the motivating reason for my action; if I say what motivates my action is my wish to help, that is the explanation of my action, whether or not a third person thinks it is because I found a dime. On the other hand, it does seem that I can be wrong about what motivates my action; at one time what explains my action is my wish to help, but at a later time, after reflection, I realise what really explained my action was selfishness. The predicates are temporally fixed. I can reflect on what I think motivates my action and at a later time find this explanation erroneous. However, this is still first personal perception and reflection. The third personal explanation seems different from first personal explanation; the third person can only get at the explanation of intentional action by quizzing the agent as to his reasons. Many of the experiments Doris cites do not involve intentional actions; for example, failing to help in the dime and Good Samaritan experiments.[24] Unintentional actions do not appear to be evidence for a lack of virtue because virtuous action has to be undertaken for the right reason.[25]

---

[22] It is unclear whether Doris can claim to have argued for a normative moral theory. Normative reasons state what an agent should do. His positive argument that instead of trying to develop characters that are largely independent of situations, we should instead concentrate more on the features of situations that influence our behaviour, has little to say about normative reasons or what an agent should do. Hence the focus of this chapter will be on motivating reasons (or the motivating use of reasons in the case of Dancy who denies a distinction between types of reason).

[23] Motivating reasons are sometimes called explanatory reasons.

[24] Of course some of those who did not help may have done so intentionally, but some simply did not notice that help was needed.

[25] We can criticise people for not having been aware of an individual needing help, but this is not to say that they would not have been disposed to help had they noticed.

Does Doris think that features of the situation explain behaviour or the agent's psychological states? On the face of it, he appears to argue that it is features of the situation that explain behaviour. He admits that 'situationism tells a disconcerting story about the way some behaviours are caused' (2002, 132). This interpretation is backed up by his discussion of responsibility. He argues that '...condemnation for ethical failure might very often be directed not at a particular failure of the will but at a certain culpable naivety or insufficiently careful attention to situations' (2002, 148). Here the emphasis is on awareness of the situation rather than the will of the agent. He also argues that the possibility of excusing conditions (e.g. coercion) mean that identifying psychological states is insufficient: '...responsibility assessment consists in establishing the presence of causal and psychological connections and the absence of excusing and exempting conditions' (2002, 130). He makes a clear distinction here between what causes an action and the psychological states of the agent.

Doris refers to several explanations of behaviour; for example, in the Good Samaritan experiment, the behaviour of the subjects is explained by the degree of 'hurry', a feature of the situation. Whether the subject was in a situation of 'low hurry', 'medium hurry' or 'high hurry' is used to explain whether or not the subject stops to help the 'victim'. Ross and Nisbett suggest that it is situational factors that give the explanations for action in Milgram's experiments on obedience.[26] In the Milgram case, these are the stepped nature of the experiment, the difficulty of stopping and the non-sensical situation. Do these explanations constitute a motivating reason for the person to shock the actor? The three explanations suggested by Ross and Nisbett could be detailed as:

1) The agent gave a 450 Volt shock because in the experiment the shocks are increased by small increments so it is difficult for the agent to justify to himself stopping at one point rather than another.

2) The agent gave a 450 Volt shock because the experimenter encourages the agent to continue when he objects. It is only on the fourth consecutive objection that the experimenter lets the agent stop, making it difficult for the agent to terminate the experiment.

3) The agent gave a 450 Volt shock because the situation no longer made any sense; he was trying to teach someone who was no longer responding.

---

[26] Ross and Nisbett (1991, 56-58), quoted by Harman (1999, 322-23) referring to Milgram (1974).

These explanations all refer to features of the situation to explain the agent's action. This suggests that the motivating reason that explains why the agent gave a 450 Volt shock is the feature of the situation and not any desire or belief of the agent.[27] Doris thinks that we should explain actions by reference to features of the situation and that we should avoid the error of explaining the actions of others in terms of their character. I here suggest that it is possible to interpret his position as compatible with the thesis that motivating reasons are features of the situation.

Doris's elaboration on what he means by a situation also suggests that it is features of situations that are important in motivating an agent's action. He refers to the distinction between nominal and psychological features of situations; the former are features of the situation that are independent of the agent and the latter are features of the situation that are salient to a particular individual at a particular time (2002, 76). He states that the features of situations that he has been considering, such as finding a dime and the number of people present, are nominal features of the situation and it is these features that he is taking to explain action. He thinks that virtue ethicists need consistency across nominal features of situations, not just psychological features of situations. He thinks that the virtue ethicist is committed to the claim that 'if a person possesses a trait, that person will engage in trait-relevant behaviours in trait-relevant eliciting conditions with markedly above chance probability p', so any inconsistency over nominal features of situations (i.e. shared ethical standards) will be problematic, regardless of whether the individual is consistent over psychological features of the situations. This makes it clear that he does think that it is features of the situation that explain behaviour, and not the individual's perception of the situation, which in the next section I will argue seems to leave him with a rather problematic theory of motivating reasons and explanation.

In summary, the position outlined above implies that features of the situation provide motivating reasons for action. For example, in the Good Samaritan experiments, the behaviour of the subjects is explained by the degree of hurry, a feature of the situation. In this case, the motivating reason for the agent is that he is in a hurry, so has no time to stop and help. This motivating reason for action makes no explicit reference to any beliefs or desires of the agent.

---

[27] I shall return later to the question of whether explicit reference has to be made to the desires or beliefs of the agent for them to be part of the explanation of an action.

2. Support for the claim that motivating reasons are features of situations

Dancy constructs an argument to support the view that reasons for action are features of the situation rather than psychological states of the agent (Dancy 2000b). This section will outline Dancy's position. The subsequent section will consider whether such a theory of reasons for action is compatible with situationism. His primary target is the view that 'the proper, philosophically revealing form (of a reason) will be "A acted because A believed that p"' (2000b, 99). He starts by defining what type of action and reason he is interested in. He states that 'intentional action is always done for a reason' (2000b, 1). He claims that this is thinking of the reasons as motivating and continues to argue that when we consider the *good* reasons for doing the action, we think of reasons in the normative sense. He does not think that there are two types of reasons, explanatory and normative, just two ways of viewing the same reason (2000b, 2). He says: 'in the best case, there is some good reason for doing the action, and the reason that motivates the agent coincides with that reason' (2000b, 3). I am here interested in the motivating use of what Dancy calls a 'reason'. He denies that either desires or beliefs are what motivate an agent and further that the fact that an agent believes something is not what motivates the agent. From these premises he draws his conclusion that when we act for a reason, our actions are not explained by psychological states. He argues that reasons for action are determined by how the world is, not how we see the world and that a reason is both normative and explanatory. However, this argument has two implications; that what explains an action need not obtain and that action explanations in terms of reasons are not causal explanations.

First, Dancy eliminates the possibility that it is desire that motivates an agent to act. He argues that desire just is the state of being motivated, rather than a part of what motivates an agent (2000b, 85). He does not think that '...because I want to' is an adequate explanation, hence some desires cannot be explained and the actions motivated by these desires cannot be explained either: 'If we cannot say why we want to do it, the fact that we want to do it offers nothing by way of explanation for the action' (2000b, 85). He thinks that such actions are 'incomprehensible'. As he identifies desire with motivation, desires (motivations) cannot explain why I was motivated; only what underlies the desire (motivation) can do this (2000b, 86). He identifies a potential problem with his position; that it appears to cause problems with desires such as those to do with the past or other desires about things over which we have no control, as it is hard to see how, in such circumstances, we can be said to be motivated (2000b, 87). He introduces a conditional

stating that: 'A is motivated to φ iff, were an opportunity of φ-ing *per impossibile* to arise, A would seize it, in the absence of contrary motivation' (2000b, 88). He argues that this conditional eliminates these problems, because it allows that an agent can be described as motivated if they would act in such a way if possible.[28]

Secondly, Dancy argues that it is not psychological states in the form of beliefs that motivate agents to act: 'it is not our believing that things are so that motivate us...but rather what we believe' (2000b, 77). Dancy thinks that there are three options for understanding reasons:

1) 'We can understand both normative and motivating reasons as psychological states of the agent' (2000b, 99).

2) 'We can understand all reasons as what the agents believe, rather than as their believings of those things' (2000b, 99).

3) '...we can hold that motivating reasons are psychological states of the agent, while normative reasons are what agents (we hope) believe' (2000b, 100).

He objects to option one, the position that both normative and motivating reasons are psychological states on the grounds that there is a difference between 'A's reason for φ-ing was that A believed that p' and 'A's reason for φ-ing was A's believing that p'. He points out that the latter is a psychological state, the former a fact, or state of affairs, so we cannot argue that only psychological states explain actions, as here is an example where a state of affairs explains the action (2000b, 102).

---

[28] Christian Piller thinks that there is a problem with Dancy's argument that desires are not what motivates an agent (2003, 419-20). He thinks that even if we grant that to desire is to be motivated. it does not establish the conclusion that Dancy needs. He gives the example that we can agree that my desire to exercise cannot explain why I exercise, but that the desire to exercise can explain why I put on my running shoes (2003, 420). We here have an action that is explained in terms of desire rather than reason. I desire to exercise, perhaps because I see some good in this end, so have a reason not based on desire. But I do not see some good in the end of putting on my running shoes, but I have reason to do this because of my desire to exercise.

Piller also thinks that there is a problem with accepting that to desire is to be motivated, as we can desire many things that we can't bring about. Piller considers Dancy's conditional reformulation that desire is to be potentially motivated, but argues that this is problematic because then there would be nothing we didn't desire i.e. this would result in A desiring to do anything you can think of (2003, 420). He gives the example of desiring to beat Tiger Woods at golf in next year's US Open. He says he has never played golf, so it would be absurd to say that he has this motivation. However, if he did find himself at the US Open next year and he was 10 strokes ahead of Tiger Woods on the final green, then he would be motivated to win. This example demonstrates that it is difficult to see how desire is being potentially motivated, because it would appear that we can be described as potentially motivated to do almost anything.

He also rejects option three, the position that motivating reasons are psychological states and normative reasons states of affairs. He thinks that if normative and motivating reasons were different things, the normative would explain why we have certain motivating reasons, and the motivating reasons explain our actions (2000b, 101). He continues to argue that all motivating reasons should be the 'right sort of thing' to be normative reasons, and hence that motivating and normative reasons cannot be different things (2000b, 103). He calls this the 'normative constraint'. Rejection of a distinction between the two types of reason leads Dancy to argue that all reasons are *what* the agents believe, rather than their believings (2000b, 99). He argues that 'no...psychological states of the agent are normative reasons; it is not normally psychological states of the agent that make his action the right one to do' (2000b, 104). He thinks this because 'the crucial point here is that believing that p is never... a good reason for φ-ing. It is what is believed, that p, that is the good reason for φ-ing, if there is one' (2000b, 107).[29]

Dancy draws upon the work of Arthur Collins to support this argument (Collins 1997). Collins uses 'he is going to the ferry because the bridge is closed' as an example of a reason for acting (1997, 108). He thinks that the motivating factor is that the bridge is closed and that we are incorrect to take the above reason to really mean that 'I am driving to the ferry because I believe that the bridge is closed' (1997, 109). He does not think that reasons are psychological, but 'objective circumstances'. He argues that 'claims about objective circumstances cannot be deleted without dropping the explanation all together' (1997, 109). He believes that a correct reason-giving explanation must appeal to the objective circumstance, not a belief that the objective circumstance obtains.

Dancy is particularly interested in Collins's objection that recasting reasons in terms of beliefs would make it possible for the agent to explain his action with no commitment to the truth of his belief (2000b, 109). Collins's central argument is that we give the same explanation of the action from both the first and third personal perspectives. He argues that from the third personal perspective we tend to recast 'Joe is going to the ferry because the bridge is closed' as meaning 'Joe is going to the ferry because Joe believes that the bridge is closed' (1997, 112). He thinks that the third person, Jack, is neutral about the state of the bridge and only refers to states of Joe. However, he argues, Joe is not neutral about the bridge; he believes it is closed and his explanation necessarily involves this non-neutrality. This is what he thinks makes the explanations appear different. He continues to argue that

---

[29] I shall consider this part of Dancy's argument in more detail later in the chapter.

when Joe restates the claim '...because the bridge is closed' to '...because I believe that the bridge is closed' he does not drop the claim about the bridge and substitute another claim (as the third person does). He says that if he did drop the claim about the bridge, then he also has to drop his belief. So if Joe restates the claim, he is not changing the subject matter; he is still talking about the objective circumstance, whereas Jack is talking about a state of Joe and can drop the claim about the objective circumstance (1997, 113-4).

Collins considers the problem raised by truth values: 'but everyone knows that 'I believe the bridge is closed' can be true even though 'The bridge is closed' is false' (1997, 114). He considers the subject's beliefs to be stored representations of how the world is. He argues that the subject matter when the agent endorses a representation is different from that when he creates a representation:

> ...we can say within this hypothetical understanding of belief, that Joe's 'I believe that the bridge is closed', differs from the reply 'The bridge is closed', in that it adverts to the stored representation rather than creating a new representation. Now Joe's statement will be false if there is no such representation in his brain. And it will be true in case there is such a stored representation even if the bridge is open (1997, 116).

He continues to argue that the truth of 'I believe that p' depends upon the existence of a representation in the agent's brain, so in this case, the agent does not endorse p any more than the third personal 'Joe believes that p'.

His two central points, endorsed by Dancy, are:

1) 'The shift from 'Because p' to 'Because I believe that p' does not delete the speaker's endorsement of p...If the report of a representation is separated from the endorsement it no longer expresses belief' (1997,118)

2) '...the explanation proffered by 'I am doing it because I believe that p' absolutely depends on the fact that these words do express the speaker's commitment to the truth of p' (1997, 118). The correct explanation refers to the 'objective circumstance', not the psychological states of the speaker.

Dancy considers whether we can make sense of explanation by making appeal to content; that it is the 'psychological state *plus* content that together constitute the motivating reason, and the content alone that constitutes the normative reason' (2000b, 113). He does not think that this is a plausible argument because of his commitment to the normative constraint, that motivating states must be the 'right sort of thing' to be normative reasons.

In this position, the motivating reason is different from the normative reason, so he does not think that the motivating reason can be a good one.

Dancy also considers whether we can make sense of explanation by making appeal to content as propositions; i.e. whether the content of our beliefs can be good reasons for action (2000b, 114). He argues that states of affairs are our good reasons, not propositions: 'it is her being ill that gives me reason to send for the doctor, and this is a state of affairs, something that is part of the world, not a proposition' (2000b, 114). He thinks propositions cannot be good reasons because they are abstract objects (2000b, 115). He thinks it would be an unusual circumstance in which it is 'the truth of the proposition that p that makes the action right, rather than more simply that p' (2000b, 116). He thinks that for such an argument to work, 'it must somehow understand contents as states of affairs...But this will require significant revisions in current views within the philosophy of mind' (2000b, 117).

Dancy's final premise is that the fact that an agent believes something is not what motivates the agent. Such a position would formulate the correct explanation as 'that he believed that p' (2000b, 121). This is different from believing being what motivates the agent. He thinks 'that he believed that p' is a state of affairs that can be the case or not be the case and can itself be believed. However, he rejects this possibility on the ground that 'if there is a significant difference between the explanation 'that he believed that p' and the simpler 'that p', the advantage is normally all on the side of the simpler version' (2000b, 125). He also thinks that 'the situations of which it is most obviously true are very uncommon ones..., so that the general thesis must be false as a general thesis just because of the peculiar nature of the cases which it correctly characterizes' (2000b, 125).

Overall, Dancy's argument has two implications. First, that not all explanations are factive. He raises the question of how an agent's action can be explained when what he believes is not the case (2000b, 146). He introduces the notion of non-factive explanations: 'There must be some form of words with which we can give the agent's explanation without committing ourselves to things being as the agent supposed' (2000b, 146-7). He thinks that we make a mistake when we think that all explanations are factive i.e. what explains an action must be the case. Because of this mistake, he says, when something is not the case, we seek something else that is the case, such as a belief, to explain the action. He thinks that we should be happy to allow that an agent's reason for action is something that is not the case. However, he struggles to define such an explanation: 'this does not tell us what

sort of thing a what-is-believed is when it is not the case – where to place such a 'thing' metaphysically. Perhaps the only answer is that it is something that may or may not be the case' (2000b, 147).

Secondly, it follows from this conclusion that not all explanations are factive and that actions can be explained by events that did not occur, that reason explanations are not causal explanations. Dancy thinks that it is possible to defend the view that reasons are not causal (2000a, 47). The common view is that of Donald Davidson; that reasons must be causes (2001, 9). Davidson argues that an action may have more than one possible explanation and that the reasons for which the agent acted can be identified as those explanations that caused the action. He thinks that 'central to the relation between a reason and an action it explains is the idea that the agent performed the action *because* he had the reason' (2001, 9). Dancy does not think that there is any advantage to thinking of reasons as causes (2000a, 47-8). He does not think that any theory of causation can 'distinguish an active from an inactive but potential cause', so does not have the advantage of being able to analyse the relation between reason and action (2000a, 47-8). He also does not think that it is an advantage of the causal view that the truth that something was a reason for action depends upon the truth that it caused the action: 'it is no help here to say that a causal story can always be unpacked, or broken down into micro-elements' (2000a, 48).

3. Problems for the situationist position

There are two problems with the argument that reasons for action are features of situations that mean that Doris cannot endorse this position. First, the situationist needs reason explanations to be factive as he cannot commit to the position that an objective circumstance that does not obtain can explain action. I think that the greatest challenge for Collins and Dancy is to answer the question of how Joe's action can be explained by the 'fact' that the bridge was closed, if the bridge was, in fact, open. Christian Piller does not think that introduction of the concept of non-factive explanations makes the problem of false beliefs with no content disappear (2003, 424). He gives the example of a crowd prematurely celebrating a winning goal that was not scored. For Dancy, it is not their false belief that explains their actions but what they believed that explains their actions: 'If the crowd celebrated prematurely, Dancy would introduce a third category: the games you lost that were also wins, just wins that didn't happen' (2003, 425). Piller feels that this makes no sense. I agree because Dancy can't say what the sort of thing that explains an action

91

based upon a false belief may be if it is not a psychological state of the agent. Without such an explanation the position appears unsatisfactory. Throughout his argument, Dancy makes appeal to the simplest explanation being the best and it seems clear that here the simplest explanation is that the action is explained by a false belief, not by a metaphysically suspect event that did not happen.

I agree with Piller that the concept of a non-factive explanation makes no sense. I think that this implication of Dancy's argument means that the situationist cannot be committed to the position that features of the situation provide motivating reasons for action because it is not always what is believed that the situationist uses to explain action. The situationist implicitly makes appeal to false beliefs to explain actions, not events that did not happen. For example, in the Milgram experiment it is not the metaphysically suspect event of an experiment-investigating-the-effects-of-punishment-on-learning-that-did-not-happen   that explains why the agent acts in a certain way, it is the agent's *belief* that he is participating in such an experiment that explains how he acts. For example, the motivating reason the agent gave a 450 Volt shock was because the situation no longer made any sense; he was trying to teach someone who was no longer responding. This motivating reason for his action depends upon his belief that he is involved in the relevant experiment. Similarly, in the Good Samaritan experiment it is the agent's belief that he is in a hurry that explains why he rushes past the person in need of help, not the event of being in a hurry that didn't happen. The fact that the agent believes he is in a specific experiment or in a hurry is important to the situationist explanation of how the agent acts.

Secondly, the situationist needs reason explanations to be causal. As established earlier, Doris's argument implies throughout that features of the situation cause action. However, it is a consequence of the type of position that Dancy argues for that action explanations in terms of reasons are not causal explanations. Hence this is another reason I do not think that the situationist can accept this theory of action explanation, because they require that explanations are causal. The motivating reason(s) the agent had for shocking the learner, such as that the situation was non-sensical, *caused* him to shock the learner. The motivating reason the agent had for rushing past the person in need of help, that he was in a hurry, *caused* him to rush past.

Can the situationist avoid the conclusions of Dancy's position, yet maintain that motivating reasons for action are features of situations? There are three approaches that the situationist

could take. First, to argue that an important role for beliefs is compatible with a theory of action that argues that motivating reasons are features of the situation. However, I will argue that this fails because beliefs are a necessary part of the explanation and not merely a background condition. Secondly, the situationist could reject Dancy's normative constraint and thus argue that normative reasons are features of situations and motivating reasons psychological states. I argue that it is possible to reject the normative constraint, but that this does not help the situationist because it requires both features of the situation *and* psychological states of the agent to explain action. Finally, I consider an argument that all motivating reason explanations are factive, but that there are situations that although can be explained, the agent had no motivating reason to act. I will argue that again this does not seem to help the situationist because the causal link is between the psychological states of the agent and the action, not the features of the situation and the action.

There is a clear role for belief in the type of explanation that the situationist offers of actions. Can this be accommodated by the theory that motivating reasons are features of situations? Dancy argues that his theory of explanation of action has a role for belief in that 'if a reason is to explain an action it must be recognized (believed) by the agent' (2000a, 49). It appears possible to interpret his view as being that belief has an explanatory role because, in explaining an action, it is necessary that an agent believes that a state of affairs is a reason for action. There is a lack of clarity in his view as to what role is held by belief. Under his view, it appears that the state of affairs plays the role of a normative reason and that it is only in conjunction with a belief, acting as a motivating reason, that action occurs. The standard belief-desire model explanations, in which both belief and desire are necessary to explain action, are incompatible with this position because it appears that in this position belief has the motivating role. However, it does not seem clear that explanations under his view are incompatible with the view that beliefs alone can motivate action.

I don't see how Dancy's argument differs from saying that beliefs can motivate actions; we have to recognise (believe) a state of affairs for it to be a motivating reason for action for the agent. For Dancy, reasons are states of affairs and these are inert until an agent has a belief that something is a reason for action and this belief motivates the agent. Davidson illustrates this point when he says 'your stepping on my toes neither explains nor justifies my stepping on your toes unless I believe you stepped on my toes, but the belief alone, true or false, explains my action' (2001, 8). Such an account may remove desire from the

theory of action, but does not remove mental states entirely, because the belief, or recognition that something is a reason, is necessary for action. Normally the two components necessary for action are taken to be a belief about the world and a desire. Instead, Dancy could be interpreted as claiming that there is a single component necessary for action, namely a belief about the world acting as a motivating reason.

Dancy examines this pressure to consider it the belief that explains action in every case (2000a, 50-1). He thinks that this pressure arises when we consider false beliefs, as a false belief can explain why an agent acted in a certain way. If true beliefs also explain action, then he thinks that there is pressure to explain all actions in terms of belief. He asks the question 'if belief is admittedly present in every case, what role can it be playing other than that of a mental state, conceived of as a cause?' (2000a, 51). He answers this question by introducing the notion of an 'enabling condition'. He argues that a belief *enables* a state of affairs to explain an action without itself being *part of* the explanation (2000a, 51). He thinks that reasons explain why an action is right but that there can be other features of situations which are not themselves reasons but are necessary for an explanation to be correct. He gives the example of thinking that 'ought implies can'. He thinks that in this case, the ability to act as I ought is not among the reasons as to why I should so act. If I can act in this way that fact *enables* this to be a reason for me to act in a certain way.

Dancy focuses on the fact that there are lots of underlying features that allow explanations to be true and that we do not have to refer to these features in our explanations (2000a, 51-2). He attacks the idea that a complete explanation has to include all the events that have or have not happened to enable this action to take place. He calls these non-guaranteeing explanations; the explanation is compatible with the event not occurring: 'if O occurs in one place where the explanation goes through perfectly, and not in another, this will be because of some feature or features present (or absent) in the first and absent (or present) in the second. In that case the presence (or absence) of those features must be an enabling condition' (2000a, 52).

Consider again Collins's example of 'he is going to the ferry because the bridge is closed' (Collins 1997, 108). When the bridge is closed, this explanation goes through perfectly, but when the bridge is open, it does not. Dancy would argue that there is a false belief absent in the first case that is present in the second case and the absence of the false belief in the first case is what enables the explanation to be correct. This seems false; it is not the

94

absence of a false belief about the bridge that makes the explanation 'he is going to the ferry because the bridge is closed' true but whether in fact the bridge is open or closed. In the case where the belief is false, it is this belief that explains the action, so beliefs must also explain the action in the cases where the belief is true. Dancy does not establish that beliefs are inert features of situations that enable states of affairs to explain actions. It is not the case that sometimes when there is no belief a state of affairs can explain an action; a state of affairs can *never* explain an action if there is no belief about it, hence beliefs are a *necessary* part of the explanation of action. It is mere linguistic convention that we do not always refer to true beliefs in explanation of action. Where the concept of enabling conditions correctly comes in is in considering whether the belief that the bridge is closed guarantees that the agent will seek an alternative route. This belief is what Dancy calls a 'non-guaranteeing explanation' because it depends on enabling conditions such as the fact that the agent's business the other side of the bridge cannot be abandoned. The fact that the bridge is closed cannot itself explain anything the agent does because if the agent had no belief about the bridge, his actions would not be affected. Dancy seems to think that *a* belief is needed for action to occur and that it is irrelevant to the explanation whether that belief is true or false. I think it is extremely relevant as to whether the belief is true or false, so it must feature in the explanation, whether implicitly or explicitly. The notion of enabling conditions fails to make room for belief in Dancy's theory because beliefs are a necessary part of the explanation of action.

A second potential approach that the situationist could take is to argue that it is the combination of the belief and its content that explains and motivates an action, but just the content (i.e. the state of affairs in the world) that provides the normative reason for action. Dancy does not think that this is a plausible argument because of his commitment to the normative constraint, that motivating states must be the 'right sort of thing' to be normative reasons (2000b, 113). He argues that we have to think of normative reasons as facts or features of the situation; it is the state of affairs in the world that makes something the right thing to do (2000b, 104). Hence motivating reasons must also be states of affairs. From his commitment to the normative constraint and his belief that normative reasons are features of situations, it follows that explanatory reasons must also be features of situations and not psychological states of the agent, such as beliefs and desires. He continues to argue that because the requirements of morality are not all relative to the beliefs or desires of the agent and that such requirements are objective, then it cannot be the case that all normative reasons are psychological states (2000b, 105). He thinks that it has to be false that beliefs are reasons for action, because it is not possible for our reasons (beliefs) to make the action

right, hence normative reasons could not be motivating reasons. He argues that when we explain an action, we are trying to show that it was done for a good reason, but psychological states (beliefs) cannot be good reasons, only states of the world (2000b, 106). He thinks that the only good reasons are normative reasons. He claims that 'the need for the potential identity is itself supported by the need for motivating reasons to be of the right sort to be good reasons. If only normative reasons can be good reasons...then only the sorts of thing that are normative reasons can be motivating reasons' (2000b, 107).[30] In the suggested position, the motivating reason is different from the normative reason, so under Dancy's view the motivating reason cannot be a good one.

To avoid the consequence of Dancy's position that explanations in terms of reasons are not causal explanations we must drop the commitment to the identity between motivating and normative reasons. He wants the reasons for doing something to coincide with the reasons for which something was done and thinks that this coincidence cannot occur unless we are talking about the same reason in two ways, rather than two separate reasons (1996, 173). He bases his argument for this position on the maxim that a normative reason for action must, potentially, be able to explain and motivate an action (1996, 173). He derives this maxim from Williams's claim that 'if there are reasons for action, it must be that people sometimes act for those reasons' (1981, 102).[31] Williams continues by saying, if they do so act, then their reasons must feature in a correct explanation of their action. For example, George may have a desire for a gin and tonic within his subjective motivational set.[32] He can deliberate between various methods of satisfying this desire and conclude that the best way to satisfy this desire is to drink the gin and tonic that is in the glass on the table. In this case, George will have a reason to drink the gin and tonic that is in the glass on the table. A correct explanation of this action will refer to George's reasons. Dancy considers the position in which an action is right for reason A and a good explanation of the action appeals to the agent's belief in A (1995, 8). He thinks that this only shows that the

---

[30] In an earlier paper Dancy rejects Michael Smith's argument that motivating (explanatory) reasons are different from justifying (normative) reasons (1996, 172). Smith claims that justifying reasons are facts in the world and that motivating reasons are psychological states (1994). Dancy thinks that this is plain incoherent, because he thinks that it means that people cannot act for reasons that are right (1996, 172). He thinks that under this theory it is impossible for the reason that motivates an agent to also justify his action. He thinks this because, again under this theory, the two types of reason are 'metaphysically different'; the justifying reason is a fact and can never be a psychological state.

[31] This is part of Williams's argument that all reasons are internal reasons. See Grover (2008) for a defence of this position.

[32] Williams thinks that a subjective motivational set is broad, containing not only desires, but 'such things as dispositions of evaluation, patterns of emotional reaction, personal loyalties, and various projects, as they may be abstractly called, embodying commitments of the agent' (1981, 105).

normative facts can make a motivating belief-desire explanation of an action a good one, but it does not show that normative facts can motivate. He does not think that the normative truth 'figures *in* the explanation' (1995, 8).

Dancy constructs an argument to show this maxim is incompatible with the claim that there are two different types of reason (1996, 173):

1) All normative reasons in favour of doing actions are normative facts.

2) All motivating reasons why agents did certain actions are psychological states.

3) If an agent acts for a normative reason, then his normative reason must 'figure as such' in a correct explanation of his action.

4) Anything that can 'figure as a reason' in an explanation 'must be able to appear as one of the reasons *for which* the agent acted' (1996, 173).

5) 'Where an action is done for a reason, that reason must be among the reasons why the agent did it' (1996, 173).

Conclusion: all reasons are psychological states.

Dancy's argument is that this conclusion, that all reasons are psychological states, is incompatible with the first premise that normative reasons are facts, so although it may look as though there are two different types of reason, in fact, combined with premise three, we end up with an incompatibility, so there must in fact be only one type of reason.

I think that there is a problem with premise three. Williams criticises attempts to make a distinction between normative and motivating reasons, because he thinks that they are closely linked (2001, 93). However, although he thinks that the two types of reason are closely linked, he does not seem to be committed to Dancy's claim that there is in fact only one type of reason. Williams thinks that if we state that an agent has a reason to act in a certain way in the normative sense, then it must be possible that the agent could act in that way for that reason; hence that reason will feature in the explanation of the action. For example, if we say that an agent ought to help a person in distress, that he has a normative reason to help people in distress, it must be possible that the agent could help the person because she is in distress and that if he so acts any explanation of the helping action must include the fact that he acted in such a way because he had a reason to help people in distress. Williams thinks that if we explain why an agent did something by citing his reason for action, then we are saying that the action makes normative sense for that agent, relative to his subjective motivational set, or S. So if asked 'why did you do that?' the agent may reply 'because the person was in distress and needed help and that was a reason

for me to act in such a way'. By explaining his action using his reason, the agent reveals that this is what he thinks he ought to do in such situations, based upon his S. Williams makes it clear that by having a certain S the agent is disposed to have thoughts such as 'the person is in distress and needs help, so that is a reason for me to act' and to then act on those thoughts. It is not the case that we should have the thought 'the person is in distress and needs help, so that is a reason for me to act in virtue of my S'. Although Williams thinks that normative and motivating reasons are closely linked, he does maintain a distinction between the two. Motivating reasons report why an agent did something relative to his S, whereas normative reasons state what the agent should do, based upon what is possible given his S; i.e. he may not currently be aware of that reason because he has not yet deliberated effectively.

Clearly Williams thinks that the existence of two types of reason is consistent with his maxim that if an agent acts for a reason, then his reason must feature in a correct explanation of his action. I think that Dancy changes the wording of Williams's maxim to suit his argument. All Williams requires is that the normative reason 'feature in' the explanation. For Williams, normative reasons state what the agent should do based upon what is possible given his S. Because Williams ties normativity to the S of the agent, he avoids Dancy's conclusion by denying premise four, that anything that can 'figure as a reason' in an explanation 'must be able to appear as one of the reasons *for which* the agent acted', because all Williams requires is that the facts feature in the explanation, hence the facts can feature as the content of beliefs in the agent's S. So although the normative facts themselves are facts about the world, they feature in explanations as the content of beliefs.

By dropping the commitment to the normative constraint it is possible to argue that when p is true: '...because I believe p' is equivalent to '...because p' and both formulations provide the agent's reason for action and explain it. However, when p is false, the equivalence does not obtain. That I believe p can explain and motivate my action, but I have no normative reason to act in that way because I have a false belief. This position seems to hold for first and third personal perspectives. Such a view is put forward by R. Jay Wallace (Wallace 1990). He argues that motivating reasons are psychological states of the agent, whose direct objects are propositions. He thinks that only normative reasons are states of affairs. Following from this, he argues that an agent's motivating reasons can be good ones when their propositional objects represent states of affairs correctly; the motivating reason does not have to *be* the state of affairs that normatively makes the action a good one.

Although I argued above that Doris's argument appears committed to the position that features of the situation are motivating reasons, his argument appears, in places, to allow space for the psychological states of agents. This would allow his position to be consistent with the claim that motivating reasons are psychological states of the agent.[33] Is the analysis of the social psychology experiments compatible with the idea that it is the agent's perception of a situation that influences action rather than the features of the situation themselves? For example, it could be argued that the Isen and Levin experiment investigating the effect of mood on helping behaviour is a case where it is the agent's perception of the situation that is important, rather than the situation itself (Isen and Levin 1972 cited by Doris 2002, 30-2). Finding a dime is clearly a feature of the situation, but it is not this feature itself that determines helpful behaviour. Finding a dime is regarded as a

---

[33] Commonly psychological states in the form of a combination of belief and desire have been taken to provide the agent's reasons for action. David Hume stated that 'reason is, and ought only to be the slave of the passions', or, in other words, that desires are a necessary part of the explanation of action (*A Treatise of Human Nature* Bk2, Pt3, Sec3). G.F. Schueler gives the example of a burning building to illustrate Hume's theory of motivation: if one discovers the fact that the building is on fire, one will only get out of the building if one has a desire not to be burnt (2009, 105). If the individual lacks such a desire, the belief that the building is on fire will not move him to act. So the Humean position holds that a desire is necessarily part of the explanation of an action in terms of the agent's reasons for that action. Smith develops a Humean theory of motivation (Smith 1994). He claims that 'R at t constitutes a motivating reason of the agent A to φ iff there is some ψ such that R at t consists of an appropriately related desire of A to ψ and a belief that were she to φ she would ψ' (1994, 92). Schueler investigates what Smith means by a desire being 'appropriately related' to a belief (2009, 106-10). Schueler thinks that this relation is necessary to resolve the problem that an individual can believe premises that entail a conclusion without the individual being aware that the conclusion is entailed. To illustrate this point in terms of practical reasoning, he uses the following example: he wants to get to campus and knows how to get there by bus. However, his car won't start and in his hurry he does not put these two facts together. He then decides to get that bus to his sister's office to see if he can borrow her car. It may turn out that in these circumstances he gets on the bus, but his desire to get to campus and his belief that the bus goes there will not be what explains his action. He does have this belief and this desire, but this explanation is not the correct one (2009, 108-9). Schueler argues that it is essential that an explanation of an action includes how the beliefs and desires are 'put together' to avoid such counter-examples. In summary, the belief and desire (the agent's reason) together with the cognitive activity that the agent undertakes in putting the beliefs and desires together to result in action are sufficient to causally explain the agent's action. (See Pettit and Smith (1990) for further defence of a Humean account of practical reason).

Although note that there are potential problems with such an account. For example, Schueler identifies a dilemma: to avoid denying psychological facts about how agents deliberate, the Humeans need to claim that the motivating desires do not feature in deliberation, but if agents are not aware of their desires, they cannot motivate in the way the Humean claims, so desires must feature in deliberation (2009, 119). He argues that if it is the case that reasoning is required for the belief-desire explanation of an action, then it is not clear that the desire is necessary (2009, 110). He argues this because reasoning involves beliefs about things rather than the things themselves; i.e. beliefs about desires rather than the desires themselves. Schueler says if: 'the Humean Theory of Motivation must be understood so as to include Smith's 'putting-together point' then the claim that the desires that motivate actions are *ever*, let alone usually or always, completely 'in the background' simply cannot be true' (2009, 118). He thinks this because for the agent to put his belief together with his desire, the agent must do some reasoning, which means that the agent has to believe that he has certain desires. This means that the desire is not in the background, but a belief about the desire is necessarily involved in practical deliberation.

piece of good fortune by the agent, thus causing them to be in a good mood and it is this good mood that influences the helping behaviour. The agent's perception of the situation, finding a dime as a good thing, is what influences behaviour.

In the example of the Good Samaritan experiment it could be argued that what is important is the agent's belief that they are in a hurry. Hurry or the thought of being a few minutes late does not seem to be straightforwardly a feature of the situation that is independent of the agent. The agent must believe that they are in a hurry, as the situation itself does not exert time pressures; these are created by human perception. So again, one could here argue that what affects behaviour are not the actual facts of the situation; this is only an experiment, so there are no actual time pressures and they are not in fact in a hurry at all. This is an experimental situation in which the environment is manipulated to appear a certain way to the agent involved, so the perceptions and beliefs of the agent must play a central role. If the agent knew the purpose of the experiment in which he was taking part, he would not believe that he was in a hurry, so may have stopped to help. What causes the subjects to feel the time pressure is their perception that they are in a hurry.

Consider again the three possible explanations of the subjects' behaviour in relation to the Milgram experiments on obedience cited by Doris. I suggested above that these explanations all appear to refer to features of the situation to explain the agent's action, but they can be reinterpreted to involve the perceptions of the agent. It is not the incremental nature of the situation that influences the behaviour, but the agent's perception that he cannot justify stopping at one point over another. The fact that the experimenter tells the subject to continue four times before allowing him to stop the experiment caused the subject to feel pressure to continue, creating the perception that the subject is powerless to stop the experiment. The third explanation also makes reference to the perception of the subject; the situation no longer made sense to him. This perceived lack of clarity about the situation seems akin to the problem in group situations. In interviews following the experiments, many subjects attributed their actions to their perception that they had a lack of responsibility for the experiment and the welfare of the learner; the experimenter has 'the biggest share of the responsibility. I merely went on. Because I was following orders...' (Milgram 1974, 50).

So, it is possible to reject the demands of the normative constraint and maintain that motivating reasons are different from normative reasons. This allows the causal link

between motivating reasons and actions to be re-established.[34] However, although the causal link is re-established, it is unclear that this helps the situationist. Doris can avoid the consequence of Dancy's argument that action explanations in terms of reasons are not causal explanations, but this does not help support his general claim that features of the situation cause action. The rejection of Dancy's position involves commitment to the claim that a motivating reason is constituted by a psychological state, so it is psychological states of the agent that are of central importance in causing his actions and not features of the situation. The situationist may counter this objection by claiming that psychological states are influenced more by the situation than any other factors. However, it is difficult to see how the experimental evidence could be used to support this claim as there is no consideration of the agent's beliefs and desires, or motivating reasons for which the agent acted. Take the example of finding a dime improving helping behaviour. In this case it is proposed the agent's motivating reason for action is not that he found a dime or that he is in a good mood, but that he has a desire to help, coupled with an appropriate belief. However, this is only an assumption as the experiment does not report upon the motivating reasons or beliefs and desires of the subject. It merely focuses upon whether one factor, finding a dime, has a causal influence on behaviour. It is not under dispute whether trivial features of situations can affect behaviour; what is disputed is whether it is *primarily* such features that influence behaviour. It is unclear from the experiments that an agent's belief

---

[34] For example, a commonly accepted account is that of Davidson. He argues that an explanation of intentional action involves statement of the agent's reasons, that is his desires, intentions and beliefs that cause the action. He has a 'causalist' view of reasons for intentional actions. In his description of intentional action Davidson states that a pro-attitude (desire) and a means-end belief combine to form the agent's primary reason for action: 'R is a primary reason why an agent performed the action A under the description d only if R consists of a pro-attitude of the agent towards actions with a certain property, and a belief of the agent that A, under the description d, has that property' (2001, 5). Under this account, the agent's primary reason gives the cause of the agent's action.

However, there are objections to this account. How does the content of a belief determine action? Fred Dretske provides an example to illustrate this problem (1988, 79). In this example, a soprano sings an aria which shatters a glass. He argues that it is properties of the singing, such as the loudness, that are relevant to the breaking and not any facts about the lyrical content of the aria. He concludes that in such cases, it is properties of the singing that will feature in an explanation, not the 'content' properties. However, he argues that the case of action explanations is different because we believe that the content of the agent's belief is causally relevant.

Alternative accounts of explanations include Anscombe's, who argues that statements of the intention with which a thing was done provide a common sense explanation of why the agent acted in a certain way, yet these explanations make no reference to the reasons of the agent as causes of the action. She argues that an explanation will cite an agent's reason for acting, such as 'I killed him because he killed my father', but that his having killed my father is not a cause of my killing him, even though this is previous to the action (2000, 10). See also George Wilson's non-causal teleological account of reason explanations (Wilson 1989). For objections to the teleological account of explanations see Sehon (1997) and Roth (1999).

about a situation is formed independently of other factors that also have an influence, for example, the physical condition of the agent and his character. To take this example, mood is not only influenced by trivial features of the situation; some people are characteristically more positive than others, as captured in the question of whether a person views the glass as half empty or half full.

Finally, the situationist could argue that all reason explanations are factive, but that there are some actions that can be explained even though the agent has no motivating reason to act in that way. For example, Maria Alvarez rejects Dancy's view that reason explanations are not factive, because she does not think that a reason explanation can have a false explanans (2008, 53-65). Alvarez discusses motivating reasons, but this term is in itself ambiguous. She takes them to be identical with explanatory (she calls them operative) reasons, rather than normative (justificatory) reasons, so she maintains the distinction between two types of reason that is rejected by Dancy. She identifies two conceptions of motivating reasons. First, the psychological conception, whereby motivating reasons are an agent's believing something, i.e. it is a fact that the agent believes that p, and these reasons are mental states of the agent. Secondly, the non-psychological conception, whereby motivating reasons are what the agent believes i.e. the fact that p. She defines motivating reasons as 'the reasons for which we act when we act for reasons' and as 'what made the action right or appropriate in the agent's eyes' (2008, 55).[35]

Alvarez argues that the non-psychological conception is to be preferred. To support this argument she gives the example of a motivating reason for giving a cousin some money is that he is in need (2008, 55). She argues that discovering that he was not in need means that she was motivated by something that was not the case. She argues that what we discover not to be the case is that he was in need, not that she believed that he was in need, so what motivates the agent is whether or not he is in need, not the belief. She says that what is believed can be true or false because what is believed is a proposition. She continues to argue that if what is believed is true, then it can be a motivating reason, but if it is false, it cannot be a motivating *reason*, but can still motivate someone to act because the agent *takes it to be* a reason: 'What is believed, if true, can be a reason. What is believed and is false cannot be a reason, but it can still be what motivates someone to act. A motivating belief that is false is not a reason, though the agent takes it to be' (2008, 58).

---

[35] She leaves aside complications as to whether the agent can correctly identify his own motives and issues around the role of desires in motivation (she thinks that the reason for which someone acts is at least partly a belief).

However, this appears to conflict with her definition of a motivating reason as 'what made the action right or appropriate in the agent's eyes'. Under this definition, it appears that a false belief *can* be a motivating reason, as even though it was false, it made the action appear correct to the agent. She introduces the notion of an 'apparent reason', using the analogy of a Vermeer and a fake (apparent) Vermeer. However, the apparent Vermeer exists, but what kind of metaphysical entity is an apparent reason, if a reason is a state of affairs that does not exist?[36] This seems to be the same problem as faced by Dancy.

Alvarez argues that 'the fact that the agent takes it to be true explains why a false belief can still motivate him to act' (2008, 58). She is here saying that beliefs can motivate, so why are beliefs not motivating reasons? If it is beliefs that motivate, what role is the state of affairs p playing? Her argument here is unclear. She argues that we have to use 'psychological explanations' when the belief is false so the agent has no motivating reason, i.e. he did it because he believed that p (2008, 59). She explains that the fact that it is possible to construct psychological explanations in both veridical and error cases is taken to provide support for the psychological conception of motivating reasons. She thinks this is incorrect as in error cases there is no motivating reason that motivates the agent. Again, this does not seem to cohere with her definition of a motivating reason. Her argument as to why false beliefs cannot be motivating reasons is not clear. It appears that in these kinds of cases we would want to say that there is a motivating reason even though there is not a normative reason. However, I would argue that this does not mean that it is not the beliefs that are doing the motivational work. The agent has to believe p, q, etc. for them to have any effect; the state of affairs p does not motivate anything. It is the belief that p that is doing the work, not p.

Alvarez thinks that these psychological explanations are useful in explaining action, but this fact does not make them motivating reasons (2008, 59-60). Their uses are i) to identify motivating beliefs ii) to convey the speaker's view about what motivated the agent and iii) to specify the epistemic relation between the agent and what motivated her. In error cases we are not motivated by a reason, which raises the question of what were we motivated by? How are we motivated to do an intentional action without a reason? She draws a distinction between reason explanations and other types of explanations, but it seems that

---

[36] However, her argument here is unclear. She confuses matters by then saying 'an apparent reason and a genuine reason are both beliefs' (2008, 58) – she seems here to be equating reason with belief (mental state) rather than the fact or what is the case – precisely the view she is trying to argue against.

in the error case, the agent did have a motivating reason, in the sense that the action made sense to the agent; it is just that the motivating reason is based upon false information.

Again, this position does not seem to help the situationist. Although it accommodates the need for all motivating reason explanations to be factive, it does not adequately explain situations involving false beliefs. Alvarez identifies three types of action explanation. First, that the agent has a normative reason (a state of affairs); secondly, that the agent has a motivating reason (a state of affairs); and thirdly, that the agent has a false belief (a psychological state). It is unclear how the situationist could coherently claim that it is only false beliefs that have an explanatory and motivational role.

Conclusion

In conclusion, I have challenged the general argument of situationists such as Doris on the grounds that the position does not adequately consider an agent's motivating reasons for action despite their commitment to explaining actions. I have argued that the situationist claim that features of situations explain our actions is undermined by an incompatibility between the claims that features of situations cause behaviour and that features of the situation are motivating reasons for action. I first interpreted the situationist argument as saying that the features of the situation provide our motivating reasons for action. I proposed that such an interpretation commits the situationist to a controversial theory of action, such as that of Dancy. However, I argued that the situationist cannot accept the implications of this position that what explains an action need not obtain and that action explanations in terms of motivating reasons are not causal.

I have suggested that these consequences of Dancy's position can be avoided by dropping the commitment to the identity between motivating and normative reasons. If we drop the requirement to meet the normative constraint then the correlation between motivating reasons and normative states of affairs is enough to make a motivating reason a good reason; good motivating reasons are those that correspond with the normative states of affairs. Dancy's further argument that reasons are states of affairs depends upon reasons being both normative and explanatory and if we distinguish two types of reason, the argument does not stand.

However, if we distinguish two types of reason I argued that it is unclear how the evidence from social psychology can be used to support the situationist argument. Using this distinction, it is psychological states of the agent that are of central importance in motivating his actions rather than features of the situation. The experiments do not show that an agent's belief about a situation is formed independently of other factors that also have an influence, such as character. To avoid this objection, the situationist may well be prepared to accept that reason explanations are not causal. As stated above, Doris himself admits that 'situationism tells a disconcerting story about the way some behaviours are caused' (2002, 132). However, even if the situationist accepts that reason explanations are not causal, he is left with the problem of how a coherent explanation can be given of actions caused by events that did not happen. Alvarez's suggestion of non-reason explanations does not seem to help because it is unclear why a false belief can motivate and explain an action yet a true belief is inert in motivating and explaining action.

This chapter has highlighted a general concern with the situationist position, namely that there is a problem with the idea that it is features of the situation that motivate behaviour. I have proposed some ways in which the situationist may avoid this problem and shall return to this theme in Chapter Seven, where I shall propose an alternative account of explanation that improves upon that provided by the standard belief-desire model yet accommodates the issues raised by the situationist. First, the next chapter will consider a problem with Doris's account of local character traits. Investigation of this problem will raise further issues regarding explanation of action.

Introduction

The argumentative aim of this chapter is to raise a new problem with Doris's positive argument for local character traits. It will challenge Doris's argument that local character traits are stable. This is important because the idea of the local character trait is central to Doris's explanation of actions. In Chapter Four, I set out Doris's argument for local character traits and explained why he wants to make such an argument. In the first section of this chapter, I consider the problems with the specific and probabilistic analysis of local character traits provided by Doris, raising three particular objections. In the second section, I consider an alternative analysis provided by Michael Fara, but argue that this resolves only one of the objections I raise in the first section. The section concludes by evaluating Hampshire's argument that conditional statements of dispositions are different from statements about character traits. In section three I develop Hampshire's argument that sentences attributing character traits are summaries of past behaviour and propose that statements of character traits are best defined as summaries of historical narratives, explaining how this resolves the objections raised in the second section. This is different from the approach of other philosophers who oppose Doris because they concentrate on whether the social psychology experiments he uses provide evidence against a certain conception of character traits. Here I do not question the experimental evidence, but his definition of a local character trait.

1. Problems with local character traits

Doris first defines a character trait in the following way: 'If a person possesses a trait, that person will exhibit trait-relevant behaviour in trait-relevant eliciting conditions' (2002, 16). There are well known problems with this simple conditional analysis of character traits as dispositions. Fara gives the example of wood being disposed to burn when heated to illustrate the problem; wood is disposed to burn when heated but it does not burn if heated in a vacuum (2005, 50). He formulates a disposition ascription as: object N is disposed to act in a certain way M when in circumstances C, which formalises Doris's definition of a person having a character trait if he is disposed 'to behave in a certain way in certain eliciting conditions'. Fara agrees with Doris that dispositions have something to do with conditionals. He uses the example of ascribing the property of solubility to salt; this

doesn't just tell us how salt behaves, but how it will behave if you put it in water (2005, 46). He characterises this as the simple conditional account: object N is disposed to act in a certain way M when in circumstances C iff, if circumstances C obtain then object N acts in a certain way M. This formalises Doris's claim that a person possesses a trait, if that person will engage in trait-relevant behaviours in trait-relevant eliciting conditions.

Fara thinks that this account is too simple, because an ascription of a disposition can be false and the conditional true and, conversely, an ascription of a disposition can be true but the conditional false. Mark Johnston provides some examples to make a similar point (1992, 232). He argues that the relationship between a disposition and its dispositional conditional is complex. He gives the example of a gold chalice that is not fragile but 'an angel has taken a dislike to it because its garishness borders on sacrilege and so has decided to shatter it when it is dropped' (1992, 232). He argues that although the conditional statement that the chalice would shatter when dropped is true, this does not make it true that the chalice has the disposition of fragility because the causal explanation of the chalice breaking is something extrinsic to the chalice. Conversely, he gives the example that 'the glass cup is fragile but an angel has decided to make the cup shatterproof if it begins to fall to the ground or if it is about to be hit by a hammer, or enter any other condition of being struck' (1992, 232). In this example, the description of the cup as fragile is true, but the associated conditional, that it would shatter when dropped, is false. Johnston again argues that the explanation for this phenomenon is extrinsic to the dispositions of the cup. This problem is further demonstrated by C.B. Martin, who calls this the problem of finks (Martin 1994). This problem centres upon objects that lack a disposition to act in a certain way in specific circumstances, but, due to an external factor, would acquire the disposition in those specific circumstances. Martin uses the example of a dead wire that does not have the disposition to conduct electricity. He introduces an 'electro fink' that attaches to the dead wire to monitor if a conductor is about to touch the dead wire. If the conductor did touch the wire, so would the fink, making the wire now disposed to conduct electricity. In such cases, the claim that the dead wire has the disposition to conduct electricity is false, but the conditional statement that the dead wire would conduct electricity if touched by a conductor is true because of the fink. He also reverses the problem, whereby a fink can remove the disposition to conduct electricity, making the disposition statement true and the conditional statement false.

There have been numerous attempts to construct more sophisticated conditional accounts that avoid such problems.[37] For example, David Lewis's argument that an object N would act in a certain way M because circumstances C obtain *and* because object N has intrinsic property B (Lewis 1997). Johnston agrees with Lewis that there is something about the intrinsic properties of the object that are important for the relationship between the disposition and the dispositional conditional. This emerges in his discussion of the problem of masking for conditionals (1992, 233). He gives the example of a fragile glass that is wrapped in such a way that it will not shatter if dropped; it retains the disposition of fragility, even though it would not shatter if dropped. He argues that the fragility of the glass is masked by packing that is extrinsic to the glass and that causes the glass to be able to withstand being dropped. He concludes that there are intrinsic features of an object that provide the 'constituting basis' of its disposition to act in a certain way in certain conditions (1992, 234). He says: 'We may therefore think of a constituted disposition as a higher-order property of having some intrinsic properties which, oddities aside, would cause the manifestation of the disposition in the circumstances of manifestation' (1992, 234).

However, Fara thinks that this more sophisticated account remains inadequate because of the problem that the disposition can be masked. He argues that the fact that a disposition can be masked means any attempt to provide a conditional account will fail because of the problems associated with specifying the 'oddities'. He discusses several attempts to do this and explains why they fail. One attempt is to specify the conditions under which the disposition will manifest; the conditions should be 'ideal'. He thinks that the same kind of problem arises if we use a notion of normal or typical conditions. David Wiggins argues that a notion of normal conditions can work, using the following example to make his point: 'Just occasionally a horse that is surefooted may miss his step. He may even trip at the point where a less surefooted animal was lucky enough to have passed recently without slipping' (Wiggins 1998, 242). He argues that we expect an explanation to be found, even though we cannot always find one: 'Nor can we write down a list of conditional sentences...such that a horse is surefooted...if and only if these conditionals are true' (1998, 242-3). Instead, 'In ordinary cases the most we can do is to specify the disposition as that disposition in virtue of which *normally*, if for instance...be the case, then__will be the case, leaving theory, anecdotal knowledge, and whatever else to lend content to the 'normally' and the 'for instance'' (1998, 243). Fara thinks this idea fails because it is

---

[37] See Fara (2005) and Manley and Wasserman (2008) for summaries.

impossible to make sense of the concept of ideal or normal conditions without appeal to the masking of the disposition i.e. the ideal or normal conditions are those in which the disposition would manifest and is not masked (2005, 52).

Another attempt Fara discusses is one which claims we do not need to add a clause about normal or ideal conditions because this is implied by the context in which we use the conditional (2005, 54). The idea here is that it is not required for the disposition to manifest in every situation, just those situations that are suitable, as defined by the context. He thinks this approach gives rise to the question of how the context defines the circumstances and thinks that the range of conditions under which a disposition could manifest is too wide for them all to be defined by the context. He thinks that we would need an explicit account of the context for us to be able to say whether the conditional account was true.

Doris discounts the metaphysical worries about the concept of a disposition. He dismisses arguments as to whether dispositions should be analysed as conditionals as unimportant to his concerns. He thinks that if behaviour conforms to the conditional, it counts for attributing the disposition and if it does not, this may be a reason to deny the person has the disposition (2002, 16). He thinks that despite the metaphysical problems, the conditional statement sums up an important link between attribution of character traits and expected behaviour. Although he claims not to be concerned about the metaphysical implications of analysing dispositions as conditionals, Doris puts some work into reformulating the conditional. He acknowledges that his original formulation implies that a trait will unconditionally result in certain behaviour (2002, 19). He redefines the conditional as: 'if a person possesses a trait, that person will engage in trait-relevant behaviours in trait-relevant eliciting conditions with markedly above chance probability p' (2002, 19).

Doris takes the approach of making the conditions more specific to define his localised character traits. His notion of local character traits differs from the notion of global character traits because it makes their attribution more specific. He uses the example of 'sailing-in-rough-weather-with-one's-friends-courageous' rather than a general attribution of 'courageous' (2002, 115). The idea here is that if we have a specific enough dispositional statement, then it will be equivalent to a conditional. A very specific ascription of a disposition can avoid being masked by other features of the situation; if the conditions are not exactly as specified, then nothing can be said as to whether the object has the specific disposition. Doris argues against more general attributions of character

traits on the grounds that they are too often masked by features of the situation, so it is not open to him to accept that dispositions can be masked by other features. This conditional analysis of a disposition seems to be at the core of his first thesis of globalism, so I think it is important to establish that this is the correct analysis. If it is not, then it is possible that the arguments from situationism are missing their target. If character traits and dispositions should be understood in a different way, do the situationist arguments still cause problems, or do they just provide empirical evidence to back up the metaphysicians claim that a conditional analysis of dispositions and, therefore, character traits is false?

Having established Doris's argument for the existence of local character traits in Chapter Four, I now pose three objections to his view. The first objection derives from David Manley and Ryan Wasserman, who argue that such an approach does nothing to explain our use of ordinary ascriptions of dispositions (2008, 64). This seems true of Doris's local character traits; that George has the local character trait of 'sailing-in-rough-weather-with-his-friends-courageous' tells us very little, only that in the future if he is sailing with his friends in rough weather he will probably be courageous. What we expect the notion of a character trait to do is to draw links between George's courageousness in a variety of situations, such as standing up for his beliefs, defending himself against a mugger and skiing the steepest slopes, as well as sailing in rough weather. We want to conclude from this variety of evidence that George is in general courageous, not that he is courageous in four specific types of circumstance. The predictive value of character traits lies in what they may tell us about how a person may act in novel circumstances, not just ones that have already been experienced. We need to integrate the fragments together into something more akin to our traditional view of character from the perspective of the third person observer, as well as from the first person perspective.

The second objection asks: how does one specify a particular situation? Doris seems to have in mind some sort of statistical condition along the lines of 'if in a high enough percentage of this specific y type of situation an individual is likely to act in x way, then they have the specific trait of being x in y'. If we have only ever been sailing with the person in Doris's example in rough weather with his friends present, don't we have evidence that he is more generally 'sailing-in-rough-weather-courageous' as we have no evidence to the contrary? Conversely, if we have been sailing in rough weather with him and Fred and found him courageous, but not so when sailing with Peter don't we only have evidence to attribute 'sailing-in-rough-weather-with-one's-friend-Fred-courageous'? It is

unclear how one specifies the situation that appears in the dispositional statement. Even these more localised traits seem open to generalisation. We appear to have no better reason to use one generalisation over another. Doris himself admits that it is difficult to specify the conditions relevant for global dispositions to manifest (2002, 27). However, he argues that the empirical investigations he relies upon do not need to fully specify the concept of a disposition and its conditions to make the point that situational features are relevant to the manifestation of traits. This may well be true, but the question raised here is whether it is any *easier* to specify the conditions relevant for a local trait to manifest, or whether a full specification of a local character trait is plagued by the same problems as global traits.

The third objection is also presented by Manley and Wasserman (2008, 75-6). They do not think that it is true of many dispositional predicates that the object would respond in *most* or *markedly above chance* stimulus-circumstances (2008, 75). They give the example of Fred being disposed to be violent in the evening, even if this only happens a third of the time. They suggest that an object needs to respond in 'some suitable proportion' of stimulus-circumstances and that how big a proportion is 'suitable' will depend upon the dispositional predicate involved and the context (2008, 76). Doris thinks that the suitable proportion for dispositions of persons is 'markedly above chance', so over 50%. Under his definition, Fred would not have the disposition to be violent in the evening because there is only a 33% chance that this will occur on any occasion, so he thinks we cannot reliably predict how Fred will act.

The difference between the two positions hangs upon the factors that are considered. Doris simply considers Fred to either have the local trait of being violent in the evening and acting in accord with it, or not. Manley and Wasserman argue that several factors will be important, including the degree of manifestation of the disposition, the degree of probability of the disposition resulting in behaviour, how similar the case is to the actual world and the importance of the situation (2008, 79). However, they point out that when several factors are combined, there may be more than one way of doing this, resulting in competing measures with conflicting proportionality results (2008, 80). This is what appears to be happening in the disagreement with Doris as to what proportion is appropriate. Doris considers one factor and comes up with one proportionality result, but Manley and Wasserman consider a set of different factors and come up with a different proportionality result.

Doris identifies four features of attributions of character: they are descriptive, evaluative, provide explanations of behaviour and can be used to predict what people will do (2002, 15). The notion of a local character trait does not seem to meet these requirements. Under Doris's definition of a local character trait, the explanation of why an individual acted in a certain way would take the form 'George did x because he is disposed to do x in these specific local conditions with markedly above chance probability p'. Returning to the example of sailing-in-rough-weather-with-one's-friends-courageous, Doris's position implies that we would explain a certain action performed by George by saying 'he did that because he is disposed to be courageous when sailing in rough weather with his friends'. When pressed to explain what we mean by this, we would explain that George acted in that way because in those circumstances he acts in that way with markedly above chance probability. This explanation is not fully satisfactory because it merely describes George rather than explaining why George has this disposition and hence why he has acted in this way on that occasion.

Doris himself admits that this is a problem for explanations involving local character traits, but argues that this is not a problem unique to local traits as it also affects explanations involving global traits (2002, 66). He gives the example of explaining why an individual is social at an office party by saying that the individual has the trait of 'office party sociability'. He suggests that such an explanation is both circular and boring because he deduces the trait from the behaviour and then uses the trait to explain the behaviour. He argues that this is also a problem for global traits, as explaining the individual's behaviour on the grounds that he has the trait of 'sociability' is equally uninformative. However, he argues that explanations in terms of local traits are non-trivial because they involve the person as well as the situation. He thinks that all explanations in terms of traits will need a wider psychological context, including appeal to motives, goals, values, attitudes and strategies (2002, 66). The next chapter will consider in further detail what it is that we demand of an explanation of behaviour.

Local character traits also have a limited function as descriptions. The description of George as having the local character trait of 'sailing-in-rough-weather-with-his-friends-courageous' tells us little in terms of describing George. We may sometimes wish to describe George in relation to very specific circumstances, but we also often want to describe George in *general*. Further, local character traits are not useful for evaluation of persons or of actions. Identifying a particular act of being courageous whilst sailing in

rough weather with his friends as a particular instance of the local trait of 'sailing-in-rough-weather-with-his-friends-courageous' does not itself do any evaluative work. The action is evaluated favourably because it is *courageous*, not because of the specific circumstances. Neither does attribution of this local trait provide an evaluation of the person. As above, we may sometimes wish to evaluate George in relation to very specific circumstances, but we also often want to evaluate George in *general*. However, under Doris's account, this type of evaluative integration of local traits is not possible.

Finally, local character traits are not useful for predicting the behaviour of individuals. Taking the above example of a local trait, office party sociability, it is unclear what work the trait is itself doing. There is evidence that past behaviour is a useful predictor of future behaviour (Ajzen 1988, 99-101). Doris argues that attribution of local traits allows us to make accurate predictions of future behaviour across narrow situational conditions. However, it is not clear that attribution of a local trait is necessary to do this. If asked to predict how an individual will act at an office party using only past behaviour, one would not cite any trait in the prediction, but just behavioural evidence from the past. The everyday use of traits in prediction is to inform predictions of how individuals may act in novel circumstances. Trait terms are useful to summarise past behaviour across situations to create a general picture of how an individual acts. This information can then be used to inform a prediction of how he may act in novel circumstances. This idea, that trait terms summarise past behaviour, will be considered in further detail in the next section.

In summary, there are three problems with Doris's standard for attribution of evaluatively disintegrated local character traits; it is not clear that this is how we use character trait attributions; it is not clear how we specify a situation and it is not clear how the probabilistic proportion should be calculated. Although local traits seem to satisfy Doris's standard for attribution, that 'there is a markedly above chance probability that the trait-relevant behaviour will be displayed in the trait-relevant eliciting conditions', these objections reveal problems with this attribution standard (2002, 66). The specification of local traits does not resolve the issues for global traits of defining probabilities and specifying conditions. Nor do local traits provide any explanation of our everyday use of trait terms.

Is there a different approach to understanding attributions of dispositions? This section considers whether Fara's approach, defining dispositions as habituals, resolves the problems identified with Doris's standard for attribution of local character traits. Fara splits ascriptions of dispositions into two parts; agent N is disposed: agent N acts in a certain way M when in situation C (2005, 63). This formulation removes the conditional from the right hand side. He gives the example of 'John is hungry': 'John eats when he is hungry', which can consistently be true even if there are situations in which John is hungry but does not eat, for example if he is ill. Fara argues that the sentences of the form 'agent N acts in a certain way M when in situation C' are generalisations, so can admit some counter-examples (2005, 66). He argues that when we attribute a disposition to an object, we are not merely describing how it behaves, but that how it behaves somehow depends on how the object is (2005, 69). There is something about the object that explains why a statement attributing a disposition is true, so there is a causal relationship between the object and what it is disposed to do.

Applying this analysis to local character traits, if we take the example of 'George is 'sailing-in-rough-weather-courageous': 'George acts courageously when sailing in rough weather'; this can consistently be true even if there are sailing-in-rough-weather situations in which George does not act courageously, for example, if he is feeling seasick. If we analysed the ascription of sailing-in-rough-weather-courage as a conditional - 'If George were sailing-in-rough-weather-courageous, then he would act courageously in relevant situations' – this evidence would count against us making the ascription, but sentences of the former type can remain true even on occasions when the disposition fails to manifest. Fara would argue that we are not just describing how George behaves when we say he is sailing-in-rough-weather-courageous, but refer to a dependency between his local character trait and his action.

Defining local character traits as habituals or generalisations seems to resolve the third objection that was raised in Section Two. Because the conditional statement is removed from the right-hand side, there is no reference to probability, so there is no debate as to what proportion of occurrences of the behaviour need to manifest for the local character trait to be attributed. The probabilistic attribution of a local character trait is replaced with a generalisation which can admit counter-examples. However, this approach does not

resolve the first objection that the notion of a local character trait does nothing to explain our use of ordinary ascriptions of dispositions. As stated before, we expect the notion of a character trait to draw links between George's courageousness in a variety of situations, leading to the conclusion from this variety of evidence, that George is in general courageous, not that he is courageous in several specific types of circumstance. It also provides no answer to the second objection that asked how one specified a particular situation. Although the conditional has been removed from the right hand side, there is no clear methodology as to how we specify the situation C. As before, if we have only ever been sailing with George in rough weather, don't we have evidence that he is more generally 'sailing-courageous' as we have no evidence to the contrary? Or conversely, if we have been sailing in rough weather with him and his friends shouldn't we specify that he is 'sailing-in-rough-weather-with-one's-friends-courageous' based upon the evidence that we have? It remains unclear how one specifies the situation that appears in the dispositional statement. Although this analysis avoids situationist critiques about the reliability of traits, it seems to face problems in how we define a situation. For example, what is a 'situation that calls for generosity'? An attribution of generosity is a one-to-many relationship. If a person was only generous in one respect, would we regard him overall as a generous person? How many examples of situations in which John acts generously have to be on the right hand side of the equation for him to count as generous?

The situationist will simply deny that the generous behaviour is caused by the character trait. John will act generously because of features of the situation, not because of his character trait. We need to explain why John acts generously; Fara thinks that there is something about the object that causes it to behave in a certain way, but this is precisely what the situationist denies. Perhaps character trait attributions are not analogous to attributions of dispositions of objects. Georg von Wright argues that virtuous traits cannot be habits or dispositions (1963, 142-3). He argues that virtuous traits cannot be defined as habitual, regular action, as proposed by Fara, because habits differ from virtues. He thinks that habits are different from virtues because habits correspond to a particular activity, whereas virtuous traits do not. He uses a similar argument against virtuous traits being dispositions. He thinks that dispositions manifest under specific conditions, so 'the appearance, with some regularity, of these signs in the appropriate circumstances *decides* whether there is a disposition or not' (1963, 142). As he thinks that a virtuous trait has no link to specific actions, observation of action cannot decide whether or not a person has a virtuous trait in the same way that action can be observed to decide whether a person has a

disposition, for example, sneezing revealing an allergy. Doris rejects this argument on the grounds that ordinary dispositional attributions, such as toxicity, can also be difficult to specify (2002, 175n10). However, there may be other methods of analysing character traits that do not depend upon a dispositional account.

Is there an alternative approach to understanding character trait attributions? Stephen Hudson argues that virtues cannot be dispositions because, if they were, they could not also function as evaluative standards (1986, 36-41). As discussed in Chapter Two, virtues, rather than acts, have evaluative priority under a standard virtue ethical account. Similarly, Hudson argues that virtues are not merely the disposition that provides the tendency to act in accord with moral principles because, if this was the case, the virtues would be secondary to the moral principles, having only a derivative value. He grounds this argument in the observation that in everyday moral practice we do make judgements about what is cowardly, generous, etc. and that such judgements guide our actions in the same way as principles might. He argues that if dispositions 'are identified with certain behaviour patterns; they are, in reality, nothing more than the behaviour patterns', hence are only a tendency to act in a certain way rather than guiding how one ought to act (1986, 39). Under such an account he thinks that the value is attributed to the acts rather than the character. For this reason, he thinks that we should reject the analysis of virtues as dispositions. He does not think that a character trait is 'uniquely identifiable with a disposition to perform some item or pattern of publicly observable behaviour' but is rather part of a 'network of generalizations used in the explanation of human action' (1986, 41). Under his account, value can be attributed to both virtuous actions and to having a certain character.

From the above examples, it is clear that character traits are in an important sense different from dispositions of objects. Hampshire provides a good discussion of this (Hampshire 1953, 5-11). He argues that a failure of character trait-relevant behaviour does not disconfirm attribution of that trait:

> To attribute a disposition to someone is never to preclude that he may on some occasion act, or have acted, in some way contrary to his general tendency or disposition: that this is always possible is part of the force of calling statements of disposition summarising statements; statements describing what in general tends to happen are in this respect very unlike universal statements (1953, 7).

He thinks that disposition statements summarise past actions and are statements of general trends rather than universals. This seems to be true of both dispositions of objects and of

character traits. Hampshire challenges the view that descriptions of character involve hypothetical statements. He thinks that attribution of a disposition requires action or deliberation: 'One cannot normally say that someone is ambitious and generous, while denying that he has ever either acted or calculated in a generous and ambitious manner' (1953, 6). In contrast, an object that has a disposition does not ever have to display the disposition for it to be attributed to it, for example, fragility. Hampshire claims that:

> ...there is no necessary link between disposition and behaviour; one can have a disposition to think in a certain manner and also to react emotionally in a certain manner. Most ordinary character-descriptions refer compendiously to a tendency discernable equally in the behaviour, and in the thought and in the feelings of the subject (1953, 6).

Hampshire continues to argue that there are three grounds for separating statements about dispositions from descriptions of the causal properties of things (1953, 7). I interpret Hampshire to mean by 'disposition' what I have been referring to as a character trait and 'causal property' to be referring to dispositions of objects. He argues that dispositions attributed to non-human objects can be paraphrased as conditional if-then statements stating 'that if certain specific operations were performed upon the objects in question, certain specific reactions would be the effects of these operations' (1953, 7). His first ground for the distinction, is that a statement describing the disposition of an object does not imply that it has ever in fact happened, but a statement describing a character trait, such as kindness, implies that a person has acted kindly. Secondly, dispositions of objects do not have to manifest themselves over a period of time. Thirdly, dispositions of objects occur under specific and stable conditions, whereas human dispositions are 'vague, summary, interpretive and indeterminate' (1953, 8).

Hampshire here identifies an important difference between dispositions of objects and character traits. It is perfectly possible that an object could have an undiscovered disposition to behave in a certain way, but it does not seem the case that a person can be fully described as having a character trait unless it results in behaviour. Hampshire thinks that when we summarise someone's character we do not commit ourselves to a conditional statement that if he is in a relevant situation, then he will behave in a certain way: 'When one makes a statement about a disposition in the present tense, one is understood to be summarising the trend of someone's behaviour and calculations up to the time of speaking, together with the normal implication that his character is so far continuing to be the same' (1953, 9). He thinks that we are merely summarising how someone has behaved in the past and assuming that in the present his character will remain the same. He thinks that factual

statements such as 'his hair is yellow' and 'his hair has been yellow up to now' are no different from 'he is helpful' and 'he has been helpful up to now'. He thinks that in making a dispositional statement you are not committed to an if-then statement, although if I believe the if-then statement to be true, is it is because I make an inference from my belief that he has the disposition.

It may be objected that there is no such necessary link between having a disposition and behaviour. For example, consider a case where Scrooge is visited in the night and undergoes a thorough character conversion, so that in the morning he is disposed to be kind. However, he dies that morning before being able to act in accord with his new disposition of kindness. The challenge is that Scrooge's dispositions could have changed without him having the opportunity to act.[38] I would argue that this case supports the distinction that Hampshire is trying to make. It is true that an object may have a disposition without ever behaving in accord with it, but for an individual to be described as having a disposition, he has to have displayed the relevant behaviour in the past. A person is not either disposed to act in one way or not; possession of dispositions are a matter of degree. In the proposed case, Scrooge may have undergone a fundamental change, so may in the morning be more disposed to kindness than previously, hence he could be described as possessing the disposition to some degree, but cannot be described as fully possessing the trait of kindness without the relevant behavioural evidence.

Doris rejects Hampshire's analysis of dispositions as he reads Hampshire to conclude from the fact that an instance of failing to act in accord with a character trait does not disconfirm the attribution of that character trait that an attribution does not commit you to a prediction of behaviour. However, I do not think Hampshire is saying anything wildly different from Doris. Both acknowledge a strong link between disposition and behaviour and that behaviour is not sufficient to identify the disposition. Doris is rejecting the possibility that it is incorrect to analyse attributions of dispositions as conditionals without explicit argument. I think he needs to provide an argument that the conditional analysis is correct as it is not clear that this assumption should go unchallenged. A non-dispositional analysis of virtuous traits may be able to accommodate the issues raised by Doris yet retain a central role for character development in moral theory.

---

[38] Thanks to Alfredo Gaete and other participants at Open Minds IV, University of Manchester and the British Postgraduate Philosophers Association, King's College London conferences in 2009 for discussion on this point.

3. More on character traits as generalisations

If dispositions do not imply conditional statements, what does this mean for the challenge from situationism? The situationist thesis is that it is features of the situation that determine behaviour, not character traits. Doris argues that this means that character traits are not reliable because you cannot predict how someone will behave in a particular situation based upon his character traits; you are better off looking at what the majority of people would do in that situation and predict that they will comply with the norm. He thinks that if you attribute a character trait to someone then you should be able to predict with above chance probability how they will behave in a particular situation. However, perhaps we can argue, along with Hampshire, that attribution of a character trait does not commit you to a prediction of behaviour in the form of a conditional statement. We need some form of prediction to make sense of the character trait being reliable, but perhaps by avoiding the problems presented by the conditional analysis the situationist argument loses some of its force.

Take the example of 'John is generous' as a simple attribution of a character trait. Hampshire takes this as a mere fact that summarises how he has acted in the past. If you have no reason to believe he might change, then you can assume that he will continue to act in this way and make predictions, so you may also believe the conditional about future behaviour. Under this interpretation, this argument that character traits are summaries of past behaviour is compatible with the arguments from situationism, rather than defeating them. The attribution of a character trait is a summary of behaviour that makes no claim about what caused that behaviour, whether it is the character trait or features of the situation. It appears that to try and defend a notion of a reliable character trait against the situationist arguments we need to retain a causal link between the disposition and the behaviour. As Flanagan argues: 'They (character traits) are intended not merely as summary statements of behavioural regularities but as names of inner phenomena that play a causal role in the generation of behaviour' (1991, 279). Both the situationist and Hampshire deny a causal link between the disposition and behaviour, but the virtue ethicist seems to need to retain this.

Is the link between character traits and behaviour a conjunction, as Hampshire suggests? This would mean that for possession of a character trait it is necessary to have certain reasons *and* feelings *and* for these to result in action. The conditional analysis of such a

119

character trait would read: 'if a person possesses a trait, that person will have trait-relevant emotions *and* trait-relevant reasons for action *and* engage in trait-relevant behaviours in trait-relevant eliciting conditions with markedly above chance probability p'. Such a formulation weakens the evidence from the situationist experiments, as they do not investigate the emotions and reasons of the subjects, so cannot conclusively say that the failure of the trait-relevant behaviour is due to features of the situation rather than failure of emotion or reason in the individual. The experiments would need to show that there was a failure of emotion and reason and behaviour to show that it is purely the situation with no influence of character that causes the behaviour.[39] We need a story about where an agent's reasons for action come from and such a story may well make reference to underlying character traits. Again, the type of explanation required for moral evaluation is of central importance here; does a conditional analysis of a character trait provide an adequate explanation of the agent's action? Or is the required explanation rather a narrative story about the agent's past behaviour, where the story causally connects with the particular action now being explained? I will explore this question of what we require from an explanation of action, particularly in relation to the requirements of moral evaluation, in the next chapter.

Conclusion

In conclusion, I have questioned the notion of a local character trait as found in the work of Doris by considering the problems raised by analysing local character trait attributions as conditional statements. I summarised some problems with his specific and probabilistic analysis and concluded that the analysis of local character trait attributions as conditional statements makes explanations of action in terms of local character traits problematic. I proposed an alternative understanding of character traits, drawing on the work of Hampshire, suggesting that statements of character summarise a historical narrative.

The advantages of local character traits proposed by Doris are unsuccessful. I will now continue to offer an account of character traits that can gain an advantage by thinking differently about character traits and provide non-trivial explanations of action. In the next chapter I will consider what sort of explanation we may demand for moral evaluation, before developing an account of narrative attributions of character traits in subsequent chapters. I propose that we need a historical narrative explanation that explains *why*

---

[39] This problem was also discussed in the previous chapter in relation to reasons for action.

George is disposed to act in a certain way. This narrative-historical explanation will go further in explaining why he has certain motives and emotions. My proposal that a full and satisfactory explanation of how an action was caused will often be a historical narrative and that this historical narrative may also describe the character of the person will be developed further in the next four chapters. Under such a description of character as a narrative, individual character traits will be integrated with other relevant traits in explaining action.

Introduction

In Chapter Five I rejected Dancy's argument that features of the situation provide the explanatory reasons for our actions on the grounds that such explanations would be non-factive and non-causal. Further, I argued that psychological states are necessary for explanations of action. However, at the end of the previous chapter I proposed that we need a historical narrative explanation that explains *why* an individual is disposed to act in a certain way. This and the next three chapters defend the position that we need more than a statement of psychological states to explain an action; we need a historical narrative explanation.

In the first section of this chapter I start to defend my conception of narratives by questioning whether Humean belief-desire explanations provide the explanatory reasons for our actions. I address the question of what type of explanation is needed for moral evaluation of action. I consider an alternative approach suggested by Goldie, that we need a fuller explanation of action such as that provided by historical narrative explanations. I argue that this type of explanation has advantages over the explanations of action provided by the belief-desire model and by Doris's local character traits, a proposal that will be further developed over the subsequent chapters. In the second section of this chapter I outline my conception of narratives that I will defend over the next three chapters.

1. What do we require from an explanation; particularly what explanations are needed for moral evaluation of our actions?

In this section I introduce Goldie's argument that belief-desire explanations are not usually sufficient to explain our actions (2007, 104). Instead, he argues that a full and satisfactory explanation is based upon further information; I am particularly interested in his suggestion that this full explanation is the historical narrative of the events leading up to the action. I agree with his argument that we need a fuller account to satisfactorily explain an action because the belief-desire explanation only tells us why a certain action made sense for a particular individual (2007, 105). He does not think that a belief-desire explanation can explain why an individual acted in one way that made sense to him rather than another that would also have made sense to him. I propose that there are three advantages to the view

that explanations of action are historical narratives; first, the account accommodates the causality of action explanations; secondly, that action explanations are narrative explanations fits with our everyday experience; thirdly, historical narrative explanations of action are compatible with a character-based virtue ethical account.

Goldie argues that we usually need 'thicker' explanations than mere causal reason explanations: 'belief-desire explanations are seldom sufficient as explanations' (2007, 104). He thinks that although intentional actions can be explained in terms of causal belief-desire explanations, these do not 'fully and satisfactorily' explain the action and nor can these full, satisfactory explanations be reduced to a belief-desire explanation (2007, 103). He argues that we require explanations of action that are broader, going beyond a mere statement of belief and desire. The central claim is that explanation of action is complex and that beliefs and desires are two causes amongst many, rarely on their own sufficient to explain action.

Goldie argues that we require a fuller explanation because all the belief-desire explanation tells us is why a certain action made sense for a particular individual (2007, 105). He uses the example of an individual being given a free ice cream in a restaurant in a flavour other than that requested. He says that:

> ...it would make sense for one to tell the waiter to take it back; it would make sense for one to eat what one is given; it would make sense for one to leave it uneaten; it would make sense for one to throw the ice cream on the floor and walk out of the restaurant; it would make sense for one to offer the ice cream to one's host; it would make sense for one to pour the ice cream onto one's host's lap, etc. (2007, 105).

He uses this example to illustrate that a belief-desire explanation cannot explain why an individual acted in one way that made sense rather than another that would also have made sense. For example, if the diner left the ice cream uneaten, the belief-desire explanation may state that he did not to eat the ice cream, because he believed it to be strawberry ice cream and he desired chocolate ice cream. He thinks that this explanation is uninformative because it does not explain why he left his ice cream, for example, because he was inconsiderate towards his host. This reflects Dancy's point mentioned in Chapter Five, that a theory of causation cannot 'distinguish an active from an inactive but potential cause', so does not have the advantage of being able to analyse the relation between reason and action (2000a, 47-8).

Goldie argues that an explanation can be thickened out in many ways, including the motive behind the action, the character of the individual, emotional influences on thinking and the historical narrative leading up to the action. All these can help explain why an individual had certain beliefs and desires that led to a particular action, but I am here most interested in the latter. Goldie argues that historical narrative explanations are important because they can go further in explaining why an individual has a certain motive, character trait or emotion (2007, 111). He says:

> ...we often seek explanations of why someone had a particular motive, or why someone has a particular character or personality trait, or why someone was drunk, depressed or angry. And the explanations that we get are narrative-historical explanations: they locate the motive, the trait, the undue influence on thinking, within a wider nexus, in a way that enables us to explain more deeply why someone did the thing that they did (2007, 111).

Goldie continues to argue that explanations of action are not sufficient if they refer only to the belief and desire that caused it (2007, 112). He thinks that there are some special situations in which we turn to belief-desire explanations, for example, if a belief involved is false or to establish the belief that someone held (2007, 105). He gives the example of a belief-desire explanation as someone taking an umbrella because she believes that it will rain and does not want to get wet. He thinks that we may turn to this explanation if in fact it does not rain and we need to make reference to the false belief to explain the action, or we may need to investigate whether she believed that the umbrella she took was hers, hence the particular belief of the individual is important (2007, 105).

Goldie thinks that except in simple types of circumstances such as that of the umbrella, 'the belief-desire explanation is usually singularly unhelpful or even redundant, true as I accept it is' (2007, 112). Here I think that he means a belief and desire may be a necessary part of an explanation of an action and that such an explanation may correctly identify the immediate cause of the action. However, especially thinking in terms of moral evaluation of our actions, this causal belief-desire explanation does not give us enough detail to make such evaluations. We need an explanation that is a narrative-historical explanation that illuminates in more detail why someone did the thing that they did.

Goldie continues to raise the question of whether these thicker explanations really explain action. Or do they identify the causes of the causes of action, i.e. the causes of beliefs and desires? He thinks that the thicker explanation is not redundant in explaining why an agent acts in a certain way (2007, 112). He says 'our everyday action explanations' go 'far

beyond what is going on in the mind of the person doing the action' (2007, 112). So, to a certain extent, the debate as to whether both beliefs and desires or beliefs alone are necessary for action is not important. The important point is that reference to psychological states alone will not sufficiently explain action; we also need to tell a story about how the agent linked the beliefs or beliefs and desires together.

Dancy, for example, may well argue that these narrative details are merely enabling conditions. I however am arguing that sensitivity to features of situations or virtuous traits are part of the full explanation of the agent's action, rather than background conditions. When is a narrative explanation complete? There is a general problem of description; which description of events is needed for explanation of an action? Anscombe argues that a single action can have many different descriptions, such as 'sawing a plank, sawing oak, sawing one of Smith's planks, making a squeaky noise with the saw, making a great deal of sawdust and so on' (2000, 11) . She argues that this means that an agent does not necessarily know that he is doing something under a certain description. For example, he may know that he is sawing a plank, but not that he is sawing an oak plank. Davidson agrees with this argument that if a person does action F by doing action G, then the act of *F*ing is identical to the act of *G*ing' (2001, 4-5). He uses the example of someone alerting a burglar by turning on a light, which he does by flipping a switch. In this case there are several identical descriptions of the action: the alerting of the burglar is the same action as the illuminating of the room which is the same action as the turning on of the light which is the same action as the flipping of the switch. The alerting of the burglar is unintentional, but the flipping of the switch, the turning on of the light and the illuminating of the room were intentional. All these descriptions are identified with the same act, but do any of these descriptions provide an explanation of what the person did? I think that what explanation is appropriate depends upon what is being explained. If it is the turning on of the light that needs to be explained, reference will be made to the flipping of the switch, but if it is the alerting of the burglar that needs to be explained, reference will be made to the turning on of the light. This problem of when a narrative explanation is complete will be further addressed in the next three chapters.

I think that there are three initial advantages to the view that explanations of action are historical narratives that can be identified. First, the account accommodates the causality usually expected of such explanations. Lewis considers the causes of events, emphasising the complexity of causal histories and the problem of talking of 'the cause'. He argues that

'any particular event that we might wish to explain stands at the end of a long and complicated causal history...in the world as we know it, the only question is whether they are infinite or merely enormous' (1987, 214). He thinks that an event is 'the culmination of countless distinct, converging causal chains' (1987, 214). Taking actions as a subset of events, a similar argument may apply to explanations of actions.[40] Causal histories of actions could also be argued to be complex and face a related problem of talking of 'the action' (Wilson 1989). George Wilson argues that there could be a causal chain of actions in the performance of a simple action, such as moving your arm (1989, 18-19). This could involve actions such as initiating neural activity, muscle contractions, and the overt arm movement. Hence, a particular action, such as turning on a light, might be at the end of a causal history which links together numerous smaller actions.

Lewis continues to argue that the complexity of the causal history of an event (or action) leads to disagreement as to 'which part of the causal history is most salient for the purposes of some particular enquiry' (1987, 215). He thinks that for an explanation to provide information about an event it must 'provide some information about its causal history' (1987, 217). He says 'the why-question concerning a particular event is a request for explanatory information' (1987, 218). Taking this definition of the explanations of events, I think that it is clear that historical narratives can be causal explanations of events.[41] The events in a narrative are causally connected and one event can explain why a later event occurs, providing an answer to the question 'why?' (Forster 1927; Carroll 2001).

Secondly, an advantage of the position that action explanations are narrative explanations is that it fits with our everyday experience and use of explanations. In Chapter Six I criticised Doris's local character traits on the grounds that they do not explain our ordinary use of character trait ascriptions. Under Doris's definition of a local character trait, the explanation of why an individual acted in a certain way would be because in those circumstances he acts in that way with markedly above chance probability. This explanation is not fully satisfactory because it merely describes the individual rather than explaining why the individual has this disposition and hence why he has acted in this way on that occasion. The proposal that explanations are narratives will provide an alternative approach that better accords with our use of character trait attributions to explain action because it will allow us to conclude from a variety of evidence that an individual has a

---

[40] Davidson argues that actions are a subset or type of events (Davidson 2001).

[41] I will put forward an argument as to why these further explanations of action have to take narrative form in Chapter Nine.

certain trait. The narrative structure will integrate the fragments together into something more akin to our traditional view of character than that proposed by Doris. Doris argues that explanations in terms of local traits are non-trivial because they involve the person as well as the situation. He thinks that all explanations in terms of these local traits have a wider psychological context, including appeal to motives, goals, values, attitudes and strategies (2002, 66). In Chapter Six I disagreed, arguing that Doris's local character traits have no advantage in this respect over more general traits. Over the next three chapters I will develop this proposal that explanations of actions and of character trait attributions are narratives, arguing that this alternative account of explanations does have an advantage over Doris's local traits because such explanations do make appeal to situations and the wider psychological context.

The third advantage is that historical narrative explanations of action are compatible with a character-based virtue ethical account. The narrative account seems to accommodate the focus on *more* than just how the person acts in one situation, as recommended by virtue ethics, identified as an advantage of the view in Chapter One. The situationist would perhaps agree that the agent has to have beliefs and/or desires about the situation that they are in for action to occur, but would, as has been suggested, perhaps argue that that it is primarily the features of the situation that influence the beliefs of the individual and therefore indirectly explain the action. This would mean that evaluation of the action is focused on the features of the current situation and not the person. However, Cheryl Misak argues that ethics is an 'experience-driven inquiry' and that experiences explained in a narrative can lead us to see something we had not previously realised, for both the narrator and the audience (2008, 621). Under such a view, we cannot evaluate a particular action largely based upon the features of that situation, but also need to take in to account the biographical narrative of the individual.

Further, the narrative account seems to accommodate the account of moral development put forward by virtue ethics, identified as an advantage of the view in Chapter One. Is experience necessary for moral deliberation and, therefore, explanation? Misak thinks that at least some of our moral judgements aim at the truth, so she thinks that our moral judgements must be 'responsive to experience' and that autobiographical narrative is a method of making our moral judgements responsive to experience (2008, 621). She says 'everything we experience is interpreted - the experiential data that we possess are not raw experiences but rather beliefs about what we experienced' (2008, 620). So, it is not the

127

features of the situation that explain what we do, but our beliefs about the situation. She argues that autobiographical narrative might contribute to our knowledge; that some insights can be gained only through experience and that others may learn something from our accounts of those experiences (2008, 615-6). She argues that narratives are subject to rational criticism and that narratives about our subjective experience can therefore play an essential role in ethics (2008, 618). Her argument lends support to the idea that the action of an agent cannot usually be explained by features of the immediate situation, but by a wider explanation that takes into account the biographical narrative of the agent. Further, we need to both experience things for ourselves and learn from the narratives of others to gain such knowledge.

Misak develops an account of why narrative may be important to ethics. She thinks that 'when we reflect upon and describe those experiences which embody moral responses, we can gain epistemic access to values or norms' (2008, 623). She continues to say that 'the interpretation or description that is involved will be thicker when what we are describing is the experience of having had a moral insight' (2008, 623). She agrees with Goldie, arguing that from the internal perspective of a participant in a narrative we can gain an insight into how an event was experienced by the participant (Goldie 2003a, 54-68).[42] She thinks that use of this type of narrative is uncontroversial in ethics because we need such methods to, for example, be able to understand the ends of others and to weigh consequences for others (2008, 625). She continues to agree with Goldie, arguing that there is also an external, thicker, perspective in a narrative, which includes the narrator's interpretation and evaluation of the events aiming at getting the audience to respond in a certain way (2008, 624). In such cases, she thinks the audience end up sharing the moral insight of the narrator, not a mere understanding of how the event was experienced by the participant (2008, 625). She concludes that there are two ways in which autobiographical narrative is important to ethics; first, she thinks that we can test our beliefs about how others experience certain events by comparing them to reports from people who have actually had such experiences and secondly, she thinks that we can 'test our moral principles against the moral insights of those who have had relevant and distinctive experiences' (2008, 626). In the next three chapters I will expand upon some of these ideas in developing an account of narratives and relating this to explanations and moral evaluation of people and their actions. First, I will outline my general account of narratives, before providing detailed support of the view over the next chapters.

---

[42] Goldie's argument will be considered in more detail in Chapter Eight.

2. What is a narrative?

In this section I will outline my view of narratives which will be defended in detail over the remaining chapters. I make three central claims about narratives in general. My first key claim about narratives is that it is not possible to specify the necessary and sufficient conditions for narratives. My second general claim about narratives is that I think that there is a certain type of knowledge given to us by narratives; emotional knowledge of how to feel about the narrated events and/or the protagonists. My third general claim about narratives is that this argument that narratives give rise to a certain sort of knowledge can be used to support Goldie's claim that narratives are needed for full, satisfactory explanations of people and their actions. I then make three claims specifically about real-life historical narratives about real events and real people. My first specific claim is that historical narratives are necessary to explain why an individual acts in a certain way in a certain situation. My second specific claim is that historical narratives of real events and people are capable of assessment in terms of truth and falsity. My final claim is that this historical narrative explanation may correspond with a character trait attribution.

There are many different words that are used to identify things that we might regard as narratives. For example: account, story, tale, chronicle, anecdote, description, plot, report, yarn, recital, score, depiction, monograph, portrayal, legend, saga, drama, fable, epic, and so on. Sometimes there are similarities between these examples, but there can also be differences. For example, a chronicle is defined by *The Oxford English Dictionary* as 'a detailed and continuous register of events in order of time; a historical record, *esp.* one in which the facts are narrated without philosophic treatment, or any attempt at literary style'. Yet one definition of a story provided by *The Oxford English Dictionary* is 'a narrative of real or, more usually, fictitious events, designed for the entertainment of the hearer or reader; a series of traditional or imaginary incidents forming the matter of such a narrative; a tale'. There are similarities in that both definitions refer to narration and to series of events. But there are differences in the definitions in that the latter aims at entertainment, but the former at historical record. This suggests that hard and fast necessary and sufficient conditions that identify all, and only all, narratives may be challenging to come by.

Similarly there are many different art-forms that we might regard as narratives. For example, novels, books, poems, films, fiction, history, comic strips, paintings, music, ballet, autobiographies, biographies, diaries, journals, memoirs, and so on. Sometimes

there are similarities between these different art-forms; for example, each narrative tells a story about a sequence of somehow related events. However, these art-forms can also have substantial differences. Consider the telling of a story through the medium of dance. For example in the ballet *Swan Lake* the story of the tragic love of Princess Odette, the Swan Queen, and Prince Siegfried is dramatized through music and dance. This is very different from the narrative of a novel such as *Sense and Sensibility*, which tells the story of the Dashwood family and the daughters' pursuit of love, happiness and marriage in early nineteenth century England. This is itself different again from the portrayal of events in a film. These differences include the complexity and realism of the narrative in that a novel will tend to be more detailed than a film or ballet because it lacks the visual component of the others. Again, this suggests that hard and fast necessary and sufficient conditions that identify all, and only all, narratives may be hard to come by.

My first key claim about narratives, therefore, is that it is not possible to specify the necessary and sufficient conditions for narratives. I claim this because the diversity in narrative art-forms indicated above suggests that it may not be possible to identify necessary and sufficient conditions for narratives. The definitions of 'chronicle' and 'story' contain as many differences as similarities. Again, a ballet, defined by *The Oxford English Dictionary* as 'a dramatic entertainment consisting of dance and mime performed to music; (in early use) a theatrical spectacle intended to illustrate the costumes and culture of other nations, or to dramatize through music and dance some myth or narrative; (later) a theatrical dance performance using precise and highly formalized set steps and techniques' is very different from a novel, defined by *The Oxford English Dictionary* as 'a long fictional prose narrative, usually filling one or more volumes and typically representing character and action with some degree of realism and complexity'. These differences are not limited to the method of narration i.e. visual, aural or written. The first section of the next chapter will expand on these observations and argue that in fact we cannot identify necessary and sufficient conditions for narratives. I think that that we can identify only general themes and commonalities between these different art-forms and different things we call 'narratives'. Not all narratives will contain every element, nor will possession of all the elements necessarily mean that something is a narrative.

I base my claim that we cannot identify the necessary and sufficient conditions for narratives on the rejection of other attempts to specify such conditions. I support my claim by considering what Aristotle, Forster and Carroll identify as the necessary conditions for

narrativity and argue that these attempts fail. I object to the claim that it is a necessary condition of narrativity that a complete narrative must have a story with an end, as argued by Aristotle and Forster (Aristotle's *Poetics* and Forster 1927). I also reject Forster's claim that it is a necessary condition of narratives that they must be told in time sequence. I continue to analyse the five-part definition of a 'narrative connection' put forward by Carroll (Carroll 2001). I argue that although he may identify some necessary conditions for narratives, I agree with the work of Velleman who argues that these conditions are not jointly sufficient to identify a narrative (Velleman 2003). However, I continue to reject Velleman's further argument that narratives are necessarily emotive. He argues that a narrative necessarily resolves an 'emotional cadence' in the audience but I reject this claim on the grounds that an audience can understand a narrative without necessarily feeling any particular emotion.

Instead I think that we can identify general themes that appear in most things we are happy to call narratives. My claim is that the themes we identify in narratives also often occur in other non-narratives, such as descriptions or causal explanations. I think that it is possible to define some general conditions for a narrative and that we do not need a rigid definition. I do not think that it is necessary for all narratives to be forward-looking or that all narratives have to be temporally-ordered, yet most narratives are commonly structured in such a way. I do agree with Carroll that a narrative *usually* involves at least two events/states of affairs; that the earlier events are *normally* at least causally necessary conditions of the later events; and that there is *normally* one unified subject. What I disagree with is that these conditions are necessary and sufficient for all narratives. I think that there are examples that do not meet these conditions yet we would normally call them a narrative; perhaps a David Lynch film is an example here. Further, I think that some things meet these conditions yet we would not normally call them narratives. I here agree with Velleman that causal explanations seem to meet Carroll's criteria but are not necessarily narratives.

My second general claim about narratives is that I think that there is a certain type of knowledge given to us by narratives; emotional knowledge of how to feel about the narrated events, and/or the protagonists, and/or the narrator. As introduced above, I agree with Misak that some knowledge can be gained only through experience, but by narrating these experiences to others, we may be able to pass on some of that knowledge to others (2008, 615-6). I claim that knowledge of how to feel about an experience falls into this

category of knowledge that can only be passed on by a narrative. I think that for an event to be fully explained the explanation needs to reveal information about the emotions involved. The audience has to know how to feel about the events portrayed, and/or know how the protagonists feel about the events portrayed, and/or how the narrator feels about the events portrayed. Hence the further explanation of the events has to take narrative form as this is the best way of passing on this information about emotions to others. This is not to say that a narrative is *necessarily* emotive, in that the audience has to *feel* a certain emotion; the audience can get a sense of how they should feel even if the narrative is not successful in achieving this aim.

My support for this second claim begins in the next chapter. I there agree with Velleman that it is the structure of a narrative that gives rise to this sort of emotional knowledge and that such knowledge is not available from another source. I agree with Velleman that the knowledge and understanding gained from a narrative has to derive from the fact that it is a narrative, otherwise the narrative structure is not necessary to gain this information and it could be derived from another source, such as a non-narrative causal explanation (2003, 18). As I note above, Velleman argues that the narrative initiates and resolves 'an emotional cadence in the audience' (2003, 18) and that this completion of an emotional sequence means that the audience understands the narrated events (2003, 19). I disagree with the specifics of Velleman's account because I do not think it necessary that an emotional sequence is completed in the audience. However, I agree that there is a strong link between narratives, emotion and understanding. It is not possible for us all to experience every type of event and the accompanying emotions, thus gaining all our knowledge through experience. Instead we gain knowledge and understanding of the experiences of others through the stories they tell us and similarly we pass on knowledge and understanding by narrating our experiences to others.

My third general claim about narratives is that the claim that the narrative structure gives rise to a certain sort of emotional knowledge can be used to support Goldie's claim that narratives are needed for full, satisfactory explanations of people and their actions, fictional or otherwise. My methodology is to argue that a narrative is generally more than a list of time-ordered events and the events normally have a causal connection that means one event can explain why a later event occurs. As explained above, I think that these general themes normally found in narratives give rise to the knowledge and understanding of the emotions of the protagonists and/or the narrator and/or the audience. That we know

how to feel about the events, or know how others feel about the events, is an indication that the events have been understood and that the explanation is satisfactory. If, for example, the audience does not know how to feel about the events portrayed in a narrative the indication is that the narrative has omitted some information or is in some other sense incomplete, so the audience cannot share the full understanding of the events as possessed by the narrator or the protagonists.

Further, I think that narratives are needed for evaluation of the people and events portrayed, not just explanation. I claim that there is a close link between understanding the events portrayed and evaluation of those events. As I think a narrative explanation of the events is necessary to gain knowledge of how to feel about the events, it follows that a narrative explanation is also needed for a reliable evaluation of the events. A bare statement of the facts does not give the audience enough information to reliably evaluate the events; the contextual details and emotional perspective of the narrator provide information about the complexity of the events needed to evaluate the events. Of course, lack of information or weak narratives do not prevent an individual from erroneously *thinking* that he understands the events portrayed or *thinking* that he is correct to evaluate the protagonist or her actions in a certain way. In fact, this is often used as a device in narrative art-forms, such as novels, to lead the audience astray in their evaluation of a character, so in fictional accounts this distortion or misunderstanding can provide an important part of the story.

Thus far I have been introducing my account of narratives in general. However, I am most interested in historical narratives about real-life events and people. This is because I think that narratives are needed to explain fully a person and what they do in real-life, not just in fictional accounts. Further, I claim that such narratives about real events will always refer to what the person has done in the past. *So, my fourth specific claim is about real-life people and events: I claim that historical narratives are necessary to explain why an individual acts in a certain way in a certain situation.* There are different types of explanatory historical narratives about people. An explanation may take the form of an autobiographical narrative about one's own history or it may take the form of a narrative about a third person. Autobiographical narratives are different from other narratives in that the narrator and protagonist is the same person. My claim is that such autobiographical narratives are important in terms of understanding and evaluating ourselves, as well as in

justifying and explaining our actions to others. Similarly, biographical narratives about others are important for understanding and evaluation of the person and their actions.

My fifth claim is that historical narratives are capable of assessment in terms of truth and falsity. There are well documented problems of distortion in constructing a narrative. The narrative may find patterns and significance to events that they did not have in reality. For example, Peter Lamarque points out that in a historical narrative the narrator selects and orders the events, so is choosing which events to emphasise and is drawing connections between the different events. He argues that narrators have to do this from some point of view, hence, despite their best intentions, will be interpreting the events portrayed and may distort the significance of the events (2004, 398).

However, I agree with Lamarque that although a narrative can distort the events it portrays, this does not prevent it giving rise to knowledge and understanding (2004, 398). I think that a historical narrative can be evaluated in terms of truth and falsity. I argue this because the historian is generally aiming at a true representation of events so cannot impose any old pattern on the events of the past, finding meaning and connections where there are none. Because the historian is aiming at a true representation of events, he will draw upon numerous sources of evidence in constructing his narrative. This means that the historian's interpretation can be corroborated by other evidence from other sources, so can be evaluated as a true or false representation of the events.

My final claim links together explanations of particular actions with evaluations of persons. A historical narrative that explains why an individual acts in a certain way in a certain situation may correspond with a character trait attribution which evaluates that person and the totality of his actions. I think that there is a close relationship between explanations of actions and evaluations of character. Not all historical narrative explanations will correspond with attribution of a character trait because there may well be no regular pattern of behaviour to be found in the historical account of the individual. Nor will there always be a historical narrative explanation of a certain action; perhaps you just want an ice-cream. Commonly, though, human action is more complex than this and often we will need the historical narrative to explain why the individual acts in a certain way and to identify patterns of behaviour consistent with possession of a character trait.

My first two claims about narratives, that it is not possible to specify the necessary and sufficient conditions for narratives and that there is a certain type of knowledge given to us by narratives, will be defended in the next chapter. Chapter Nine will develop an argument to support my agreement with Goldie that narratives are necessary for a full explanation of an individual and his actions, before narrowing the focus to historical narratives about real events and real people and presenting an argument to support my claim that such narratives about real events and people are capable of assessment in terms of truth and falsity. Chapter Ten will present my arguments to support my final claim that historical narrative explanations may correspond with character trait attributions.

Conclusion

In this chapter I have proposed that the action of an agent cannot usually be explained by features of the immediate situation. An explanation of an action usually needs to be wider, drawing on the biographical narrative of the agent. In a simple case, a belief-desire explanation may be sufficient. Consider again Goldie's example of where a belief-desire explanation works when someone taking an umbrella because she believes that it will rain and does not want to get wet, functions as an explanation if in fact it does not rain and we need to make reference to the false belief to explain the action. However, I have raised several issues with the belief-desire model providing reasons for our explanations of actions. I concluded by agreeing with Goldie that usually we require a fuller explanation in more complex situations. It may be that in some simple cases an action can be explained with a short narrative; for example, finding a dime puts you in a good mood and therefore you have a belief-desire combination to help an individual pick up the papers that he dropped thus help the individual pick up the papers. In other cases the narrative explanation may be less clear or more complex. For example, in the Milgram case a simple explanation could be that that the situation pressurised the participant to comply or could be a belief-desire explanation along the lines of a desire to complete the experiment successfully and a belief that the best way to do this was to do as you were told. However, such simple explanations do not satisfy the audience; they require a fuller explanation of why the individual succumbed to the pressure or believed that the best thing to do was to obey the experimenter. My proposal is that a fuller explanation will take narrative form. I have here outlined my general claims about narratives that I will defend over the next three chapters. In the next chapter I shall start to defend this position by considering whether we can define the necessary and sufficient conditions of a narrative.

Introduction

The argumentative aim of this chapter is to establish what, *in general*, a narrative is. My central thesis is that we cannot define the necessary and sufficient conditions of a narrative. My argument is that we can identify themes that occur in many things that we are happy to call narratives, but that also occur in others that we would perhaps not, such as descriptions or causal explanations. Any attempt to specify the necessary conditions of narrativity will be futile due to a proliferation of counter-examples from a wide range of different art-forms. I will support this claim by considering what Aristotle, Forster and Carroll identify as the necessary conditions for narrativity and argue that these attempts fail.

In section one of this chapter I object to the claims that a complete narrative must have a story with an end and that it must be told in time sequence are necessary conditions for a narrative. In section two I analyse the definition of a narrative put forward by Carroll. I argue that although he may identify some necessary conditions for narratives, I draw upon the work of Velleman to argue that these conditions are not jointly sufficient to identify a narrative. In section three I argue that narratives are not necessarily emotive. I outline some objections to Velleman's argument that narratives are distinctive because they provide the audience with an understanding of how to feel about the events portrayed. I respond to these objections on behalf of Velleman but reject the idea that we can specify the necessary and sufficient conditions of narrativity. I conclude that a narrative is distinctive on the grounds of its general features rather than a set of necessary and sufficient conditions.

A narrative tells a story about events that have occurred. Goldie says 'a narrative is something that can be narrated; it need not be narrated though, for a narrative can be just 'thought through', as, for example when one remembers or imagines a sequence of events' (2003a, 54). However, a narrative does more than merely list the events as it portrays them as a coherent whole, explaining how the events are related. Goldie acknowledges this when he describes a narrative as 'more than just a bare chronicle, but a representation of the events, and of the people involved, which is organized, shaped and coloured in a way that gives coherence, meaningfulness, and emotional import to the represented events' (2003a, 54). Is it possible to identify what is characteristic of all and only all narratives?

And, what is a good narrative? Velleman suggests that 'what makes a story a good story specifically as a story – what makes it a good example of storytelling, or narrative – is its excellence at a particular way of organizing events into an intelligible whole' (2003, 1). What is meant by this 'particular way of organising events' that creates the 'coherence, meaningfulness and emotional import' characteristic of a narrative? Do narratives have to contain causal connections? Why are coherence, meaningfulness and emotional import important to a narrative? This chapter will address these questions.

1. Endings and time sequences

It is argued that a complete narrative must have a story with an end, in the sense that there are no further questions for the audience as to what happens next. Both Aristotle and Forster argue that this is a necessary condition for a narrative (Aristotle *Poetics*; Forster 1927). Aristotle argues that the coherence of a narrative derives from the plot, which organises the events within a structure. He considers how a narrative may be identified by examining the structure of a story:

> A whole is that which has a beginning, a middle and an end. A beginning is that which itself does not follow necessarily from anything else, but some second thing naturally exists or occurs after it. Conversely, an end is that which does itself naturally follow from something else; either necessarily or in general, but there is nothing else after it. A middle is that which itself comes after something else, and some other thing comes after it (*Poetics* 1450b26-33).

He claims that a plot has a definite beginning and end, starting with an event that is not a necessary consequence of something else and ending with an event that has no subsequent necessary consequence. The idea seems to be that for a plot to be complete under the Aristotelian picture there should be no further questions for the audience as to why the first event happened and as to what happened next. Aristotle gives an example of a plot: 'the statue of Mitys in Argos killed the man who was responsible for Mitys' death by falling on top of him as he was looking at it' (*Poetics* 1452a8-9). This story implies that this chance event happened for a purpose, drawing connections between one event and another. He says that 'There is an important difference between a set of events happening *because* of certain other events and *after* certain other events' (*Poetics* 1452a11-21). Here he identifies that a plot involves necessary or probable links between events; one event has to follow another necessarily or probably to lend the events coherence and to explain how they are related to each other. This shows the importance he placed upon causal connections within a plot. For Aristotle, a plot is a causal explanation of events.

Similarly, Forster identifies two features, story and plot, that he thinks essential for a narrative to be intelligible (Forster 1927). He suggests a distinction between a story and a plot, identifying a story as 'a narrative of events arranged in their time sequence' (1927, 42). He says that the story is good only if it makes the audience wonder what happens next. He identifies a plot as follows:

> A plot is also a narrative of events, the emphasis falling on causality. "The king died and then the queen died," is a story. "The king died and then the queen died of grief" is a plot. ... Consider the death of the queen. If it is in a story we say "and then?" If it is in a plot we ask "why?" That is the fundamental difference between these two aspects of the novel (1927, 87).

He draws a distinction between a story as a mere chronology of events and a plot that involves an explanation of the events and the connections between them. It is the story, for Forster, which gives a narrative its temporal order.

Forster seems to agree with Aristotle, saying: 'We must not ask 'And then?' too often. If the time-sequence is pursued one second too far it leads us into quite another country' (1927, 49). This implies that Forster thinks that for a *story* to be complete there should be no further questions for the audience as to what happened next. I do not think that it is clear that a *plot* has to end with no further questions left for the audience; all that seems essential of a plot is that all the events have been explained. A plot can leave open questions of what happens next, without the audience feeling that it is incomplete. I think that Forster's argument is that a narrative must have both story and plot, so a complete narrative must leave no further questions for the audience, either as to what happens next or as to why something in the narrative occurred.

However, this does not always seem to be the case, because the point of a story can be to raise questions and leave the audience speculating. For example, the film *The Italian Job* is not criticised for having no ending, or for being incomplete, even though the audience does not know if the robbers get away with it, managing to rescue both themselves and the gold. This is a cliff-hanger in the most literal of senses, but does not seem to be an ending in the sense required by Forster and Aristotle. They require that there are no further questions, yet in *The Italian Job* many questions remain: How could they get the gold out of the bus? Does the bus fall over the edge? Do they end up eating spaghetti in an Italian jail?

Such cliff-hanger endings appear to be less common in novels. Novels also appear more likely to be criticised for having no ending or being incomplete. For example, *Suite Francaise* will forever remain incomplete because its author died in Auschwitz after

completing only two sections of the novel. The narrative is set in the year that France fell to the Nazis and follows the journeys of several inhabitants of Paris as they escape from Paris, moving on to depict the lives of those in a village under occupation. Again, many questions remain: Will Benoit escape to Paris? Will the Michauds help him? Does Lucile find happiness? Does the German Lieutenant return from the war with Russia? This narrative is not complete and, although praised, critics argue that the work would be better if circumstances had allowed it to be finished. This provides a contrast with the example of *The Italian Job*, where there is no critical view that the film would be better if the remaining questions were answered. Is a cliff-hanger of the type found in films an ending of sorts? Perhaps it is not necessary that a complete narrative have a story with an end, in the sense of there being no further questions as to what happened next. The fact that a story does not have an ending, in the sense that there are questions as to what happened next, does not mean that the story is not a narrative: the events within the narrative still explain one another.

It is further argued that a story has to be arranged in time sequence. However, I think that a story can gain its merit from the fact that it is *not* told in time sequence. Forster appears to argue that the story aspect of a narrative is only any good if the events are arranged in time sequence. This does not seem to be true. For example, in *The Time Traveller's Wife* Audrey Niffenegger weaves intertwining narratives together. Clare's narrative follows in normal linear fashion, but the narrative of Henry, a time traveller, leaps backwards and forwards in time and both narratives are joined together to form a narrative of their relationship. The novel starts in 1991, when Henry first meets Clare, but then moves onto the first time Clare met Henry when he was time travelling back in 1977. The narrative then continues to weave back and forth through time, yet retains coherence.[43]

2. Narrative connections

The attempts to define single necessary conditions for a narrative outlined in the previous section are problematic. Carroll instead argues that narrative connections are an essential part of what it is for something to be a narrative (2001, 119). He thinks that there are five necessary conditions of a narrative connection (2001, 126):

1) At least two events/states of affairs are represented.

2) The narrative is forward-looking

---

[43] Later in this chapter I will consider whether it is necessary for it to be possible to extract the time sequence rather than that the story is told in time sequence.

3) There is one unified subject

4) The narrative is temporally-ordered

5) The earlier events are at least causally necessary conditions (or contributions towards a causally necessary condition) of the later events.

Does a narrative have to be about events? First, Carroll defines a narrative connection as necessarily involving at least two events or states of affairs. He gives an example to illustrate the difference: a statement that is not a narrative is "There was an old lady who lived in a shoe", but the following statement is a narrative: "There was an old lady who lived in a shoe that was very small, so she went looking for a boot". This does appear to put some limits upon what can count as a narrative. The former sentence is a description of a character and her circumstance and does not seem to be a narrative. The latter sentence describes how things were and the relations between that state of affairs and the actions of the individual. This gives us a more detailed description of the person in relation to things that they have done. I agree with Carroll here that a narrative must involve at least two events or states of affairs as otherwise it is just a description.

Secondly, he argues that narratives have to be forward-looking. It is not clear exactly what is meant by 'forward-looking' but I take it to mean that the narrative has to progress forwards through time, ending at a later point than it started. He allows that the presentation of a narrative may involve using devices such as a flashback, but argues that the time-ordering should be retrievable and that the flashbacks or flash-forwards themselves have to be told in a forward-looking manner (2001, 126). For example, the film *Eternal Sunshine of the Spotless Mind* commences in the present, with what appears a normal boy-meets-girl story. Whilst waiting in his car outside her apartment, Joel is questioned as to what he is doing there by an apparent stranger. This event is explained through a flashback to the recent past, which itself contains flashbacks further into the past to explain the whole history of the event, before returning to the present. This is an example of when, although the events are related out of time order, it would be possible to extract the temporal ordering and the flashback episodes themselves are told forwards.

Carroll continues to consider what makes a series of events or states of affairs into a narrative. He does not think that a list of separate events can be a narrative, so thirdly claims that a narrative must be about a unified subject: 'the events and/or states of affairs must be *connected*; they cannot simply be a list of *disconnected* events and/or states of

affairs' (2001, 120). He argues that even if a single subject is clear, a statement is not necessarily a narrative. He gives the example of: 'The President talked to his adviser; the President ate a piece of cheese; the President jogged; the President waved to reporters'. He argues that this is not a narrative because a narrative needs a temporal order, whether explicitly stated or contextually apparent (2001, 120). I agree that this is not a narrative, as it is a list of descriptions of the President, with no plot making links between the descriptions. It is not entirely clear what is meant by a 'unified subject'. Carroll gives the example of a person, the President, but I take him to mean any subject, including more complex examples such as the Second World War.

Fourthly, Carroll continues to argue that temporal ordering is a necessary but not sufficient condition for a narrative. He gives the example of a temporally-ordered chronicle: 'First the President talked to his adviser, then he ate a piece of cheese; then he jogged; and, finally, he waved to the reporters' (2001, 121).[44] It is less clear that a time-ordered chronicle is not a narrative; what is missing? In this example, the descriptions are now linked in a time order, but again is still not a narrative. It may describe what the President did today, but it does not tell us, either implicitly or explicitly, why any one event is related to the others. Carroll suggests that a narrative does not have to be told in time sequence, just that it should be possible to extract the temporal ordering (2001, 126).

This argument appears stronger than Forster's claim that a narrative has to be arranged in its time sequence. Carroll's idea is that a narrative may be constructed from several episodes that are not themselves in time sequence, but those episodes must themselves be told forwards. Although these episodes are not presented in time ordered sequence, they could be rearranged into such a sequence. For Carroll, it is a necessary condition of a narrative that we are able to extract the time-ordered events, yet this is not sufficient to identify a narrative, as chronicles also list time-ordered events. He gives the example of an explanation of a lost battle: 'the battle was lost for want of a horse and the horse was wanting for lack of a horseshoe' and argues that this is just an explanation, of the form 'x at t3 because of y at t2 because of z at t1' (2001, 126). He argues that this is not a narrative because it is not forward-looking. He says 'King Philip could find no shoe for his horse and could not ride into battle and as a result the battle was lost' is a forward-looking narrative of the loss of the battle (2001, 126).

---

[44] Carroll is here using 'chronicle' as a descriptor of a non-narrative timeline. Although 'narrative' is often taken to be a synonym for 'chronicle' – for example, *The Chronicles of Narnia* is a narrative – here 'chronicle' is being used to label the non-narrative.

Finally, Carroll continues to consider what may separate a narrative from a chronicle, arguing that the answer lies with the causal links between events in the narrative (2001, 122). He argues that the causal connection between the events in the narrative is not sufficient, as he does not think that the later events in most narratives are entailed by the earlier events. He gives the example of reading 'that a thief enters a bank and robs it; in the next scene, as he exits the bank, he is apprehended by the police whom we subsequently learn have been watching him all along' (2001, 122). This example is a narrative, but robbing the bank does not causally entail that the robber will be caught. He argues that the connection between events in a narrative is that the earlier event is a causally necessary condition for the subsequent event (2001, 123). He refers back to the robbery example to illustrate his point; if the robber had not robbed the bank, then he would not have been arrested, but robbing the bank did not guarantee that he would be arrested. The robbing of the bank is a causally necessary, but not sufficient, condition of the arrest. He argues that different events can share the same causally necessary conditions, so a causally necessary condition is compatible with a range of different events (2001, 124).

Carroll gives the following example to demonstrate that a causally necessary condition is needed: 'Aristarchus hypothesized the heliocentric theory thereby anticipating Copernicus' discovery by many centuries' (2001, 125). In this example, there is a unified subject, it is time-ordered, but there is only a necessary connection, not a *causally* necessary connection, between the events. He does not think that the example should be considered a narrative because there is 'no line of influence' between the events; hence he thinks that this example is of a coincidence rather than a narrative (2001, 125). He argues that even if the example is reworded as 'Aristarchus hypothesized the heliocentric system and then centuries later Copernicus discovered it again', although the relevance of mentioning the two events is made clear, it is not a narrative because there is no causal connection between them (2001, 125). He implies that the narrative has to explicitly make a connection between two events or states of affairs. Perhaps it is possible that a narrator or audience could themselves imagine the causal connection between them, linking the events or states of affairs together? In this case, the narrator or audience would be making the links themselves, hence creating a narrative, creating a plot, out of the two events that are randomly selected. However, this suggests that the original telling of the events was not a narrative, as the narrator or audience is doing work to create a narrative about the events. This suggests that an implied or explicit causal connection is necessary for something to be a narrative. I think he is wrong about this. A good writer could make implicit suggestions

to the audience as to how to complete the narrative. The members of the audience will create links (and different audience members will create different links) and therefore might make what is written or told a narrative.

Carroll concludes that narrative is a common form of explanation (2001, 128). Because he argues that the events in a narrative have to be causally related, he thinks that the causal relations give narratives their explanatory power. He thinks that for narratives to be explanatory, they must be connected to causation and not just be a list of temporally ordered events. The only reason, he thinks, for citing an earlier event in a narrative is that it plays some role in the causal network of the event we wish to explain (2001, 128).

There are three problems with Carroll's attempt to define the necessary conditions of narratives. First, it is not clear from Carroll's argument as to why we are able to work out the time-ordering of several episodes that are told forwards, yet cannot extract the time-ordering of an episode that is not told forwards and fit that into its time-ordered place amongst the episodes. Consider again his non-narrative explanation of a lost battle: 'the battle was lost for want of a horse and the horse was wanting for lack of a horseshoe'. I think that the forward-looking narrative that can be extracted from this example is 'there was no horseshoe, which meant that there was no horse available, which caused the loss of the battle'. Although the explanation is told backwards, it is possible to extract the time-ordered events. Telling an episode within a story backwards does not seem to exclude the whole story from being a narrative; perhaps it is just a more complicated narrative.

Secondly, demanding that narratives be forward-looking seems to limit their use as explanations. Explanations look back to the past to explain why something occurred, so usually start with the event and explain backwards, rather than starting at some point in the past and explaining forwards to the event being explained. As argued above, it is not clear that flashbacks have to be told forwards, as the time-ordering can be extracted even if they are told backwards. It is also possible for whole narratives to be told backwards. In *Time's Arrow* Martin Amis tells the story of a Nazi doctor backwards in time. Not only does this narrative make sense, but it derives its whole meaning from the fact that it is narrated backwards. Although it is possible to extract the forward-looking time sequence from this narrative, if the story was narrated forwards, it would be an entirely different narrative. It does not appear necessary for all narratives to be forward-looking. Even when a story is narrated backwards using forward-looking flashbacks, the story can gain additional interest

from being told in such a way. For example, the film *Memento* gains its interest from the fact that it is told backwards using flashbacks.

Thirdly, is it an essential characteristic of a narrative that it has a temporally ordered story aspect? Temporal ordering may not be sufficient to identify a narrative, but it is also not clear that temporal ordering is necessary for a narrative. Is it difficult to see how a temporal ordering of events could be extracted in the example of *The Time Traveller's Wife*. Here, the occurrence of the first meeting of the central characters occurs both in 1977 and 2000. It is possible to create a temporal ordering of the events for each individual, but it does not seem possible for their joint narrative. In fact, the narrative of their relationship only makes sense by being told out of order. It does not appear necessary for all narratives that it be possible to extract a temporally ordered sequence of events.

In summary, I do not agree that it is necessary for a narrative to be forward-looking. Or that the narrative has to be temporally-ordered. Yet I think that it is possible to define some general conditions for a narrative as they will usually connect events together causally. I agree with Carroll that a narrative usually involves at least two events/states of affairs; otherwise it is just a description. The earlier events are normally at least causally necessary conditions (or contributions towards a causally necessary condition) of the later events. And there is normally one unified subject. However, these conditions are not singly sufficient for a narrative. If a single subject is clear, a statement is not necessarily a narrative as a chronicle can have one subject. If at least two events/states of affairs are represented the statement is not necessarily a narrative as a chronicle lists more than two events. If the earlier events are at least causally necessary conditions of the later events the statement is not necessarily a narrative because a causal explanation does not have to take narrative form e.g. 'I ate the ice cream because I wanted to'.

Are the conditions outlined above jointly sufficient for a narrative?[45] Are there examples that meet all these conditions but are not a narrative? I agree with Velleman that Carroll's conditions are not jointly sufficient for all narratives because an explanation can meet all of his criteria without being presented in a narrative (2003, 1-25). He questions whether narrative has some explanatory force over and above that of causal explanation. He aims to 'show that something other than causality or probability serves the function of differentiating narrative from other genres and endowing it with its peculiar explanatory

---

[45] Carroll himself does not specify what proportion of narrative connections must be possessed for something to be a narrative (2001, 119).

force' (2003, 4). He interprets Carroll, as I have above, as arguing that the explanatory power of a narrative derives from the causal relations between the events in the narrative. Velleman thinks that this reduces the explanatory force to information that would be explanatory if presented in a non-narrative format. He does not deny that narratives relate the causal relations between events, but wishes to show that the narrative form itself has explanatory force, not just its causal content. He does not think that the causal content of a narrative distinguishes it from any other type of explanation. For example, 'the cue hit the white ball, which then hit the red ball, which caused the red ball to move' describes some causal relations, but is not a narrative.

Thus far, I have identified themes that occur in many things that we are happy to call narratives, but that also occur in others that we would perhaps not, such as descriptions or chronicles. There appear to be examples that are borderline cases, such as Lynch's film *Mulholland Drive*. Is the narrative structure there, but just in an obscure way, or is one simply not present? Some people may regard such films as narratives and others may not. It appears that there can be disagreement as to what is a narrative and that some works have a stronger narrative structure than others; could narratives have degrees of strength rather than simply possessing a certain set of characteristics?[46]

3. Are narratives necessarily emotive?

In this section I will continue to develop my argument that we cannot identify the necessary and sufficient conditions of narratives. I shall consider Velleman's argument that a narrative is necessarily emotive and that this is sufficient to identify a narrative.[47] I argue that he does not identify the necessary and sufficient conditions of narrativity. Instead, I argue that the connection with emotion is a general feature of most narratives. Most narratives will give us some sort of emotional understanding, whether of the protagonists or of the narrator. My argument, in agreement with Velleman, is that this understanding would not be available from a non-narrative chronology or chronicle of events. This will be important to my later argument that a narrative explanation of character is necessary for us to have a full understanding of character.

---

[46] For example, Gregory Currie argues that although within a narrative we find causal relations between events, 'we need something more subtle than an all-or-nothing category of narrative; we need something that admits of degree' (2006, 310).

[47] The next chapter will consider his more general scepticism about the knowledge and understanding that can be gained from a narrative.

Velleman's positive argument is that a narrative is necessarily emotive and that this is sufficient to identify a narrative. He draws a distinction between Carroll's example of 'Aristarchus hypothesized the heliocentric system and then centuries later Copernicus discovered it again' and Aristotle's example of 'the statue of Mitys in Argos killed the man who was responsible for Mitys' death by falling on top of him as he was looking at it' (2003, 6). He says that both these stories involve causally unrelated events, but that the latter pair of events seems more of a story, and not just because of an imagined causal connection, such as an avenging force, as he thinks it still a story even if read as an accidental death. He thinks 'the crucial difference between these examples...is that in Aristotle's the sequence of events completes an emotional cadence in the audience' (2003, 6). He thinks that although the events have no causal connection, the audience has 'emotional resolution' because the murderer suffers suitable consequence. He thinks that Carroll's example has no such emotional resolution, but that it might be possible to tell a story about Aristarchus and Copernicus, despite there being no causal connection, if we could find 'something that they might mean to an audience in emotional terms' (2003, 6).

I suggested earlier that the idea behind a complete plot as characterised by Aristotle and Forster is that there should be no further questions for the audience as to why the first event happened and as to what happened next. I take Velleman to be providing an explanation of why the audience has no further questions as to why the first event happened and as to what happened next. He thinks that a plot is complete when there is no emotion that precedes the beginning or follows the end: 'the story begins with the circumstances that initiate some affect, or sequence of affects, and it ends when that emotional sequence is in some way brought to a close' (2003, 14). He takes as an example Aristotle's requirement that a tragedy cause the audience to feel fear and pity. Velleman thinks that this requires that the plot leads the audience through an emotional sequence, starting 'from an essentially initiatory emotion to an essentially conclusory one', where fear would mark the beginning of the plot and pity the end (2003, 15).

Velleman puts forward two premises:
1) 'That the understanding provided by narrative should be attributable to the nature of narrative itself – to that in virtue of which a recounting of events qualifies as a story' (2003, 18).
2) 'That a description of events qualifies as a story in virtue of its power to initiate and resolve an emotional cadence in the audience' (2003, 18).

From these, he asks how a narrative renders events intelligible. He argues that a story enables the audience to assimilate events to 'familiar patterns of how things feel' (2003, 19). He thinks that:

> ...the audience of a story understands the narrated events, first, because it knows how they feel, in the sense that it experiences them as leading it through a natural emotional sequence; and second, because it knows how it feels about them, in the sense that it arrives at a stable attitude toward them overall...The audience may or may not understand how the narrated events came about, but it understands what they mean – what they mean, that is, to the audience itself, in emotional terms (2003, 19).

For Velleman, a narrative does not merely explain how events occurred, but gives us an understanding of how to feel about the events.

However, it is not clear that Velleman's argument is successful. Paisley Livingston has recently raised four objections against Velleman's view (Livingston 2009). The first objection Livingston raises against Velleman attacks the claim that a narrative is *necessarily emotive*. He asks whether there are some instances of narrativity that do not reliably arouse and resolve emotions (2009, 32). He thinks that 'some narratives, and good ones too, are not designed to arouse affect, but are meant to have a predominantly if not exclusively cognitive or intellectual impact' (2009, 32). He gives the example of Aesop's fables as falling into this category. Carroll also objects to Velleman on the grounds that 'one may tell an affectless narrative...I see no reason to deny that such a story could be a narrative, even though it arouses no emotion. Many scientific and historical narratives are dispassionate, but they are still narratives' (2007, 14). He argues that the completion of an emotional cadence is not necessary for a narrative.

The second objection raised by Livingston again challenges Velleman's argument that being emotive is necessary for narrativity because a narrative does not always give rise to the same emotion in different audiences. He says 'if a film leaves some audiences cold, consistently gives rise to derisive laughter among others, but reliably stirs up pity, fear, and like emotions among naive spectators, does it have narrativity?' (2009, 32). Ismay Barwell proposes a similar counter-example to Velleman's view because she thinks that some narratives arouse ambivalent emotions; for example, it could be appropriate to feel a mixture of sadness and gladness (2009, 55).

The third objection raised by Livingston challenges Velleman's argument that narratives resolve an 'emotional cadence' in the audience because Livingston thinks narratives do not

necessarily resolve or bring closure to the emotion that they raise: 'Even when a narrative has successfully been designed to stir up specific emotions, it need not resolve the climatic feeling or affect, but may instead be designed to leave the members of the audience in a state of intense arousal, such as feelings of political indignation or anger' (2009, 32). Barwell also challenges Velleman's argument that a narrative is distinctive because it offers 'emotional understanding of the events represented' and that a good narrative organises events into a whole that makes emotional sense (2009, 54). She suggests the following counter-example: 'A reader...might feel increasing dismay for which, again, the outcome provides no resolution. Often the conclusion of a story arouses contempt, anger, sadness, or fear rather than dispelling emotions aroused in response to earlier events' (2009, 55). She continues to say 'how events feel, the sequence of emotions experienced as responses to narrative representations, need not have a tripartite form whose last stage is a resolution' (2009, 55).

The final objection Livingston raises attacks the view that being emotive is *sufficient* for narrativity. He asks whether some non-narrative discourses arouse and resolve emotions (2009, 32). He gives the example of 'a standard notation of the moves in a particular chess game would not usually be taken as a narrative, but it could have a marked emotional cadence for chess cognoscenti' (2009, 32).[48]

Barwell's solution to the objections she raises against Velleman is to argue that narratives give us reasons for evaluative judgements of the events rather than a causal explanation or emotional understanding of the events. She argues that narratives allow us to make evaluative sense of the events portrayed rather than emotional sense. This proposal, however, seems to be open to similar objections to those that face Velleman. It is not clear that a narrative necessarily provides reasons for evaluating the events portrayed. One can imagine a light-hearted novel where one is not invited to evaluate the events but just to be entertained. For example, in a P.G. Wodehouse novel we do not really need to 'evaluate' the events. In one sense we could be said to evaluate the events, but that sense is only a low level one, satisfied by more or less every narrative and representation of events that there is. 'Evaluate' here does not merely mean 'judge' or 'comprehend', but has more to do with judging *and* deciding whether something good or bad, say, has happened. We can of course discern whether something good or bad has happened if we choose to do so, but the point of such a light-hearted narrative is that it can entertain without us having to worry

---

[48]Carroll also argues that the completion of an emotional cadence is not sufficient for a narrative (2007, 14).

too much about whether something good or bad has happened.[49] It is also not clear that providing reasons for evaluating events is sufficient for narrativity, as non-narrative discourses can invite evaluation of the events. Using the same example Livingston uses against Velleman, a standard notation of the moves in a particular chess game would not usually be taken as a narrative, but it could invite evaluation of the moves. Barwell's own counter-example to Velleman's view, that some narratives arouse ambivalent emotions, also appears to hold against her view, as it could be appropriate to evaluate the same event in a variety of ways. It is neither necessary nor sufficient for narrativity that it give us reason for evaluating the events.

I agree with objections one, two and four above, that being emotive is neither necessary nor sufficient for narrativity. However, this is not to say that emotion is unimportant to narratives. Goldie takes a similar approach when he argues that emotional import or cadence is not a necessary or sufficient condition for something to be a narrative. As outlined above, he thinks that *in general* a narrative reveals coherence, meaningfulness and emotional import (Goldie 2009, 98). He thinks coherence is important because it causally connects representations of actions, events and states of affairs. It is characteristic of most narratives, but not of chronicles of events or lists of propositions. He thinks meaningfulness is characteristic of most narratives because the narrator presents the actions, events and states of affairs in such a way that the audience can understand the thoughts, feelings, and actions of the protagonists. He gives the example of a child being told the story of Goldilocks and the Three Bears in such a way that the child understands the perspective of Goldilocks, imagining what it might be like to experience what Goldilocks experienced. Goldie calls this the internal perspective of a narrative. He thinks that in such cases we are centrally imagining the perspective of one of the people internal to the narrative (2003a, 57-8).[50] He thinks emotional import is characteristic of narratives because the telling of the story reveals the narrator's evaluation of, and emotional response to, what happened (2004a, 159).

Goldie argues that the emotional import of a narrative derives from a different sort of imaginative engagement with the narrative, what he calls the external perspective or point of view (2004a, 159). This is the perspective of the narrator, so is external to the narrative, he says, because it is not part of the content, but is a perspective on the content. He thinks

---

[49] Thanks to Simon Kirchin for this example.
[50] Goldie borrows the terminology of 'central' and 'acentral' imagination from Richard Wollheim (Wollheim 1984).

that when we engage imaginatively with the perspective of the narrator we are acentrally imagining his perspective (2003a, 57-8). From this perspective, we are not imagining what it might be like to experience what Goldilocks experiences, but are engaging with the narrative in the way encouraged by the narrator. This type of engagement with the narrative may give rise to different emotions than imagining what it is like to be one of the protagonists. He argues that:

> ...the external perspective is always there, in spite of sometimes seeming evanescent, always shaping and colouring the narrative, and indicating the narrator's own evaluation and emotional response to what happened. A narrative, when it is narrated, can thus reveal internal and external perspectives, and the narrative will thereby invite the reader or audience to share in these perspectives and to respond emotionally to the events as they are represented in the narrative (2003a, 55).

How the narrator links the different events together reveals how the narrator feels about the events, so a narrative can tell us something about the emotions and values of the narrator.

Goldie gives a further example of a narrative:

> Two strangers, a woman and a man, were the only passengers on a tube train late one night. The man had been drinking all evening, and, as sometimes happens to us when we have had too much to drink, suddenly, and for no good reason, he became very angry with the woman, and got up and began to tell her in no uncertain terms just what he thought of her disapproving glances (2003b, 209).

In this example, he explains that the lower-level perspective is that of the man, inviting the audience to identify with his anger. He argues that the higher-level perspective invites the audience to share in the acentral perspective of the narrator, inviting the audience to share the narrator's response of compassion for the woman. He argues that a successful narrative will give the audience an understanding of the emotions of both the character in the narrative (anger) and the narrator (compassion). He argues that the audience does not itself necessarily have to feel the same emotions as the narrator; for example, the narrative might invite compassion, but the audience may respond in another way (2003b, 209-10).

The argument is that it is *generally* characteristic of a narrative that the way the events are structured and linked together gives the audience an understanding of how to feel about the events.[51] At most it must be necessary for a narrative to be *capable* of producing an emotional response, but this does not mean that it will produce an emotional response in a particular individual. Matthew Kieran argues that we do not need to imagine the emotions of the protagonists or the emotional perspective of the narrator to understand a narrative

---

[51] The claim that a narrative gives the audience an understanding of how to feel about the events will be further elaborated in the next chapter.

(2003, 69-87). He thinks that 'at best one can simulate emotional states consistent with certain character traits, but then there are many emotions consistent with a particular trait, and any given emotional state is consistent with many character traits' (2003, 75). He continues to argue that some character traits are not connected with an emotional state, giving the example of being considerate without having any particular feeling. He gives the example of Darcy in *Pride and Prejudice*, saying that we do not have to imagine having the character traits attributed to Darcy to understand his character traits, even though understanding the character traits is essential to our understanding of the novel. He argues that a narrative leads us to view a character in a certain way but that we do not have to imaginatively identify with the character to understand this. He does not deny that imaginative identification is possible and can be meaningful, just that it is not necessary to feel anything to understand a narrative.

I agree with Kieran that although a narrative leads us to view a character in a certain way, we do not have to imaginatively identify with the character to understand this; it is not necessary for the audience to feel anything to understand a narrative. For example, I can understand the narrative 'the statue of Mitys in Argos killed the man who was responsible for Mitys' death by falling on top of him as he was looking at it' without feeling anything, because I can understand what the narrator is trying to make me feel, even if he is not successful in this. Causal stories or chronologies of events do not necessarily reveal anything about the emotional perspective of the narrator and of the protagonists. For the most part, narrative seems to causally connect two or more events together in such a way that the audience learns something about the emotions of the characters and the narrator, even if they are not felt by the audience. Emotion is important for narrativity because a narrative will usually reveal something about how the narrator feels or how the audience should feel that is lacking from a chronology of events and this information is in addition to the casual connections between the events.

Consider this example of Goldie's to illustrate: an individual relates a narrative about securing some funding for a research project, intending to relate the story in an undistorted fashion. However, he argues that it is possible for the narrator to unintentionally reveal his vanity: 'my action – my act of narration – is expressive of reasons that can appropriately be classified as boastful and vain, even though it was no part of my intention to boast in telling you about what happened' (2004a, 162). The argument is that this understanding of the narrator as boastful and vain would not be gained from a chronology of events. This understanding can be gained only from the relation of events in a narrative. So, I accept

objections one, two and four, that being emotive is neither necessary nor sufficient for narrativity, yet argue that emotional understanding is still important as it is characteristic of most narratives, but not chronologies of events. What the writers are failing to consider is that 'narrative' is a concept that can cover many different types of artwork or human creation. To expect that one could define, explicitly and precisely, the necessary and sufficient conditions for something being a narrative seems optimistic. However, we can still spot important trends.

Consider again objection three. This stated that the emotion provoked by the narrative is not necessarily resolved. In such cases, one can argue that these *are* narratives, but that they are not very *good* narratives because the emotional import is unclear. And further those narratives with no clear emotional resolution are bad narratives or incomplete narratives, rather than not being narratives at all. Barwell ties the completion of an emotional cadence to the necessity of a narrative having a beginning, middle and end, but I do not think this is central to Velleman's view (2009, 49). As argued above, it is not a necessary condition of a narrative that it has an end. Similarly, it cannot be a necessary condition of narrativity that it resolves an emotional cadence, again because the position would not be able to accommodate cliff-hangers. To return to the example of *The Italian Job* the emotional sequence for the audience is not complete, as we do not know how to feel at the end because we do not know what happens. At best, we are left with a feeling of anticipation or intrigue as to how events turned out. It does not appear that the lack of a conclusion to the emotional sequence means that the film is not a narrative. There is a question here as to what it is for an emotional sequence to be brought to a close; can anticipation (or uncertainty, or disquiet, or suspense, or disbelief, or...) be the end of an emotional sequence? It is unclear what Velleman may think about this, as if we are thinking of narrative explanations, the end of the narrative is specified by the action or event that we are trying to explain. Perhaps it is not necessary that a complete narrative have a story with an end, in the sense of the emotional sequence for the audience being complete.

Or do we find films such as *The Italian Job* satisfactory because we enjoy, as the audience, being able to imagine our own endings to the events portrayed and thus feel what we want to feel about the imagined ending? We imaginatively complete the narrative ourselves and an incomplete narrative is a good one if it gives us the resources to do this; not just any incomplete narrative will do. The fact that a story does not have an ending, in the sense of

emotional closure, does not mean that the story is not a narrative: the events within the narrative still explain one another and suggest how we should feel about them. Is emotional closure rather a sign of a good narrative rather than necessary? Aristotle's example of 'the statue of Mitys in Argos killed the man who was responsible for Mitys' death by falling on top of him as he was looking at it' is regularly referred to as an example of a narrative, but it is hard to feel any emotion without knowing more about Mitys and the man responsible for his death. The implication is that the audience should feel glad that the statue killed the man, but again, this is the audience completing the narrative. So perhaps the narrative is good because it gives us the resources to imagine this, but perhaps the narrative would be better if we knew more about Mitys and the man responsible for his death so that we know how to feel.

In summary, it is not possible to identify the necessary and sufficient conditions that successfully identify all narratives. My view differs from that of Velleman because I agree that there is a problem with the claim that a narrative is necessarily emotive because it is possible to tell a narrative that does not arouse an emotion. Further, it cannot be a necessary condition of narrativity that it resolves an emotional cadence because the position would not be able to accommodate cliff-hangers. I also agree with Livingston's objection that being emotive is not sufficient for narrativity because there are some non-narrative discourses that arouse and resolve emotions (2009, 32). However, it is possible to identify what is characteristic of most narratives; for example, they reveal causal connections between events that explain how the event happened and they reveal how the narrator wants the audience to feel about those events. The narrative structure itself generally gives rise to an understanding of the events narrated over and above the causal relationships between the events. This further understanding is based upon emotional understanding of the events. Consideration of these issues has raised some interesting questions about the extent of imagination and interpretation done by the audience in understanding a narrative and by the narrator in telling the story. If narratives are subject to imagination and interpretation, do they give rise to any knowledge at all or are they always biased and distorted? These questions will be considered in more detail in the next chapter.

Conclusion

In conclusion, I have argued that a narrative is more than a list of time-ordered events, as the events have a causal connection that means one event can explain why a later event

occurs. A narrative can gain its meaning when we are able to emotionally engage with the protagonists or from our engagement with the perspective of the narrator. However, it is not necessary to share in the emotions of the characters or narrator because the audience can understand what the narrator is trying to make them feel, even if he is not successful in this. Narratives are causal explanations, but provide further understanding over and above the causal connections they illuminate. Narratives achieve this by relating to an emotional response; they reveal something about the emotions of the characters and the narrator. A narrative will generally reveal something about how the narrator feels or how the audience should feel that is lacking from a chronology of events and this information is in addition to the casual connections between the events. Is it possible to identify what is characteristic of all and only of narratives? I think that some things are generally characteristic of narratives but not only narratives; for example, all narratives involve implied or explicit causal connections, but so can other types of explanation.

Introduction

The aim of this chapter is to establish what type of knowledge we can obtain from historical narratives. Establishing that we can obtain knowledge from a historical narrative will be important for my argument in Chapter Ten that narratives are essential for understanding of character traits. Historical narratives are real and therefore character traits are real, not fictional. This chapter addresses the question of what we can learn from narratives, concentrating upon narratives about real events, not fictional narratives. This is important because it is common practice to explain historical events and human action by using a story or narrative.

In the first section of this chapter I will defend the idea introduced at the end of the last chapter that narratives are essential for a certain sort of emotional understanding. I there agreed with Velleman that the narrative structure itself generally gives rise to an understanding of the events narrated over and above the causal relationships between the events. I agreed with him that narratives do not render events intelligible just on the basis of the causal connections they illuminate. However, I disagreed with him that a narrative is necessarily emotive, claiming that the narrative can give rise to emotional understanding of the events without necessarily evoking a particular emotion in the audience. In this chapter I will take an example of a biographical narrative about a historical character to support my claims and argue that this example better supports my claims than those of Doris.

In the second section of this chapter I respond to the general sceptical claim that narratives do not provide us with knowledge because they simplify things too much and find connections where perhaps there are none. I consider whether there are any specific problems raised by autobiographical narratives and argue that narratives can be evaluated in terms of truth and falsity. I conclude by arguing that narratives are truth-apt and that they can provide the audience with knowledge.

1. Historical narratives and emotional understanding

This section will argue that Doris's use of the example of Schindler does not provide support for his view that people have a collection of narrow local traits that cannot often be

evaluatively integrated into wider traits. Instead, I will argue that the example of Schindler offers a useful case study that demonstrates that a detailed narrative about an individual can reveal consistent, reliable traits. Even if an individual is not fully virtuous, possessing the unity of all the virtues, the sorts of localised trait behaviours identified by Doris can be evaluatively integrated under thick trait terms such as 'courage' and 'generosity'. Further, I will argue that such a narrative is necessary for a full explanation of the events and that without this full explanation, a reliable evaluation of the events is not possible, so narrative is also necessary for evaluation of the events.

Doris appeals to the example of Schindler to support his view that character is fragmented. He says: 'in fact, some rescuers exhibited strong inconsistencies. Schindler saved over a thousand Jews in Poland from deportation and murder, but he was also a manipulative, hard-drinking, and womanising war profiteer who did not particularly distinguish himself either before or after the war' (2002, 59). Doris takes this inconsistency to suggest that Schindler had no global character trait of altruism, supporting his view that an individual may be consistently altruistic in some specific set of circumstances but not others and that we are used to accepting such fragmentation.

However, I would argue that the example of Schindler does not quite establish the conclusion desired by Doris and can in fact be used as an example for the opposing view that there *are* global character traits. Consider these extracts from Thomas Keneally's account of how he gathered the evidence for *Schindler's Ark* and from *Schindler's Ark* itself:

> He paused before a *chocolatier's* store, where there was an enormous heart-shaped box of chocolates in the window. This was, clearly, not a box for sale – it was the *choclatier's* trademark. But Oskar, with characteristic exuberance could not see the distinction. "I would like to get that for dear Mrs Gosch," he said (Keneally 2009, 32-33).

And:

> A number of Schindler's friends would claim later – though it is not possible to prove it – that Oskar had gone looking for the dispossessed family at their lodgings in Podgórze and had given them a sum close to fifty thousand zloty in compensation. With this sum, it is said, the Nussbaums bought themselves an escape to Yugoslavia. And however good a light this rumoured action places Oskar in, it has to be said that it is probable. Fifty thousand zloty signified substantial dissent, but there would be other similar acts of dissent by Oskar before Christmas. Some friends would in fact come to say that generosity was a disease in Oskar, a frantic thing, one of his passions. He would tip taxi drivers twice the fare on the meter. This has to be said, too – that he thought the Reich housing authorities were unjust and told Stern

so, not when the régime got into trouble, but even in that, its sweetest autumn (Keneally 1982, 57).

These extracts do not reveal that Schindler had numerous specific local traits, such as buying-chocolates-in-post-war-Paris-generous, compensation-to-Jews-who-have-their-property-confiscated-generous, and tipping-taxi-drivers-generous. These examples, along with countless others recounted in *Schindler's Ark*, reveal that Schindler was consistently and reliably generous throughout his life. I think that this demonstrates that we can evaluatively integrate the local character traits in a way denied by Doris. As a general rule, there may be no evaluative integration between say buying-chocolates-in-post-war-Paris-generous and tipping-taxi-drivers-generous. However, in a narrative about a particular individual it is clear where there is evaluative integration; generosity in one area, such as chocolate buying, does imply generosity in other areas, such as tipping taxi drivers. So Schindler's types of generous behaviour are evaluatively consistent.

Doris claims that the only type of evaluative integration possible is the grouping of local traits under thin evaluative headings such as 'good' and 'bad'. He says: 'it does not imply that the entirety of a person's behaviour cannot merit an on balance "evaluative score"; Josef Mengele is far in the red, and Schindler far in the black, despite the fact that their behaviour was not evaluatively consistent' (2002, 115).[52] Instead, I argue that local character traits can be grouped under thick evaluative terms, such as 'generous', 'courageous' etc., which begin to resemble remarkably the attributions of global character traits that Doris argues do not often exist. My argument is that this evaluative integration is discovered by narrating a story about an individual over a long period of time. It will not be revealed by observing an individual's behaviour at, for example, one summer camp, as in Newcomb's study into introversion and extraversion (Newcomb 1929). This is not to say that the construction of such a narrative is *easy*; I will return to this issue in section two below.

What the narrative of *Schindler's Ark* in fact reveals is that Schindler was not fully virtuous, not that he possessed a character comprising numerous narrow traits. The narrative reveals that he had as many negative global character traits as positive. For example there are as many stories about his reliable, consistent womanising as there are of his reliable, consistent generosity:

---

[52] Josef Mengele is the infamous doctor who conducted horrific experiments on the child prisoners of Auschwitz.

The Brinnlitz Oskar was still the Oskar old Emaila hands remembered. A *bon vivant*, a man of wild habits. Mandel and Pfefferberg, at the end of their shift and overheated from working on the pipe fittings for the steam, visited a water tank high up near the workshop ceiling. Ladders and a catwalk took them to it. The water was warm up there, and once you climbed in, you could not be seen from the floor. Dragging themselves up, the two welders were amazed to find the tub already taken. Oskar floated, naked and enormous. A blonde SS girl, the one Regina Horowitz had bribed with a brooch, her naked breasts buoyant at the surface, shared the water with him. Oskar became aware of them, looked up at them frankly. To him sexual shame was a concept, something like existentialism, very worthy but hard to grasp. Stripped, the welders noticed, the girl was delicious.

They apologised and left, shaking their heads, expelling their breath, laughing like schoolboys. Above their heads, Oskar dallied like Zeus (Keneally 1982, 364).[53]

And:

Victoria Klonowska, a Polish secretary, was the beauty of Oskar's front office, and he immediately began a long affair with her. Ingrid, his German mistress, must have known, as surely as Emilie Schindler knew about Ingrid. For Oskar would never be a surreptitious lover. He had a childlike sexual frankness. It wasn't that he boasted. It was that he never saw any need to lie, to creep into hotels by the back stairs, to knock furtively on any girl's door in the small hours. Since Oskar would not seriously try to tell his women lies, their options were reduced; traditional lovers' arguments were not possible (Keneally 1982, 69).

The actions and events portrayed in the narrative show Schindler to act consistently and reliably in accord with both virtuous and non-virtuous traits. However, these are traits defined in a global sense, so this narrative does not support the claim that his character is fragmented at a local level. The only fragmentation is at the global level; he does not possess the unity of all and only all virtuous traits and cannot be described as fully virtuous. As Keneally comments, it is this moral ambiguity that is interesting: 'Some people have always been troubled by Oskar's ambiguity. To me it was from the start the whole point of the tale' (Keneally 2009, 26).

I stated above that evaluative integration is discovered by narrating a story about an individual. Doris may respond by questioning why explanation of the attribution of a character trait has to take narrative form. The inadequacy of conditional statements or belief-desire explanations of behaviour suggested by Doris's view have been a recurrent theme throughout the previous chapters. Presumably Doris would not deny that there needs to be some further explanation as to how someone comes to have a certain localised trait or why they have a particular belief-desire combination. What he may challenge is why this further explanation has to take narrative form, rather than that of some other sort of causal

---

[53] 'Emalia' is the informal name of Schindler's factory.

explanation. And, further, why does this narrative resolve the empirical problems of attributing consistent, reliable, evaluatively integrated traits that he thinks is best resolved by the notion of local traits?

As introduced at the end of the previous chapter, I am supportive of the argument that we gain a certain sort of emotional knowledge from a historical narrative that we would not gain from other sorts of explanation. Lists of causal chains of events give us only a limited understanding of what took place. Compare Keneally's narrative of events:

> Next, the brothers Danziger, who cracked a metal press one Friday. Honest bemused men, half skilled, looking up with staring shtetl eyes from the machine they had just loudly shattered. The Herr Direktor was away on business and someone – a factory spy, Oskar would always say – denounced the Danzigers to the administration in Plaszów. The brothers were taken from Emalia and their hanging advertised in the next morning's rollcall in Plaszów. *Tonight*, (it was announced) *the people of Plaszów will witness the execution of two saboteurs*. What of course qualified the Danzigers above all for execution was their orthodox aura.
>
> Oskar returned from his business trip to Sosnowiec at three o'clock on Saturday afternoon, three hours before the promised execution. News of the sentence was waiting on his desk. He drove out through the suburbs to Plaszów at once, taking cognac with him and some fine kielbasa sausage. He parked by the Administration Block of Plaszów and found Goeth in his office. He was pleased not to have to rouse the commandant from an afternoon nap. No one knows the extent of the deal that was struck in Goeth's office that afternoon, in that office akin to Torquemada's, where Goeth had had ringbolts attached to the wall from which people could be hanged for discipline or instruction. It is hard to believe, though, that Amon was satisfied simply with cognac and sausage. In any case, his concern for the integrity of the Reich's metal presses was soothed by the interview, and at six o'clock, the hour of their execution, the Danziger brothers returned to the back seat of Oskar's plush limousine to the sweet squalor of Emalia (Keneally 1982, 235-6).

With my re-telling as a chronology of events:

> The Danziger brothers cracked the metal press; then this was reported to the administration in Plaszów; then the brothers were taken from Emalia; then their hanging was advertised in the next morning's rollcall in Plaszów; then Oskar returned from his business trip to Sosnowiec; then he read news of the sentence on his desk; then he drove out to Plaszów with some cognac and sausage; then he found Goeth in his office; then he struck a deal with Goeth; then the brothers were driven back to Emalia.

Drawing upon the work of Velleman and Goldie outlined in the previous chapter, my argument is that the latter causal chain of events gives us some understanding of what happened but does not give us any understanding of how to feel about the events. Here I claimed that chronologies of events do not reveal anything about the emotional perspective

159

of the narrator, the emotional perspectives of the protagonists, or what the narrator is trying to get the audience to feel. The second example above provides support for my view. This lists the events in the order that they happened and suggests causal links between the events; for example, the brother were taken from Emalia because someone reported that they had broken a machine. It gives us information about the events that happened, but gives the reader no sense of what the narrator feels about the events or what the narrator wants the audience to feel about the events. Although all the events listed are factually correct the audience does not know how to feel about the events because neither the emotional perspective of the narrator nor the emotional perspective of Schindler is apparent.

In the last chapter, I claimed that, for the most part, narratives seem to causally connect two or more events together in such a way that the audience learns something about the emotions of the characters and the narrator, even if those emotions are not felt by the audience. The first extract above provides support for my view. The same events as in the second example are described, but in this extract we, the audience, get a sense that we are supposed to admire these events. Further, we get a sense that the narrator feels admiration of Oskar's actions in rescuing the brothers and that he feels disgust at the actions of the spy and Amon, along with general disgust at the whole regime. We get a sense of these emotions because the events are portrayed in a narrative, with the events linked in such a way to reveal these emotions.[54]

As argued in the previous chapter, I do not think that narratives are necessarily emotive, in the sense that a narrative will necessarily give rise to a certain emotion in the audience. Nonetheless, emotion is important for narrativity because a narrative will usually reveal something about how the narrator feels, or how the protagonist feels, or how the audience should feel that is lacking from a chronology of events. This additional information is needed for the events to be fully explained for the audience. If the audience does not know how to feel about the events it is unclear that they fully understand the events. As Misak argues, some knowledge can be gained only through experience, but by narrating these experiences to others, we may be able to pass on some of that knowledge (2008, 615-6). I would argue that knowledge of how to feel about the experience falls into this category. For an event to be fully explained the explanation needs to reveal information about the

---

[54] Later in this chapter I will consider whether this additional knowledge can be assessed as true or false.

emotions involved. Hence the further explanation of events has to take narrative form as this is the best way of passing on this information to others.

Without a full narrative explanation of the events that contains the relevant information to gain knowledge of how to feel about the events, it is difficult to see how someone can reliably evaluate the events. The level of detail contained in a narrative is important to aid evaluation of the events. Compare again the above examples. The bare statement of the facts does not give the audience enough information to evaluate the events reliably; the contextual details and emotional perspective of the narrator in the first extract provide more information about the complexity of the events. In summary, narratives are important for both explanation and evaluation of events. The world is complex and it is only through a detailed narrative that it is possible to explain fully the events to others, in the sense that to fully understand and evaluate the events, they also need to know how to feel about the events. Further, it is only through study of such narratives that reliable, consistent, integrated traits can be revealed. In the next section I will consider what the difference is between fictional narratives and historical narratives; if we are to gain knowledge from historical narratives they must be capable of truth and falsity.

2. *Autobiographical narratives, knowledge and truth*

Thus far, this chapter has considered narratives in general, focusing on an example from history. The attention in this section will continue to focus on such narratives about real events to investigate whether there is anything specific that separates these narratives from fictional narratives and whether real-life narratives raise any particular problems. This section first concentrates upon a certain species of real-life narrative, autobiographical narrative about one's own history, to raise some questions about the extent to which a narrative is subject to interpretation and distortion. This highlights a general scepticism as to whether a narrative can provide knowledge and understanding. This section concludes with an argument that narratives are truth-apt and can provide knowledge.

Autobiographical narratives about one's own history are different from other narratives in that the narrator and protagonist is the same person. Such autobiographical narratives will be important in the next chapter in terms of understanding and evaluating ourselves, as well as in justifying and explaining our actions to others. Misak argues that autobiographical narratives are important because she thinks that some features of events

161

or states of affairs can be understood only if they have been experienced. She argues that autobiographical narratives are distinct from fictional narratives because the narrator is also a participant in the narrative. She thinks that if an individual learns from these experiences then they can narrate these experiences to others to give them some understanding of what has been experienced (2008, 615). However, that the narration of a story involves interpretation and motivation leads her to raise a concern: 'We all know that in recounting our own experience we can mislead, embellish, and even self-deceive' (2008, 616).[55] Can autobiographical narratives be true and objective?

Goldie identifies some further complexities with autobiographical narratives. He thinks that narration is a type of intentional action performed for reasons and that these reasons are important in explaining why an individual relates '*this* particular narrative at *this* particular time in *this* particular way' (2004a, 157). He thinks that therefore the audience has 'a double interpretive task of considerable complexity in understanding an autobiographical narrative' because they have to interpret both the content of what is narrated and the narrator's reasons in performing the act of narration (2004a, 157). He identifies a further complication in that the audience themselves will not be free from bias. This bias can distort the audience's understanding of the narrative even if the narrator himself tells an undistorted story about what happened. He argues that the audience must take into account its own tendency to see things in a certain way when interpreting a narrative (2004a, 163). He thinks that his observations explain why convergence and divergence of evaluation occur in a clinical setting. He thinks that divergence between the clinician's and the patient's evaluation of the patient's autobiographical narrative does not arise from 'disagreement about 'the facts' – about the truth of the content of the narrative' (2004a, 164). He argues that the divergence arises from the differing evaluations of the narrative by the clinician and the patient (2004a, 164). He does not think that this means there is no objectivity when it comes to understanding narratives because he equates the objectivity of narratives with having an appropriate evaluation of what happened.

Can an autobiographical narrative be true and objective given that it is reliant upon evaluation and interpretation? One solution is suggested by Misak who thinks that the experiences or insights reported in an autobiographical narrative can be the subject of scrutiny (2008, 625). She does not think that the fact we have to scrutinise the evidence we

---

[55] Barbara Tversky draws upon social scientific evidence showing that people admit to omitting or exaggerating information in stories about what happened to them (Tversky 2004). She explains that these alterations can change the individual's memory and become the individual's memory of how the events occurred.

gather from narratives distinguishes narrative evidence from other kinds of evidence. She gives this example: 'the newspapers are full of examples of how pressure to come up with striking results can lead to the falsification or the withholding of data, and the fact that trials are often paid for by pharmaceutical companies can amplify this pressure' (2008, 627). Hence even empirical results need scrutiny. She argues that 'both our own reports of experience and those of others are open to scrutiny' (2008, 627-8). She thinks that narratives do not only conflict on the grounds of exaggeration, omission, or self-deception, but they can also conflict because one of the narrators *has got things wrong* (2008, 629). She does not think that the fact narrative evidence has to be scrutinised discredits it as a form of evidence; it is, after all, still used in the courtroom. Some methods she identifies of scrutinising autobiographical narrative include internal coherence, consistency with other evidence, simplicity, and explanatory power (2008, 630). She also thinks that our background beliefs are important in our evaluations of autobiographical narratives. We will have beliefs about the reliability of sources: she gives the example that 'I may distrust the narratives of one friend, as I know her to be a terrible exaggerator, especially when she is talking about the talents of her children, yet trust the narratives of another friend, whom I have never known to exaggerate' (2008, 630).

For Goldie, the difference between factual autobiographical narratives and fictional narratives is that the former refer to how things were (2004a, 160-1). Under his approach the content of a narrative is true if it corresponds to the facts. If an individual narrates something, that narration is true if things were as the individual narrates (2004a, 161). He gives the example of: '"My doctor told me that I had VD", as uttered by Robert, is true just if Robert's doctor told him that he had VD' (2004a, 161). He argues that we should think of a narrative as a collection of propositions rather than as a sequence of events. By thinking of a narrative as a collection of propositions, 'the metaphysical notions of reference and truth have application in factual autobiographical narrative, just as they do in, for example, scientific explanation, whereas they have no application in fiction' (2004a, 161).[56]

---

[56] The central focus of this thesis is not to defend a correspondence theory of truth. Richard Kirkham provides an outline of different types of correspondence theory (Kirkham 2001, Chapter 4). He identifies that the commonality between such theories is 'the claim that one of the individually necessary and jointly sufficient conditions for the truth of a belief (proposition, or whatever) is that the very fact the belief is a belief in (or the very fact that the proposition expresses) must obtain' (2001, 139). A commitment to a correspondence theory of truth raises questions such as 'what do the facts look like?' and 'are there such things as thick evaluative facts?' In this case, propositions containing thick evaluative terms such as courageous would have to have thick facts to which they should correspond. Kirkham identifies three main objections to correspondence theories of truth. First he identifies objections raised as to what is the correct

These concerns as to whether an autobiographical narrative can be true and objective seem to apply to historical narratives in general, as all evaluations of factual narratives are open to interpretation and distortion. Velleman raises such general scepticism about narratives (Velleman 2003). He argues that the understanding and explanation of events that is provided by a narrative is only knowledge of how the audience *feels* about the events portrayed in the narrative. He says that having 'sorted out its feelings towards events, the audience mistakenly feels that it has sorted out the events themselves' (2003, 20). For Velleman, narratives are problematic because they lead the audience to believe that having decided how to feel about the events portrayed, it has knowledge of the events portrayed. He thinks that the audience is mistaken if it thinks it has such knowledge.[57]

Such a view is in opposition to that of Goldie and Misak, who consider a narrative to be true when it relates events as they happened; i.e. the narrative corresponds with the facts. This opposing view does not think that the truth of a narrative can be evaluated in such a way. Instead, a historical narrative, in terms of which events have been selected and combined, should be evaluated using criterion common with fictional narratives.[58] An example of such a view is that of Hayden White (White 1973 and 1978). White argues that a historian organises events into narrative patterns familiar in literature, such as the pattern of a tragedy. He argues that these narrative patterns are familiar to the audience, so narrating historical events in accord with such a pattern means that the audience can understand the story of the events. He does not think that the past events themselves have the structure that is evident in the narrative. So, under this view, a historical narrative does not simply relate what happened, but is a construction that adds meaning to the events portrayed. In a later paper, Lamarque neatly summarises such positions as stating that a narrative creates the events, in the sense that there are no plot-like structures of events independent of narrative (2004, 400).

truth bearer; whether it is beliefs, propositions, statements, etc. Secondly, he identifies objections that are raised as to whether part of reality, such as facts, situations, states of affairs, etc., can be correspondents to the truth bearers. Thirdly he identifies objections that there is no relation between reality and truth bearers, or that the theory fails to explain the relation. However, he is optimistic that these common objections do not pose a problem for correspondence theories: he says that 'none were found to be particularly telling' (2001, 140). What is important for my account is that fictional accounts are different from narratives about historical events. The difference is that historical narratives can be true or false depending upon how accurately they express what happened; fictional narratives are not subject to such criteria.

[57] Livingston also identifies this general scepticism in Velleman's argument (2009, 31).

[58] See Carroll 1990 for a summary of this type of view.

However, I agree with Lamarque that it is possible to resist this view.[59] There is a problem of truth and reference faced by rival accounts, such as White's, that argue that character traits are wholly constituted by a narrative. Lamarque thinks that narrative is a representation of something because if it were merely the narrative '...history would be reduced to a *product* rather than a subject of narrative' (2004, 393). He argues that the sentences in historical narratives have truth-values and that the names denote according to ordinary criteria. He quotes a passage from Gettysberg to illustrate his argument:

> Robert E Lee planned to cross the Potomac himself this morning (June 25th 1863). Just before and just after he did so, he penned letters to Jefferson Davis. In the first (written 'opposite Williamsport'), Lee worried that his thrust into the North might prove too successful (Trudeau 2002, 71 cited by Lamarque 2004, 397).

Lamarque argues that the names 'Lee', 'Potomac', 'Jefferson Davis' used in this narrative refer in a straightforward manner. He thinks that the truth of the narrative can be established by contemporary official records, which include copies of Lee's letters. He thinks that although there may be philosophical problems with the ideas of reference and truth, 'it seems clear that no additional problems arise from the appearance of names and assertions in narratives of this kind' (2004, 397).

Lamarque argues that although a narrative can distort the events it portrays, this does not prevent it giving rise to knowledge (2004, 398). He thinks that historical narratives involve selection and ordering of the facts by the narrator, so the narrator is going to emphasise some events over others, make causal relations between the events and impose some kind of ending on the story. He argues that narrators have to narrate from some point of view or other and that however impartial they might intend to be, the narration will involve some interpretation and evaluation. He thinks that at a minimum, the narrator will have chosen some facts as being more important than others (2004, 398). There is a clear similarity here with Goldie's concern that a narrator may distort an autobiographical narrative. Goldie and Lamarque share a concern that even though someone intends to relate a truthful and undistorted narrative about what happened, he might unintentionally distort the events portrayed.

---

[59] Carroll also suggests that we can resist this view because he thinks that although narratives are in some sense constructed, they can provide accurate information. He argues that narratives can provide accurate knowledge about past events, including background conditions, causes and effects, social contexts, the logic of situations, practical deliberations and the resulting actions (Carroll 1990).

Despite the fact that a narrative can be distorting and that it involves interpretation and evaluation, Lamarque does not agree with White that historical narratives are a type of fiction, and gives three reasons for dissenting. First, as outlined above, he thinks that in a historical narrative, names denote and sentences have truth-values according to ordinary criteria, so historical narratives are not subject to different 'fictional' truth criteria. Secondly, he thinks that the Gettysburg passage he cites gives an example of a historical narrative that has no distorting perspective, so historical narratives are not necessarily distorting. Thirdly, he does not think that it is distorting to give historical significance to an event, even if that event was not recognised as significant at the time (2004, 398). He thinks that a narrator will have to select which facts to portray in the historical narrative, but how these facts are selected is subject to evaluation. He says: 'Historians are rightly taken to task where lack of objectivity, neglect of detail, and misleading emphasis are evident' (2004, 399). He argues that assigning significance to historical events is not to create the very events themselves. He says: 'The structures that historians impose on clusters of events, like the Middle Ages, are markers of significance not inventions of fact' (2004, 400). Again, there is a clear similarity here with the case of autobiographical narratives. For example, Misak argues that experiences or insights portrayed in autobiographical narratives can be the subject of scrutiny (2008, 625).

A tension is emerging between considering narratives as set of propositions that refer and have truth conditions and considering narratives as interpretive constructions that confer meaning on the events portrayed. Mark Day thinks that there is so much potential causal information in considering an individual event that it can be difficult to ascertain what will be most important (Day 2008). He thinks that a historian who has the benefit of hindsight can more easily make judgements about the relative importance of different causes (2008, 109). He says: 'Judgements of importance are implicit in all narratives, particularly the sort of complex narratives typical in history. In narrating historians select, arrange and establish significant themes, and connect causal processes' (2008, 109). He thinks that often it is assumed that the witness to the event knows best which causes are significant. However, he thinks that 'far from being at the end of a chain of unreliable whispers', the historian has multiple pieces of evidence so can often gain a more accurate idea of what happened than those who were witnesses of the event being explained (2008, 203). Under this view, the historian is not merely providing a causal explanation, but is making a judgement. He continues to argue that even if the narrative-historical explanation is causal, it is not necessarily the causal connections within the narrative that are informative. He says:

'While "the cause of Archduke Ferdinand's death was the shot that killed Archduke Ferdinand" is quite true, and locates a causal relationship, it is spectacularly uninformative, and in particular moves us nowhere towards making sense of my actions and motives' (2008, 126). This explains how we move from the basic facts that are clearly truth-apt, to linking them together to form a narrative.

Day considers how a historian decides what events are explanatory and so should be included in a narrative. He argues that the historian must use all available evidence and that something should be included in a historical narrative if it made a difference to what is being explained; if it made no difference to that to be explained, it should not be included. He does not think that something either simply makes a difference or it does not; it is a matter of degree as to whether an event is more explanatory because it makes more of a difference (2008, 161). He says: 'An event which is more important, relative to what is being explained, is therefore more worthy of inclusion in a historical account. Thus the selection and omission in a historical account can be held to task, according to the requirement that it should mention the important events, and omit the unimportant' (2008, 161).[60]

Day introduces W.H. Walsh's idea of a 'colligatory concept', for example 'the Cold War', to see if this resolves the problems of reference (Walsh 1970). Day defines a colligatory concept as a concept that refers to a complex of events that extends over time (2008, 172). He thinks that these colligatory concepts refer to a narrative development using a single phrase. Walsh gives an example of a colligatory concept as 'the expansion of Nazi Germany'. Day explains that the use of this concept explains a more specific part of the overall event, for example, the invasion of the Rhineland, by placing it in the wider event. In such cases, he argues that the specific is not explained by a general concept, but by placing the specific event under the wider context (2008, 172). However, Day questions whether the colligatory concept adds anything to our understanding. He thinks that in this example, it is the Nazis' plan to expand that does the explaining rather than the colligatory concept.

---

[60] Such argument lends weight to the distinction between historical narratives and fiction. For example, Carroll thinks that discovery of new information may lead to more detailed historical narratives that revise and improve on earlier accounts in the same way that scientific theories are adjusted as more evidence comes to light (Carroll 1990).

Day continues to consider the view that narrative-historical writing is interpretive. He casts doubt on the claim that historical narratives are true because they correspond to reality. He defines such a view as stating that '...the past existed, our historical accounts are about that past, and those histories are true to the extent that they correspond to past facts' (2008, 188). He explains that a correspondence theory of truth states that a statement is true if it corresponds to the facts. This is the sort of view that Carroll and Goldie suggest; that a proposition in the narrative is true if it refers to the past fact as it was.

However, Day poses a challenge to this view, in that it needs '...to give an account of facts such that they are sufficiently similar to statements to permit correspondence, and yet sufficiently different so as not to be simply the projection of language onto the world' (2008, 189-90). He argues that this problem arises because the historical narrative can be true or false 'in a way that goes beyond the truth or falsity of their individual components' (2008, 196). By this, he means that the events that the historian chooses to narrate and the way that he puts them together are also capable of truth and falsity. He says: 'the challenge is to find something in the real past to which that holistic truth can correspond' (2008, 195). He explains his challenge by considering whether narratives have holistic truth. He argues that the set of statements within the narrative are representative of the person or event as a whole, so have holistic truth. He thinks that if all the statements are critical, then a negative overall conclusion is implied, but if all the statements are positive, then a positive overall conclusion is implied. However, he thinks it possible that the implication can be false even if the statements in the narrative are true. He gives the example of a person who is worthy of criticism, when a historical narrative about the person implies that they are not (2008, 196).

Can the problem that there are no plot-like structures of events in the past, so there is nothing to which the narrative as a whole can refer, so narratives cannot be subject to normal rules of reference and truth, be resolved? I think that perhaps there may be *a* narrative that, when considered *in isolation*, could contain true statements with an overall false implication. However, throughout this discussion it has been emphasised that the historian should use all available evidence when constructing his historical narrative. A narrative where all the statements are true yet the overall implication false would not be corroborated by other evidence. Such a narrative is false because the historian has not used all available evidence and there does not need to be a plot-like structure existing in the past for this holistic truth to correspond to. As Goldie argues we do not so often disagree about

the truth of the content of the narrative, but tend to disagree with how the narrator has evaluated that content (2004a, 164). That there is disagreement over the implication or evaluation of a narrative does not imply that there is no true or false narrative portrayal of the events; perhaps we just do not have enough conclusive evidence to fully understand the events portrayed.

The narrative has a meaning over and above that of its parts. Consider the following passage, returning us to the case of Schindler:

> Oskar took part in a similar transaction when the Gestapo raided the apartment of a forger and discovered, among other false documents completed or near completed, a set of Aryan papers for a family called the Wohlfeilers, mother, father, three adolescent children, all of them workers at Schindler's camp. Two Gestapo men therefore came to Lipowa Street to collect the family for an interrogation which would lead, through Montelupich prison, to that grim hill fort. Three hours after entering Oskar's office both men left Deutsche Email Fabrik, reeling on the stairs, beaming with the temporary bonhomie of cognac and, for all anyone knew, of a pay-off. The confiscated papers now lay on Oskar's desk and he picked them up and put them in the fire (Keneally 1982, 235)[61].

The events portrayed are:

1. The Gestapo raided the apartment of a forger and found a set of Ayran papers for a family of workers at Schindler's camp.
2. Two Gestapo men came to Lipowa Street to collect the family for interrogation, followed by death.
3. The men entered Oskar's office.
4. Three hours later the men left, appearing drunk.
5. Oskar burned the confiscated papers.

And the overall implication of the narrative is that Schindler should be viewed in a positive light. The narrative implies that Schindler bribed the two Gestapo men, but that this should be viewed as an admirable action to save the lives of five of his workers.

It is difficult to see how the five events listed in the narrative could be true, but the overall implication false, as suggested by Day. My argument is that a historical narrative can be true or false. This is not to say that the entire narrative has to correspond to a pattern of events in the past. However, the historian cannot impose any old pattern on the events of the past, finding meaning and connections where there are none, because his account has to be based upon a multitude of evidence from the past. If the account is *not* based upon

---

[61] The 'grim hill fort' is where prisoners were taken to be shot.

multiple sources, then it cannot be said to be a reliable account and we should question its truth. However, where a historian has corroborated a sequence of events with evidence from multiple sources, both the individual events and the linking together of those events can be said to be a true representation of what occurred.

Support for my view is provided by Keneally in the preface to *Schindler's Ark*, where he outlines the numerous different sources he used to inform his account:

> This account of Oskar's astonishing history is based in the first place on interviews with fifty Schindler survivors from seven nations...It is enriched by a visit...to locations which figure prominently...But the narrative depends also on documentary and other information supplied by those few wartime associates of Oskar's who can still be reached, as well as by the large body of his postwar friends. Many of the testimonies regarding Oskar...further enriched the record, as did written testimonies from private sources and a body of Schindler papers and letters... (Keneally 1982, 13).

Keneally's later account of how he gathered the evidence for *Schindler's Ark* demonstrates why he used multiple different sources to write his account; the fact that individual sources can be unreliable (Keneally 2009). He reports: 'Now Bejski, a scholar, a man of serious intent and more than a little worried about the projected book, warned me against accepting all of Poldek's exuberant tales unless they were corroborated by other prisoners' (2009, 162). Poldeck was the man who introduced the story to Keneally, but he could not rely solely on Poldeck's account. In *Searching for Schindler* Keneally provides a detailed account of how he went about gathering the required evidence to write an accurate narrative. Early in his research he comments: 'Dr Rosleigh had a professional gravitas Poldeck lacked, and to hear the same stories emerging from both kinds of men, so far apart geographically, impressed me greatly' (2009, 51). He refers on several occasions to the importance of verification of facts: 'I knew, too, that things that were said by one interviewee would have to be matched or weighed against what the historic record said, against context and the memories of other former *Schindlerjuden*' (2009, 60-1). He further protected his account from the unreliable narratives of others: 'When three or four survivors told the same story, though, and the story was supported by documents...well, I believed it had just about earned its place in the ultimate record' (2009, 87).

On my view, the same level of detailed research should be undertaken in the case of autobiographical accounts. Our own memories of events in which we were involved are no more reliable than those of others. So, again, autobiographical accounts should be supported by interviews with witnesses, visits to locations, documentary evidence such as photos and videos, and written sources such as letters and diaries. This is exactly the same

types of evidence that Keneally refers to in his explanation of his sources for his account of Schindler. So, although the author and the protagonist are the same person in the case of an autobiographical narrative, the supporting evidence should be no less rigorously scrutinised to assess the validity of the account. An autobiographical narrative is still a type of historical narrative and therefore capable of truth and falsity.

Keneally's account highlights the difficulty of constructing a narrative account of an individual and the level of research that goes into uncovering the events of the past. I agree that it is not easy to tell an accurate narrative about a person, be that yourself or a third party. It is not often that we will have access to the required level of evidence to tell an accurate story without making the effort to do some research into the person and his past. Hence I agree with Doris's claim that in everyday life we tend to attribute global character traits on too little evidence (2002, 101). He argues that the range of evidence on which a character trait attribution is commonly based is limited, both in terms of number of instances of behaviour observed and in the range of circumstances in which the behaviour is observed. However, I disagree that this is because people 'presuppose the existence of character structures that actual people do not very often possess' (Doris 2002, 6). The difficulties are grounded in the complexity of the data required to construct an accurate narrative, but difficulty in gathering the evidence does not mean that the character structures do not exist.

A longitudinal study, such as that provided by Keneally of Schindler, demonstrates how such a narrative can be constructed. It is not the scientific method proposed by Doris. Such a method necessitates the observer watching numerous behaviours of a particular individual and noting whether or not that behaviour was in accordance with a certain trait. That observer then uses that evidence to draw conclusions about the robustness and stability of the character of the observed subject. However, the example cited demonstrates how it is possible to use multiple sources to narrate the story of an individual's life and draw conclusions about the robust, stable character traits that the individual reliably displays. It is of course too demanding that one individual will have access to all the necessary primary sources of information. But we do not have to construct these narratives in isolation from others; instead we pool the secondary evidence of many to draw an accurate picture of the individual, removing the bias of the individual observer.

In summary, I think there are three conclusions that can be drawn about historical narratives. First, historical narratives are not a type of fiction because they are factual accounts about what happened in the past. We do not need to commit to an anti-realist position that claims that it is the narrative that gives meaning to the events rather than the events themselves being meaningful. Second, because they are factual accounts, the sentences in historical narratives are true or false according to normal criteria, not the sort of criteria that may be relevant when thinking about fiction. Similarly, the names in a historical narrative refer in the same way as other names. Third, the difficulty the historian faces in selecting which facts are significant and evaluation of those facts does not mean that there is no true way of presenting the facts. An event was historically significant, to a greater or lesser degree, whether or not a narrative is constructed about it. How the historian emphasises this event and the justification for his evaluation of the significance of the event is open to scrutiny. Giving an incorrect account of the significance of events can make the narrative false even if the factual statements made are true. Such a claim does not commit the realist to the existence of a narrative structure to reality. The truth or falsity of the narrative interpretation of the events depends instead on corroboration of that interpretation by other factual evidence.

Conclusion

I have argued that such historical narratives can provide us with knowledge. I considered the specific cases of autobiographical narratives and argued that factual autobiographical narratives are different from fictional narratives on two grounds. First, in the former the narrator is also a participant in the narrative and secondly, the former refer to how things were. I considered how narratives can be evaluated in terms of truth and falsity, drawing three conclusions about historical narratives. First, historical narratives are not a type of fiction because they are factual accounts about what happened in the past. Second, because they are factual accounts, the sentences in historical narratives are true or false according to normal criteria, not the sort of criteria that may be relevant when thinking about fiction. Third, the difficulty the historian faces in selecting which facts are significant and evaluation of those facts does not mean that there is no truth of the matter.

In the next chapter I will draw upon these conclusions to develop an account of narrative attributions of character. I will develop the idea that sentences attributing character traits

are summaries of past behaviour and those statements of character traits are best defined as summaries of historical narratives.

Introduction

In Chapter Two I proposed that individual actions add up to general dispositions because character trait attributions are of narrative form. At the end of Chapter Six I proposed that sentences attributing character traits are summaries of past behaviour rather than conditional statements. This chapter develops my narrative account of character traits. I argue that character traits are not constituted by a narrative, but that this does not lead to the conclusion that narrative is inessential to understanding character traits. My argument depends upon a development of an account of narrative whereby narrative is not constitutive of character traits i.e. the narrative is about something, namely character traits, which have independent existence. I argue that that narrative is important for understanding of character traits because of the knowledge and understanding that is derived from a narrative that could not be derived from any other means. This account provides an alternative to the conception of character trait attributions as conditional statements.

In this chapter I focus upon developing an alternative argument that gives narrative central importance in the understanding of character trait attributions. In the first section I consider the different features of attributions of character traits. I argue that the evaluative and explanatory features are more important for moral evaluation than the descriptive and predictive. In the second section I concentrate on the argument that character traits are not constituted by the narrative but rather the narrative is about independently existing character traits. In Chapter Two I proposed that a character trait would not exist over and above the narrative that may be told about it. Here I consider an objection to that view, arguing that we should consider a narrative to be *about* character traits rather than *constitutive* of character traits. I argue that a narrative about character is necessary to give us an emotional understanding of that trait that could not be derived from another source. In the third section I argue that this account of character trait attributions as historical narratives allows that evaluation of acts is in one sense primary, yet maintains that virtuous character traits overall have greater intrinsic value. In the fourth section I argue that this account is preferable to the account of character trait attributions as localised conditional statements, as put forward by Doris. In the fifth section I answer three potential objections to this account. First, I reject the claim that it is committed to a type of anti-realism about

character traits. Secondly, it could be claimed that the narrative account is psychologically unrealistic, but I provide evidence to the contrary. Thirdly, I argue that real-life narratives can give us the required level of detail about character, rejecting the claim that only fictional narratives provide an adequate account of character. I conclude that my account of character trait attributions as historical narratives provides a genuine alternative to the conditional account and that my account better coheres with our normal moral practice.

1. The features of attributions of character traits

This section will argue that the use of character traits to explain behaviour is of central importance for moral evaluation. In Chapter Four I highlighted four features of attributions of character identified by Doris: character traits are supposedly descriptive, evaluative, provide explanations of behaviour and can be used to predict what people will do (2002, 15). By a descriptive use, I mean that we may describe ourselves or others as having a particular trait or set of traits. We may use this description in a predictive way; once we have identified that a person has a particular trait, we may use this information to predict how they will behave. Character traits are also used to explain why people act in certain ways (Goldie 2004b). This is a common device in fiction. For example, in Jane Austen's *Sense and Sensibility* a contrast is drawn between two sisters. Elinor is described as possessing 'a strength of understanding, and coolness of judgement'; in addition, her 'disposition was affectionate, and her feelings were strong; but she knew how to govern them' (2008, 6). These traits are used to *explain*, for example, her actions of advising and counselling her mother. In contrast, Marianne is attributed traits of being 'sensible and clever; but eager in everything: her sorrows, her joys, could have no moderation. She was generous, amiable, interesting: she was everything but prudent' (2008, 6). To take one example, these traits are used to explain her actions following her father's death of giving herself wholly up to her sorrow, 'seeking increase of wretchedness in every reflection that could afford it, and resolved against ever admitting consolation in future' (2008, 6). Attributions of character traits also have an evaluative component, for example when we talk of virtuous and vicious character traits as a way of praising or blaming someone.[62] It is also used in historical narratives. To return to the example of Schindler cited in the previous chapter, in places the narrative explicitly states that he is generous and that this trait explains his behaviour. For example:

---

[62] Not all traits have an evaluative component. For example, extraversion is not commonly used as an evaluative term. However, virtuous or vicious traits certainly have an evaluative component.

The Oskar Schindler who comes down from his office on the frosty mornings of an Aktion to speak to the SS man, to the Ukranian auxiliary, to the Blue Police and to the OD details who would have marched across from Podgórze to escort his nightshift home; the Oskar Schindler who, drinking coffee, rings Wachtmeister Bosko's office near the ghetto and tells some lie about why his nightshift must stay in Lipowa Street this morning – that Oskar Schindler has endangered himself now beyond the limit of cautious business practice. The men of influence who have twice sprung him from prison cannot do it indefinitely even if he is generous to them on their birthdays (1982, 153).

This narrative describes Schindler as 'generous' and this trait is used to explain his confidence in breaking the rules as his generous behaviour protects him from arrest.

When attributing character traits in a descriptive sense we tend to think of particular episodes rather than in terms of a narrative whole. For example, if asked whether a particular person is generous, to answer this question we would tend to think of episodes in his past that display this trait (or display an opposing trait). Similarly, if asked to predict how a particular individual would act in a certain situation we would draw upon relevant past episodes. However, in doing so, we are just picking out the relevant part of a narrative. To attribute a character trait we imply that the trait displayed in the past is the same trait as in the present and the future. Hence descriptions of character traits are not 'episodic' in Galen Strawson's sense of the term (Strawson 2004).[63] By 'episodic' Strawson means a descriptive, empirical thesis that states that not all ordinary human experience is of life as a narrative (2004, 428). He argues that there is not only one good way for people to experience their lives. He thinks that there are people who experience their lives in a non-Narrative, episodic way and that their lives are equally valuable. However, the same character trait has to persist through time; it makes no sense to say that the character trait in the present is different from the character trait displayed in the past. Strawson says with regard to the self: 'what I care about, in so far as I care about myself and my life, is how I am now. The way I am now is profoundly shaped by my past, but it is only the present shaping consequences of the past that matter, not the past as such' (2004, 438). He is happy to say that there is a strong relation between his current self and the other parts of Galen Strawson's (GS's) life; these have emotional significance and give GS certain responsibilities. He however does not identify these other parts of GS's life with his current self. However, if thinking about character, then the past *does* matter, not just because it shapes the present but because my character is partly constituted by my past. The trait does persist through time, so the trait displayed in the past is the same trait as in

---

[63] Note Strawson is talking about a narrative conception of personal identity rather than of character, but some of the issues he raises are relevant to a narrative account of character traits.

the present and the future. This differs from Strawson's account of the self where the self does not persist through time, so it is not necessary to construct a narrative about it to understand it.

As argued in the previous chapters, there is a difference between a causal explanation and a narrative. If any causal chain were defined as a narrative this would be a very weak definition and would lead to the weak claim that there is a causal explanation for any character trait. I am making a stronger claim that to understand and have knowledge of virtuous and vicious character traits requires a narrative explanation, not merely a causal explanation. This is needed because virtuous and vicious character traits have an evaluative element which can be fully understood only if we know how to feel about virtuous and vicious actions and people. As argued in the previous chapter, a narrative will reveal this emotional knowledge whereas a causal explanation will not.

The only way we can fully understand our character is to construct a narrative about it because it is only through constructing a narrative we gain this perspective on how to feel about the actions and events that form patterns of character-based behaviour. An explanation of why we have attributed a certain trait needs to take narrative form because a mere causal chain of events only gives us some understanding of what happened and no understanding of how to feel about the events. In the previous chapter I compared two explanations of the Danziger brothers being saved. This event could be part of an explanation of why Schindler is attributed the trait of kindness, or generosity, or altruism, and it is the narrative explanation of the events that gives us a full understanding of why these actions are relevant to the attribution of, for example, kindness because it reveals how to feel about the events. 'Kindness' has an evaluative component as well as an explanatory component. Hence, as argued in the previous chapter, the narrative explanation is also needed for a reliable evaluation of the character because there is an inter-dependency between how to feel about events and evaluation of those events.

2. A narrative theory of character traits

This section develops my alternative view of understanding character traits and its relation to explanation of actions. In the previous section I argued that narrative is essential for a full understanding of character. However, this is not to say that the narrative creates the character. In the previous chapter I set out my argument that a narrative is a representation of *something*. It may be that in the case of character traits we need to detach our trait from

177

the narrative. Character traits have to exist so that we can evaluate the narratives about them as good or bad, true or false, and thus use them to explain action. For such evaluation to occur, character traits must exist independently of the narrative. If there are no independent character traits, the only types of truth and falsity that accounts of character traits can aspire to are those found in fiction, such as internal coherence. In the case of a character trait it seems that we are telling a story about something rather than arguing that the character trait is the narrative. My argument is supported by the analogy with historical accounts outlined in the previous chapter. I think that this analogy supports the view that character traits are not constituted by a narrative. I claim that this separation does not mean that narratives are insignificant to understanding of character traits. Instead I argue that narratives about character traits gain their significance from the emotional understanding they give us about the events and traits portrayed.

Why should we adopt my account? In Chapter Nine I identified a tension between considering narratives as a set of propositions that refer and have truth conditions and considering narratives as interpretive constructions that give meaning to the events portrayed.[64] Under the latter view, the historical narrative of a character trait is not a representation of past events but constructs character traits by selecting which events to include in the narrative and constructing links between these events. Under such a view a historical narrative of a character trait does not relate events as they happened, so the truth of the account will have to be evaluated in some other way, for example using criteria used for fictional narratives. The narrative creates the character traits because there is no trait-like structure of events that is independent of the narrative.

The analogy drawn with historical accounts in the previous chapter is useful here because any story explaining a character trait will be based upon events and actions of the past. Here I agreed with Lamarque that a historical narrative is a representation of something because if history were the narrative, then it would be the product of the narrative, not the subject (2004, 393). If history were a product of narrative, it could not be evaluated as true or false, but would be more akin to a fictional account. It is clear that we would wish to resist this claim as we normally expect historical accounts to aim at a true representation of the events that occurred. I think that it is unclear that there is any reason for historical accounts of character traits to be different from other historical accounts. We usually expect accounts of character traits to aim at the truth in the same way as other historical

---

[64] See White 1973 and 1978 (81-100) on the similarities between historical accounts and fiction.

narratives and are critical of those who have delusional or false accounts of their character traits. This points towards character traits being a *subject* of narrative in a similar way to other historical artefacts rather than a product of the narrative.

As argued in the previous chapter, historical narratives can be distorting and involve interpretation and evaluation, but this does not mean that historical narratives are a type of fiction. I there argued that this is not necessarily a problem for the truth conditions of historical narratives because the narrated facts are subject to evaluation and corroboration with other sources. In the case of narratives about character traits, such narratives can be evaluated in a similar way. The sentences in an autobiographical or biographical narrative will have truth-values according to ordinary criteria, so as with historical narratives, narratives about character traits are not subject to different 'fictional' truth criteria. Character narratives are similarly not necessarily distorting and, even if they are, the narrated facts about character traits are subject to evaluation. For example, the facts can be corroborated with other evidence such as the accounts of others, photos, videos, diaries, and so on. The advantage of this view that narratives are about character traits rather than constitutive of character traits is that it allows an account of the truth-values of such narratives because it characterises them as factual accounts rather than fictional constructs. So the selection of narrated facts is also important for character, but is objective; to omit all Hitler's bad deeds in an account of his life is to give a false narrative.

In summary, drawing upon my general account of narratives, I think there are three conclusions that can be drawn about narratives of character traits. First, narratives about character traits are not a type of fiction because they are factual accounts about what happened in the past. We do not need to commit to an anti-realist position that claims that it is the narrative that creates character traits from the events and actions rather than character traits existing and influencing the events and actions. Secondly, because they are factual accounts, the sentences in narratives of character traits are true or false according to normal criteria, not the sort of criteria that may be relevant when thinking about fiction. Thirdly, the difficulty the individual faces in selecting which facts are significant and evaluation of those facts does not mean that there is no true way of presenting the facts. An event was historically significant for the development or attribution of a character trait whether or not a narrative is constructed about it. How the individual emphasises this event and the justification for his evaluation of the significance of the event can be scrutinised and evaluated. The truth or falsity of the narrative interpretation of the events depends on

corroboration of that interpretation by other factual evidence. As discussed, there will be matters of emphasis and selection, just as in the case of historical accounts, or film reviews. But, that should not undermine the more fundamental point: just as historical events happen, and a film moves from one scene to another, so character traits exist and can be described.

The remainder of this section develops an account of the significance of narratives about character traits, based upon the general features of narratives that were identified in Chapter Eight. I there argued that although a narrative leads us to view a character in a certain way, we do not have to imaginatively identify with the character and feel his emotions to understand this; it is not necessary for the audience to feel anything to understand a narrative. I argued that, for the most part, narrative causally connects two or more events together in such a way that the audience learns something about the emotions of the characters and the narrator, even if they are not felt by the audience. I further argued that emotion is important for narrativity because a narrative will usually reveal something about how the narrator feels or how the audience should feel that is lacking from a chronology of events. This information is in addition to the casual connections between the events. Emotional understanding is important as it is characteristic of most narratives, whilst it is lacking from non-narrative chronologies of events.

Drawing upon my characterisation of narratives in the previous chapters, it is possible to identify what is characteristic of most narratives.[65] For example, they reveal causal connections between past events that explain how the current event happened and they reveal how the narrator and characters feel about the events, along with what the narrator wants the audience to feel about those events. The defence of the use of narrative in understanding character traits centres upon finding a middle ground between describing narratives of character traits as trivial lists of historic events and complete literary narratives with a unifying theme that construct the character trait. To avoid claims of triviality I claim that the narrative about a character trait generally gives rise to an understanding of the events narrated over and above the causal relationships between the events. This further understanding is based upon emotional understanding of the events.

Doris criticises the view that the notion of character is necessary for the narratives that people use 'to make sense of their lives' (2002, 119). He says, for example, we often say

---

[65] These are not a set of necessary and sufficient conditions. In Chapter Eight I argued that we cannot define such a set of conditions for narratives.

things such as 'early disappointments in his career are what fuel Donald's ruthless ambition, while Angelina's betrayal at the hands of Maxwell is what makes her so slow to trust' (2002, 119). He considers whether clinical psychology provides any argument for the indispensability of character: 'therapeutic transformation is predicated on narrative intelligibility, and narrative intelligibility is predicated on character discourse, so therapeutic transformation is predicated on character discourse' (2002, 119). However, he argues that narrative need not be character driven. He cites Aristotle's characterisation of a tragedy as an example: he tells us in the *Poetics* tragedy 'is essentially an imitation not of persons but of action and life' (*Poetics* 1450a15-25). Doris takes this to show that tragedy essentially involves action and that the character of the people involved is secondary. Further, he thinks that moral narratives do not have to 'be narratives invoking character...; a story might teach fairness, equality, or any number of other central moral ideas without proceeding in the discourse of character' (2002, 126).

Given the general characteristics of narratives that I have identified in the previous chapters, I would agree with Doris that a narrative does not have to be *driven by* the character traits of the people involved. For example, Aristotle's narrative about the statue of Mitys, which is commonly agreed to be an example of a narrative, does not involve strong character development. However, although the character of the author of Mitys' death is not developed, it is implied that this action makes him a bad person, allowing us to understand the story. Daniel Hutto draws attention to the level of sophistication required, when he compares the characters in children's fairy tales to those characters found in more complex literature; he says 'Little Red Riding Hood is no Madame Bovary, to be sure' (2007, 64). The level of character development necessary depends upon the audience and is perhaps a mark of a good narrative rather than characteristic of all narratives. A good narrative will provide enough information for the audience to imagine what it is like for the character to experience the events related, but this will vary from little characterisation in the case of the author of Mitys' death or Little Red Riding Hood, to extensive characterisation in the case of Madame Bovary. However, although a narrative does not have to develop explicitly the characters involved in it, it appears that there has to be *some* characterisation implied by the story for the audience to be able to identify with the protagonists. It is difficult to imagine a story such as that Doris suggests that teaches fairness or equality where the audience does not care about any of the characters involved.

In summary, a narrative has to be about something. Character traits have to exist so that we can evaluate the narratives about them as good or bad, true or false. For such evaluation to occur, character traits must exist independently of the narrative, not be constituted by a narrative. If there are no independent character traits, the only types of truth and falsity that accounts of character traits can aspire to are those found in fiction. Narratives about character traits gain their significance from the emotional understanding they give us about the events portrayed. This claim does not require that the narrative develop the character of the protagonists. All that is required for the narrative to be significant is that it reveals something about the emotions of the protagonists, the narrator and what the narrator wants the audience to feel about those events. This further emotional understanding derives from the narrative structure itself not from a recounting of causal relationships between the events.

3. Character traits and acts

This section turns to the evaluative role of virtuous and vicious character traits. In this section I will expand upon the objection to virtue ethics outlined in Chapter Two that what is primary is the evaluation of local actions and not the evaluation of global traits. I will argue that defining character trait attributions as summaries of past behaviour, or historical narratives, provides a solution to this objection. I will argue that defining character trait attributions as narratives explains both our ordinary evaluations of actions and of dispositions.

In Chapter Two I explained that under Hurka's 'occurrent-state view' the local use of the concepts of virtue and vice is primary and virtuous dispositions are identified as those that give rise to such occurrent virtuous acts and desires (Hurka 2006). To recap, I contrasted this view with the dispositional view; according to which the global use of the concepts of virtue and vice is primary and what are considered virtuous acts and desires derive from these dispositions. Hurka thinks that in everyday use moral concepts refer only to the occurrent states of the perpetrator and not his longer-term traits (2006, 71). He argues that when making an everyday judgement we do not take into account the behaviour of the person at other times when making a judgement about the virtuousness of this particular act. He thinks that when we make everyday global judgements about virtue, such as saying that a person is brave or generous we derive those judgements from our local judgements about the virtuousness of particular acts, desires, and feelings of the person (2006, 74). He

argues that the disposition may have some intrinsic value because, drawing upon Ross, an individual who has a disposition to be unselfish, for example, may be considered more virtuous than an individual who lacks this disposition even if not currently acting upon that disposition (Ross 1939, 291-92 cited in Hurka 2006, 73). However, his claim is that this value is always less than the value of individual virtuous acts (Hurka 2006, 73). He identifies the core disagreement between the two views as being over whether it is dispositions that are primarily good or whether it is the occurrent states.

In Chapter Two I argued against this view on the grounds that it is not clear that the virtuousness or viciousness of acts has conceptual priority. Further I argued that if dispositions have intrinsic as well as instrumental value, it is unclear what role such dispositions are playing in the occurrent-states view. I claimed that in everyday moral discourse we often do want to attribute general traits to people and that we regard these general traits as having value themselves. Finally I argued that it is not clear that the value of the occurrent acts and desires that arise from virtuous dispositions will always be greater than the value of the dispositions themselves.

My narrative account of character traits provides a solution to these problems. First, it provides an account of the link between acts and character traits. Under the narrative view dispositions have the role of contributing to our evaluation of particular acts. The narrative view explains the inter-dependency of acts and dispositions; the individual acts are essential for the disposition to exist in the individual, but the disposition is essential to link the individual acts together into a structured whole. Secondly, it provides an account of why character traits have intrinsic value. According to my account, the narrative explanation of a character trait usually connects two or more events together in such a way that the audience learns something about the emotions of the protagonist(s), or the narrator, or how the narrator wants the audience to feel. This knowledge is lacking from consideration of singular acts. Character traits have to be explained by a narrative rather than some other type of causal explanation because this is the best method of gaining a full understanding of the trait, including knowledge of how to feel about the character that could not be gained from some other type of causal explanation. The argument that character trait attributions are summaries of historical narratives suggests that the character trait has intrinsic value because it facilitates deeper evaluation of the events portrayed because of this additional emotional knowledge that it gives us.

Thirdly, it explains why the value of the overall character traits are greater than that of the individual component acts. As argued in the previous chapter, a narrative about character does more than list the acts done by an individual. Each of these individual acts has a value, but the overall value of the composite is greater. The narrative about character has value over and above the value of its individual components. The events that are chosen and the way that they are put together in the narrative gives the overall narrative a value. If all the statements in a narrative are positive, then a positive overall conclusion is implied that can be greater than the sum of its parts. For example, the same individual having acted generously twice is better than that person having acted generously once. Considering descriptions or explanations of a person's character as a historical narrative demonstrates how the aggregation of particular virtuous actions can give the person a virtue that is greater than the sum of the parts of the narrative. The narrator selects and orders the facts about the person, emphasising some events over others and making causal relations between the events. The narrator will interpret and evaluate the actions or events portrayed in the narrative and draw an overall evaluation of the person. This overall evaluation of the person that is implied by the narrative is different from the evaluation of the individual events or actions that are the content of the narrative. The evaluation of the individual actions or events comes first, but the full value of these events can be appreciated only as part of a narrative. The narrator structures individual actions or events to give an evaluation of the person and this evaluation illuminates the value of the individual acts. The narrative view gives an account of how an occurrent motive that is one of many featuring in a narrative rather than a one-off occurrence can have more value.

In summary, the narrative account suggests that there is an inter-dependency between individual actions and character traits with neither having evaluative priority. A narrative account of character traits will not be independent of an account of particular actions and events.

4. Character traits and conditional statements

In this section I will expand upon the objection first raised in Chapter Six that the notion of a local character trait does not explain our ordinary use of ascriptions of dispositions. I will argue that defining character trait attributions as summaries of past behaviour, or historical narratives, provides a solution to this objection. I will argue that defining character trait attributions as narratives does explain our ordinary use of ascriptions of character.

Doris would presumably agree that there will need to be a further causal explanation of how someone comes to have a certain fine-grained trait. My argument is that such an explanation has to take narrative form and, *because it takes narrative form*, this allows us to integrate these fine-grained traits into coarser, global traits in a way denied by Doris. As explained in the previous chapter, he claims that the only type of evaluative integration possible is the grouping of local traits under thin evaluative headings such as 'good' and 'bad' (2002, 115). Under my view fine-grained character traits can be grouped under thick evaluative terms, such as 'generous', 'courageous' etc., allowing us to attribute coarser, global traits to people.

Doris identifies a potential problem with how situationist motives would fit into the narrative of an individual life:

> ...my unconscious filial piety admits of narrative integration precisely because it is meshed in a temporally extended story encompassing, among other things, my history with my family, the development of my relationship with my father, the development of other significant attachments and so on – in short, the story of my life. Determinative motives of the situationist kind, it seems to me, are not readily enmeshed in such biographies (2002, 143).

Doris argues that situationist motives can be linked to individuals by considering their broader plan. He gives the example of an individual not identifying with the callousness resulting from haste, but who does identify with the hectic lifestyle that leads him to hastiness and therefore callousness. He also suggests that particular behaviours do not occur in isolation, but have causal histories, in which situational factors may be relevant. Of course situational factors have a causal role; the disagreement is that it is also believed that character has a substantial causal role in narrative explanations. For example, Hutto says an understanding of an individual's reason for action 'requires a more or less detailed description of his or her circumstances, other propositional attitudes (hopes, fears), more basic perceptions and emotions and perhaps even his or her character, current situation and history' (2007, 43).

A narrative-historical explanation of character will go further in explaining why an individual has certain motives and emotions. As explained in Chapter Seven, in discussing the efficacy of belief-desire explanations, Goldie argues that we need 'thicker' explanations to explain human actions than mere belief-desire explanations and that these explanations will take the form of historical narratives (2007, 104). A similar argument applies to explanations of behaviour in terms of character traits. For example, under the standard definition of character trait attributions as conditional statements, the explanation

of why an individual acted in a certain way would take the form 'George did x because he is disposed to do x in certain relevant conditions'. Applying Goldie's argument, the suggestion is that such an explanation will not always be a full and satisfactory explanation; we need a historical narrative explanation that explains *why* George is disposed to act in that way. This narrative-historical explanation will go further in explaining why he has certain motives and emotions. Returning to the example of courage, the standard position implies that we would explain a certain action performed by George by saying 'he did that because he is disposed to be courageous in situations with relevant features'. I argued that when pressed to explain what we mean by this, we would explain that George acted in that way because in those circumstances with those features he is most likely to act in that sort of way. This explanation is not fully satisfactory because it merely describes George rather than explaining why George has this disposition and hence why he has acted in this way on that occasion.

Under Doris's definition of a local character trait, the explanation of why an individual acted in a certain way would take the form 'George did x because he is disposed to do x in these specific local conditions with markedly above chance probability p'. In Chapter Seven I suggested that such an explanation will not always be a full and satisfactory explanation; we need a historical narrative explanation that explains *why* George is disposed to act in that way. To recap, this narrative-historical explanation will go further in explaining why he has certain motives and emotions. Returning to the example of sailing-in-rough-weather-with-one's-friends-courageous, Doris's position implies that we would explain a certain action performed by George by saying 'he did that because he is disposed to be courageous when sailing in rough weather with his friends'. I argued that when pressed to explain what we mean by this, we would explain that George acted in that way because in those circumstances he acts in that way with markedly above chance probability. Again, I claimed that this explanation is not fully satisfactory because it merely describes George rather than explaining why George has this disposition and hence why he has acted in this way on that occasion.

Consider now what we may say under the definition of character traits as summaries of past behaviour. If using a character trait to explain George's action, we would say 'George did that because he is courageous when sailing in rough weather with his friends'. If pressed to explain further what we meant by this, we would not merely say that he is disposed to act that way in those circumstances with above chance probability, but would

explain his past behaviour that has led to this action. Perhaps George used to be very timid about sailing in rough weather, but went on a sailing course with a group of friends last summer. By doing this, he now has enough confidence in his skills and in his friends to act courageously when encountering rough weather, hence why he acted in this particular way on this particular occasion. This narrative-historical explanation gives a much fuller, satisfactory explanation of how George came to act in this way on this occasion. And we can assess that this narrative is true because it corresponds to how things were, as verified by multiple sources of evidence. Generally, when pressed to further explain how an individual's character explains a particular action, we will give a historical narrative explanation, rather than a statistical statement, indicating that perhaps the definition of character trait attributions as summaries of past behaviour is more akin to our normal use of such attributions. I think that an understanding of character traits as a historical narrative, based upon the account of narrative from the previous chapter, allows a character trait to be more than a summary of events and maintains the causal link between character trait and behaviour. In the example I gave, perhaps the audience feels admiration for George in overcoming his fears, as well as understanding the events that led up to this point. The narrative structure gives the audience the resources to evaluate George's character and actions, as well as explain them.

Thus far, I have still been considering character traits localised to particular types of circumstance. However, my main objections to Doris remain that the notion of a local character trait does nothing to explain our use of ordinary ascriptions of character traits and that it is unclear how one specifies the situation that appears in the character-trait statement. How might we integrate the fragments together into something more akin to our traditional view of character? My argument is that a full and satisfactory explanation of how an action was caused will often be a historical narrative and that this historical narrative may also describe the character of the person. Under such a description of character as a narrative, individual character traits are integrated with other relevant traits and other relevant information in explaining action.

For example, if challenged to explain why we have described George as courageous, we will tell a story about his past behaviour in a variety of situations, such as standing up for his beliefs, defending himself against a mugger, skiing the steepest slopes, sailing in rough weather and so on, and link these past actions to how he is now. The definition of a character trait as a historical narrative allows us to conclude from this variety of evidence

that George is in general courageous, not that he is courageous in four specific types of circumstance. Doris does not allow us to draw this conclusion because he argues that local traits are not evaluatively integrated. This definition also resolves the problem of how to specify the situation that features in the local character trait because the historical narrative will describe the relevant situational factors, yet allow the integration of the locally evidenced traits into a more general statement of character, such as courage. Such a narrative will not merely be a biographical list of actions that George has done in the past because these actions are linked together to explain why he is now courageous. The narrator is choosing which events are significant to this explanation, so George's courage is not just a function of several actions, but a disposition with its own worth.

This narrative construction appears to explain our ordinary use of character trait ascriptions. This explanation based upon past behaviour is a description of the person as having a certain character trait. Based upon this description we can make predictive inferences about what the person may do next, thus retaining the predictive nature of character traits without specifying percentages and relevant conditions. Of course, it is not necessary to base predictions upon a narrative-historical explanation of an individual's character. We frequently make predictions of behaviour on far less evidence. However, these predictions will not be reliable. A good prediction will be based upon a narrative-historical account. The causal link between character trait and behaviour is also retained because a story about an individual's past behaviour can causally explain a current action. My argument is not that the narrative explanation contains reference to global character traits, because in this instance Doris would simply argue that the explanation was false. To return to the example of Schindler cited above, in places the narrative explicitly states that he is generous and that this trait explains his behaviour. In this case, Doris would simply argue that the narrative is false because character traits do not function in this way. My argument is that a narrative about, for example, Schindler's behaviour in response to the seizing of Jewish property can be used to (partially) explain his subsequent actions in saving the Danziger brothers from being hanged. Such a narrative explanation of his actions also links the events together to give an overall description of his character traits; it does not simply refer to his traits to explain the actions.

In summary, the narrative construction appears to explain our ordinary use of character trait attributions. If considering how an individual's character explains a particular action, we will give a historical narrative explanation, rather than a statistical statement. If

considering a description of someone's character, when explaining or justifying this description we would also give a narrative-historical explanation that goes beyond merely listing some relevant actions from the individual's past.

5. Objections

This section considers possible objections to my argument. First, I revisit the objection that a narrative cannot be true and objective because historical narratives are open to interpretation and distortion to see if this causes particular problems for narratives about character traits. Secondly, I consider whether a narrative conception of character traits is psychologically realistic. Finally, I consider the argument that although fictional narratives give us information about character, real-life narratives do not function in this way. [66]

---

[66] A further potential objection is that an account of historical narratives is dependent upon identification of events as a distinct category. If facts and events are synonymous, the facts that feature in a historical narrative have no event to which to refer. This again would result in a type of anti-realism about character traits, as the events that constitute the narrative would not be distinct from the propositions about those events. Events are often defined as 'things that happen' but is this a metaphysically distinct category from facts or statements about things that have happened? A view that seems to conflate events with facts is that of Jaegwon Kim, who argues that an indefinite number of events can occur at the same time and place (Kim 1998). This is because he argues that events are the objects of explanations, where how the object has (or comes to have) a certain property is what is being explained. Because there are numerous individual properties, there are also numerous different events that can occur at any particular time and place. Lawrence Lombard gives the example 'if at a certain time an object, o, acquires the property of being red, it also acquires the property of being coloured, of being red and such that p (for any true proposition p), etc.; and since all those properties are different, the events which are o's acquiring of those properties must all be distinct' (1998, 284-5). He argues that the problem with this theory is that because pairs of properties are not identical 'no stabbing is a killing, no walk is a stroll, no party is a celebration and no run is a getting of some exercise' (1998, 285). He suggests that Kim's treatment of events is more like a theory of facts or true propositions. Should we conflate events with facts? There may well be numerous true propositions about a particular time and place that are not identical, but the events may be identical − Lombard gives the example of his running being his escape from the mugger, but the proposition that he escaped from the mugger is not the same as the proposition that he ran (1998, 285-6).

If not facts, what are the actions and events that provide the components of a historical narrative? F.P. Ramsey argues that there is a distinction between events and facts (1927, 156). He takes descriptions of events, such as 'that Caesar died', to assert the existence of that event at a certain place at a certain time. He does not think that the event of the death of Caesar should be confused with the fact that Caesar died. If they were equivalent, then any description of the event should be able to be substituted for another. He gives the example of an individual being aware that Caesar died. If facts were the same as events, 'that Caesar died' could be substituted for 'that Caesar was murdered', but clearly someone can be aware that Caesar died without being aware that he was murdered. The facts do not occur at a certain place and time whereas the events that are the subject of the facts do occur at a fixed point in time and space. W.V. Quine's view is that events occur in a particular place and time, so are no different from physical objects (1960, 170). Under his view one event is identical with another if it has the same spatiotemporal location. Lombard argues against this view on the grounds that it entails that if an object simultaneously changes colour and shape, this is the same event, whereas he thinks that these are two different events because colour and shape are different properties (1998, 283-4). He thinks that events do not occupy the locations where they occur in the same way that physical objects do, so two events can occur in the same place at the same time.

a) If character traits are the subject of a narrative, can such a narrative be objective or true? To recap, in the previous chapter I outlined Day's concern that narrative-historical writing is interpretive. He casts doubt on the claim that historical narratives are true because they correspond to past facts (2008, 188). He poses a challenge to a correspondence theory of truth, arguing that a problem arises because historical narratives can be true or false 'in a way that goes beyond the truth or falsity of their individual components' (2008, 196). By this, he means that the events that the historian chooses to narrate and the way that he puts them together are also capable of truth and falsity. He argues that the set of statements within the narrative are representative of the person or event as a whole, so the whole has a truth and that 'the challenge is to find something in the real past to which that holistic truth can correspond' (2008, 195).

I argued in Chapter Nine that this is precisely the sort of worry that leads to anti-realism about historical narratives, but that this tension can be resolved. It can be resolved by arguing that the narrator should use all available evidence when constructing his historical narrative. Hence a narrative about a person's character where all the statements are true yet the overall implication false would not be corroborated by other evidence. Such a narrative is false because the narrator has not used all available evidence. It can also be false if later evidence comes along which tells against it. There does not need to be a plot-like structure existing in the past for this holistic truth to correspond to. That there is disagreement over the implication or evaluation of a narrative about a person's character does not imply that

---

Jonathan Bennett does not think that there can be a systematic metaphysical answer to this problem (Bennett 1988 cited by Lombard 1998, 289). He thinks that an event such as an avalanche has many properties and the fact that it is an avalanche is an explanation of only one of many facts about the properties of the avalanche. He thinks it is indeterminate as to what properties are included in the explanation of an event. He argues that we are engaged in a semantic debate about the co-reference of event names, rather than a metaphysical debate about what it is for one event to be identical with another. He argues that no metaphysical theory will on its own determine the truth or falsity of an event identity statement. However, Bennett argues that despite this problem, events and facts are categorically distinct. He thinks this because we use different language to talk about events and to talk about facts. He uses the example of Bernard's journey to Calais and Bernard's cross-Channel swim to illustrate the distinction (1988, 10). He argues that clearly the swim was the journey, so these are names of the same event. This event corresponds to facts such as that Bernard journeyed to Calais or that he swam the Channel. He argues that these are two different facts because substituting one for the other can change the truth value of a statement, for example 'I am surprised that Bernard journeyed to Calais, but not that he swam the Channel' (1988, 11).

This brief summary shows that it is possible to argue for a distinction between events and facts. Facts and events are not synonymous, so the facts that feature in a historical narrative do have events to which to refer. This enables realism about character traits to be defended as the events that constitute the narrative are distinct from the propositions about those events. The identification of an event will not depend wholly on a metaphysical account of events, but also upon the narrative context in which it is placed.

there is no true or false narrative portrayal of the events; perhaps we just do not have enough conclusive evidence to understand the events portrayed. The same solution as for historical narratives appears to also apply to narratives specifically about character.

b) One of Doris's main complaints against virtue ethics is that it is based upon a psychologically unrealistic account of character; dispositional traits do not influence action as much as we believe. Is my alternative narrative account of character psychologically realistic? Does the narrative account better explain the interaction between trait and situation?

As explained in Chapter Four, Doris's central argument is that dispositional traits do not influence action as much as we might think and in fact the situation is more influential. Hence, he claims we should not predict what an individual will do based upon the character trait that we attribute to him. I will defend the narrative account of character as being psychologically realistic on two grounds. First, there is psychological evidence that character traits persist through time. Secondly, I argue that a narrative conception best reflects the doctrine of 'modern interactionism' (McAdams 2002, 292).

Doris argues that individual differences based on character traits are not consistent over time. However, there is evidence that suggests he is wrong. There are several longitudinal studies following the same group of people over a period of time that do find consistencies in the traits of individuals over time (McAdams 2002, 377). These studies find that trait scores at time t1 positively correlate with the scores at later times. Dan McAdams presents data from longitudinal studies into several traits such as neuroticism, extraversion, openness to experience, agreeableness and conscientiousness. Across these studies, he identifies a strong correlation of around +.65 between the identification of a trait at one time and the subsequent identification of that trait at a later time. As he points out, there is of course not a perfect correlation of +1 because we would expect there to be some change to an individual's traits over time. These data do seem to cast some doubt on the claim that we cannot predict behaviour based upon traits because they are not stable from one situation to the next. Sure, a trait may not be active or be over-ridden in a particular situation, but this is not to say that it does not persist through time.

The idea that traits persist or change through time lends itself to a narrative explanation of that change or stability. McAdams argues that dispositional traits outline some of the most

important features of an individual, but that when we hear an account of episodes within that individual's life we may feel that the trait attributions leave something out (McAdams 2002, 620). He argues this because he thinks that although trait attributions give us an understanding of behavioural consistencies in the individual, it gives us no understanding of how these episodes relate together or of what they mean to the individual. He thinks that: 'A person's internalized and evolving life story is as much part of his or her personality as is his or her dispositional traits and characteristic adaptations' (2002, 621). At the level of an understanding of the character of the individual, a narrative gives us an account of how the different individual traits integrate with each other. The story that the individual tells about his character reveals which traits have the most importance to the individual. At the level of an individual trait, a narrative gives us an account of how the different individual events contribute to the attribution of that character trait. The story that the individual tells about his character trait reveals which events he thinks have had the most importance to the individual in the formation of this trait.

The narrative account of character and character traits will not be independent of an account of the situations that the individual finds themselves in. The doctrine of 'modern interactionism' endorses a similar relationship between the individual and situations (McAdams 2002, 292). Under this approach, behaviour is 'a function of a continuous process of multidirectional interaction or feedback between the individual and the situation he or she encounters' (2002, 292). The individual has certain cognitive and motivational factors that influence behaviour and the situation influences behaviour depending upon how the individual interprets the situation. A narrative explanation of behaviour will similarly contain relevant elements of both the individual's character and features of the situation.

c) A final possible objection is put forward by Gregory Currie (Currie 2009). He argues that there is a difference between fictional character and real-life character based upon the fact that in fiction the protagonists act in order to inform the reader about their character, whereas in real-life people rarely act with the intention of informing others about their character, except perhaps when they are trying to mislead others (2009, 63). He continues to argue that fictional narratives allow the audience to make 'more inferential connections between actions, events, and Character' than we could ever hope to make when considering real people (2009, 63). He thinks that in real-life we will only have a 'vague, working understanding of a person's Character' whereas in a fictional narrative the

audience can confidently make judgments about the characters portrayed (2009, 63-4). In considering character studies of real people, Currie thinks that we use 'the particular to illustrate general traits in a life', so such character studies 'do not fit well with the particularizing, sequential aspirations of narrative' (2006, 311).

I do not think that we have to draw this conclusion when considering the character of real people. I think that there are two different uses of narrative in thinking about the character of real people. First, we can use a narrative to describe a person's character. This is akin to the use of particular events to illustrate general traits referred to by Currie. However, we can also use narratives to explain why a person has a particular character trait. This chapter has argued that a study of the character of a real person does fit well with the particularizing, sequential aspirations of narrative. If we are explaining why a person has a particular character trait or why he has acted in a certain way in a certain situation this explanation will be a historical narrative that takes particular events and actions from the individual's past and puts them in a sequence that explains the current event or action. Further this chapter has argued that real-life narratives about character can give us knowledge. Although it may be true that it may be easier to derive information about a character from a fictional narrative that has been designed for this purpose, this is not to say that it is not possible to construct an informative narrative about a real-life character. Certainly it is true that people make sweeping attributions of character without having the requisite evidence. For example, on the basis of one action a person may have the trait of courage attributed. Here the underlying narrative will not contain enough information to evaluate whether the person is courageous, but this is not to say that such a narrative could not be constructed.

Conclusion

In conclusion, I have argued that a character trait is an independent entity about which narratives can be told. I drew upon the characterisations of narratives from the previous chapters to argue that a narrative about a character trait is significant because it usually imparts an emotional understanding that is not available from a mere chronology of events. Such narratives about past episodes are important for understanding character traits because they not only explain how past actions and events have a causal affect on how we are now, but also give us an understanding of how to feel about those past actions and

events, giving us the resources for emotional understanding and evaluation of character traits.

1. Character traits

In summary my view is that character traits are best defined as historical narratives. Such narratives provide explanations of actions. Thus character traits provide explanations of particular actions. Character traits are not constituted by a narrative, i.e. the narrative is about something, character traits, which have independent existence. Narratives are necessary for understanding of character traits because of the knowledge and understanding that is derived from a narrative that could not be derived from any other means. This account provides an alternative to the conception of character trait attributions as conditional statements. It also provides an account of how a narrative attribution of a character trait can give a full explanation of an action. An individual character trait attribution is best defined as a summary of a historical narrative rather than as a statement of a conditional disposition. The narrative about the person as a whole will describe how all the individual character traits relate to each other. This narrative gives the audience, whether the person themselves in the case of an autobiographical narrative or a third party, the information it needs to evaluate the person. The narrative provides the detailed information about the person, including his actions, feelings and reasons, recommended by virtue ethics as important to evaluation of the person.

2. Virtue ethics

The definition of character trait attributions as summaries of historical narratives provides a sound foundation for a character-based ethics. This definition provides such a foundation for a character-based ethics because it is consistent with all the advantages of virtue ethics identified in Chapter One, as will be outlined below. In some areas, such as the broader focus of virtue ethics on emotions and reason as well as actions, the narrative explanation of character trait attributions as a summary of a historical narrative is a better fit than the traditional interpretation of character trait attributions as conditionals. The historical narratives account of character trait attributions better explains the interplay between actions and traits. Under this account the past actions of the individual aggregate together into virtuous or vicious traits and these traits explain why the individual acts in a certain way in the present or may do so in the future. The narratives of particular character traits are parts of an overall narrative of the person, providing an explanation of how the individual traits integrate in the fully virtuous person. Where the person has a coherent

narrative comprising of a network of virtuous traits, which themselves comprised various individual actions, emotions and reasons, the individual can be considered fully virtuous.

A narrative conception of character traits has to be coupled with an account of practical wisdom and *eudaimonia* to identify which of those character traits are virtues or vices. As set out in the introduction, the main aim of this thesis has not been to set out a list of virtues and vices. However, I think it possible that such a list can be developed from the traits for which there is evidence in psychology. For example, McAdams presents data from longitudinal studies into several traits, including neuroticism, extraversion, openness to experience, agreeableness and conscientiousness (McAdams 2002). I think that it is possible to use these character traits for the foundations of a narrative character-based ethics. Can these character traits be mapped to virtues and vices? Aristotle lists the virtues as courage, temperance, generosity, magnificence, magnanimity, pride, patience, truthfulness, wittiness and friendliness (Aristotle *Nicomachean Ethics*). Each virtue has a corresponding vice, either an excess or lack of the characteristic. The Aristotelian virtues are more fine-grained than the broader modern traits that have been identified, but the Aristotelian virtues can be grouped under these broader headings. For example, temperance could be part of extraversion; generosity, magnificence, magnanimity, wittiness, friendliness could parts of agreeableness; and pride, patience, truthfulness could be parts of conscientiousness. Neuroticism relates to emotional stability and the neurotic person tends to experience negative emotional states. One manifestation of this is that the individual is more likely to think that normal situations are threatening, so the Aristotelian virtue of courage has links to neuroticism. The 'openness to experience' trait would be needed to gain a wide range of experience to develop practical wisdom and the ability to choose the right action, for the right reasons in any given situation.

A virtue ethics based upon a narrative account of character trait attributions can best accommodate the complexity of everyday moral life. It is only through the construction of a narrative that the detail of a situation and why the person acted in a certain way can be fully explained. These narratives may correspond to the attribution of a particular trait and will aggregate to provide an account of the person overall. The standard definition of a character trait as the tendency to act in a certain way in relevant situations does not accommodate this complexity. Under that account a person either acts in accord with the relevant trait or not. However, it does not reveal anything about the person's reasons for acting in that way or his feelings about acting in that way, or about how his action on this

occasion integrates with how he has acted in the past. All these elements are relevant to a full understanding of why a person acts in a certain way in a certain situation. Merely observing behaviour and attributing this to the possession of a tendency to act in that way does not accommodate this complexity, nor does it aid evaluation of the person as a whole.

A virtue theory based upon a narrative account of character trait attributions accommodates the focus of virtue ethics on more than how people act. The actions of the person are important because they provide the material for the narrative about the person's character. However, the narrative will also contain references to many other features, such as the situation, the emotions of the person, his reasons for acting, and so on. A narrative account of character adequately captures all the elements that an Aristotelian account requires of a fully virtuous person. The narrative about the character of the fully virtuous agent will demonstrate that he is disposed to think, feel, choose, and act 'at the right times, about the right things, towards the right people, for the right end, and in the right way' (*Nicomachean Ethics* 1106b21). The traditional account of character that defines traits in terms of the action that they cause focuses only upon action. The narrative definition provides a better account of how actions relate to feelings and reasons, as well as the interdependency between traits and actions. It also explains the close relationship between reasoning and motivation, identified in Chapter One as an advantage that virtue theory has over other theories.

Virtue theory can accommodate degrees of moral understanding. A further advantage of a narrative view is that it allows for degrees of possession of character traits. Do we either have a character trait or not have a character trait? Daniel Dennett, when discussing narrative theories of the self, thinks that such a position allows for this to be a matter of degree (1991, 422). The traditional conditional view of character traits means that we either have a character trait or we do not; either we act in a certain way in the relevant situations with above chance probability or we do not. A narrative conception of character traits allows for a story to be told about that trait and how it interrelates with other character traits, allowing for a degree of possession of a trait. For example, a story can be told about a person that describes him as more or less kind, explaining how his kindness is affected by his other traits. Schindler could be described as possessing a degree of kindness. A narrative about his life illuminates kindness towards the Jews he saved but unkindness towards his wife whom he ultimately left. A question as to whether he is kind does not admit a 'yes' or 'no' answer; attribution of kindness is a matter of degree and this

is accommodated by the narrative view. The narrative account of character allows a narrative to be constructed about individual traits or about the person as a whole. The narrative about the person as a whole will reveal the degree to which the individual is virtuous. The narrative about an individual trait will reveal the degree to which the person possesses a specific trait. The traditional interpretation of character trait attributions as conditionals does not so easily accommodate degrees of possession of traits; the individual either acts in accord with the trait or not.

A virtue ethics based upon a narrative account of character also provides an appealing account of moral development. The narrative reveals a clear developmental story as to how the person became virtuous. The fully virtuous person will have coherence to his narrative because he will consistently act, feel and reason in a virtuous way. The less than fully virtuous person will have a less coherent narrative because his behaviour, emotions and reasoning will not follow regular patterns. This narrative account, therefore, accommodates some of Doris's unease about the fragmentation of character that can be observed. When an individual appears fragmented, this is because he is less than fully virtuous. There may be a degree of coherence to the narrative of such a person, but where there are inconsistencies, the character of the person will appear fragmented. The development over time of this narrative will reveal whether the person becomes more virtuous over time. The narrative will explain how the person reacts to past events and how he uses past experiences to inform his future actions. It is not necessary that the fully virtuous person experience his life as a narrative with this type of coherence, but it is the case that it would be possible to tell a coherent narrative of the fully virtuous person.

In conclusion a virtue ethical account can be based upon a narrative account of character trait attributions. Not only *can* virtue ethics be based upon such an account, this account provides a better fit with some of the demands of virtue ethics than a more traditional account of character traits as conditionals. There are four main advantages of my account. The first advantage is that it better accommodates the complexity of everyday moral life because a narrative contains more detail than a conditional statement. The second advantage is that it provides a better account of the importance of emotions, feelings, reasons, motivation, as well as actions, to the evaluation of an individual or his action as virtuous. The third advantage is that the narrative provides an explanation of the degree to which a person is virtuous. The final advantage is that the narrative provides a story about the moral development of the individual. The narrative account of character trait

attributions, coupled with an account of practical wisdom and *eudaimonia*, provides a stronger basis for a virtue ethics based upon Aristotelian principles than the account of character trait attributions as conditional statements. My account respects the differences between persons, as well as the differences between situations and the different perceptions of those situations, and the complexity of moral life. Only through recounting the story of my life can my virtue be evaluated.

Ajzen, I., 1988, *Attitudes, Personality and Behaviour*, Chicago, Ill.: Dorsey Press

Alvarez, M., 2008, 'Reasons and the Ambiguity of Belief', *Philosophical Explorations*, Vol. 11, Issue 1, 53-65

Amis, M., 1991, *Time's Arrow*, London: Penguin

Annas, J., 2003, 'Virtue Ethics and Social Psychology', *A Priori*, Vol. 2, 20-32

Anscombe, G.E.M., 1957, *Intention*, London: Blackwell. Reprinted 2000, *Intention: Second Edition*, Cambridge, Massachusetts: Harvard University Press

Anscombe, G.E.M., 1958, 'Modern Moral Philosophy', *Philosophy*, Vol. 33, No. 124, 1-19

Aristotle, 1996, *Poetics*, translated by M. Heath, London: Penguin

Aristotle, 1999, *Nicomachean Ethics*, translated by T. Irwin, 2nd Edition, Indianapolis: Hackett

Athanassoulis, N., 2000, 'A Response to Harman: Virtue Ethics and Character Traits', *Proceedings of the Aristotelian Society*, Vol. 100, 215-221

Austen, J., 1813, *Pride and Prejudice*, Penguin Classics edition, 2003, London: Penguin

Austen, J., 1811, *Sense and Sensibility*, Oxford World's Classics edition, 2008, Oxford: Oxford University Press

Badhwar, N.K., 1996, 'The Limited Unity of Virtue', *Noûs*, Vol. 30, No. 3, 306-329

Barwell, I., 2009, 'Understanding Narratives and Narrative Understanding' in N. Carroll (ed.) *The Poetics, Aesthetics, and Philosophy of Narrative*, Oxford: Blackwell, The American Society for Aesthetics, 2009, 49-59

Bennett, J., 1988, *Events and Their Names*, Indianapolis: Hackett

Burnyeat, M.F., 1980, 'Aristotle on Learning to be Good' in A.O. Rorty (ed.) *Essays on Aristotle's Ethics*, Berkeley: University of California Press, 1980, 69-92

Carroll, N., 1990, 'Interpretation, History and Narrative', *Monist* 73: 2, 134-67

Carroll, N., 2001, 'On the Narrative Connection' in his *Beyond Aesthetics*, Cambridge: Cambridge University Press, 118-132

Carroll, N., 2007, 'Narrative Closure', *Philosophical Studies* 135, 1-15

Collins, A.W., 1997, 'The Psychological Reality of Reasons', *Ratio*, Vol. 10, Issue 2, 108-23

Currie, G., 2006, 'Narrative Representation of Causes', *The Journal of Aesthetics and Art Criticism*, Vol. 64, Issue 3, 309-316

Currie, G., 2009, 'Narrative and the Psychology of Character' in N. Carroll (ed.) *The Poetics, Aesthetics, and Philosophy of Narrative*, Oxford: Blackwell, The American Society for Aesthetics, 2009, 61-72

Dancy, J., 1995, 'Why There is Really no such Thing as the Theory of Motivation', *Proceedings of the Aristotelian Society*, Vol. 95, 1-18

Dancy, J., 1996, 'Real Values in a Human Context', *Ratio*, Vol. 9, Issue 2, 171-83

Dancy, J., 2000a, 'Recognition and Reaction' in R. Crisp and B. Hooker (eds.), *Well-Being and Morality: Essays in Honour of James Griffin*, Oxford: Clarendon Press, 2000, 39-52

Dancy, J., 2000b, *Practical Reality*, Oxford: OUP

Dancy, J., 2004, *Ethics Without Principles*, Oxford: Oxford University Press

Darley, J.M. and Batson, C.D., 1973, '"From Jerusalem to Jericho": A Study of Situational and Dispositional Variables in Helping Behaviour', *Journal of Personality and Social Psychology* 27, 100-108

Davidson, D., 1963, 'Actions, Reasons, Causes' reprinted in his *Essays on Action and Events*, 2nd Edition, Oxford: Clarendon Press, 2001, 3-19

Day, M., 2008, *The Philosophy of History*, London: Continuum

Dennett, D., 1991, *Consciousness Explained*, New York: Penguin Books

Doris, J.M., 1998, 'Persons, Situations and Virtue Ethics, *Noûs*, Vol. 32, No. 4, 504-530

Doris, J.M., 2002, *Lack of Character*, Cambridge: Cambridge University Press

Dretske, F., 1988, *Explaining Behaviour*, Cambridge, Massachusetts: MIT Press

Eternal Sunshine of the Spotless Mind, 2004, film, directed by Michel Gondry, USA: Anonymous Content

Fara, M., 2005, 'Dispositions and Habituals', *Noûs*, Vol. 39, No. 1, 43-82

Flanagan, O., 1991, *Varieties of Moral Personality*, Cambridge, Massachusetts: Harvard University Press

Forster, E.M., 1927, *Aspects of the Novel*, O. Stallybrass (ed.), London: Penguin

Goldie, P., 2003a, 'Narrative, Emotion, and Perspective' in M. Kieran and D. M. Lopes (eds.), *Imagination, Philosophy and the Arts*, London: Routledge, 54-68

Goldie, P., 2003b, 'Narrative and Perspective: Values and Appropriate Emotions' in A. Hatzimoysis (ed.) *Philosophy and the Emotions*, Royal Institute of Philosophy Supplements Series, Cambridge: Cambridge University Press, 201-220

Goldie, P., 2004a, 'Narrative, Emotion and Understanding', in B. Hurwitz, T. Greenhalgh and V. Skultans (eds.), *Narrative Research in Health and Illness*, London: Blackwell, BMJ Books, 156-67

Goldie, P., 2004b, *On Personality*, London: Routledge

Goldie, P., 2007, 'There are Reasons and Reasons' in D.D. Hutto and M. Ratcliffe (eds.) *Folk Psychology Re-Assessed*, Dordrecht: Springer, 103-114

Goldie, P., 2009, 'Narrative Thinking, Emotion and Planning' in N. Carroll (ed.) *The Poetics, Aesthetics, and Philosophy of Narrative*, Oxford: Blackwell, The American Society for Aesthetics, 2009, 97-106

Grover, L., 2008, 'A Defence of Internal Reasons', *Richmond Journal of Philosophy*, Issue 18, http://www.richmond-philosophy.net/rjp/rjp18_grover.php (last accessed 29th July 2010)

Hampshire, S., 1953, 'Dispositions', *Analysis*, Vol. 14, No. 1, 5-11

Haney, C., Banks, W. and Zimbardo, P., 1973, 'Interpersonal Dynamics of a Simulated Prison', *International Journal of Criminology and Penology* 1, 69-97

Haney, C. and Zimbardo, P., 1977, 'The Socialization into Criminality: On Becoming a Prisoner and a Guard' in J. Tapp and F. Levine (eds.), *Law, Justice, and the Individual in Society: Psychological and Legal Issues*, New York: Holt, Rinehart and Winston

Harman, G. and Thomson, J.J., 1996, *Moral Relativism and Moral Objectivity*, Oxford: Blackwell

Harman, G., 1999a, 'Virtue Ethics without Character Traits', http://www.princeton.edu/~harman/Papers/Thomson.html (7 pages, last accessed 29th July 2010)

Harman, G., 1999b, 'Moral Philosophy Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error', *Proceedings of the Aristotelian Society* Vol. 99, 315–331

Harman, G., 2000, 'The Nonexistence of Character Traits', *Proceedings of the Aristotelian Society* Vol. 100, 223-226

Harman, G., 2003, 'No Character or Personality', *Business Ethics Quarterly*, Vol. 13, Issue 1, 87-94

Hartshorne, H. and May, M.A., 1928, *Studies in the Nature of Character: Studies in Deceit Vol. 1*, New York: Macmillan

Hudson, S.D., 1986, *Human Character and Morality: Reflections from the History of Ideas*, Boston, London and Henley: Routledge and Kegan Paul

Hume, D., 1739-40, *A Treatise of Human Nature*, L.A. Selby-Bigge (ed.), Oxford: Oxford University Press

Hurka, T., 2001, *Virtue, Vice, and Value*, Oxford: Oxford University Press

Hurka, T., 2006, 'Virtuous Act, Virtuous Dispositions', *Analysis*, Vol. 66, No. 1, 69-76

Hursthouse, R., 1991, 'Virtue Theory and Abortion', *Philosophy and Public Affairs* 20, 223-246

Hursthouse, R., 1999, *On Virtue Ethics*, Oxford: Oxford University Press

Hutto, D.D., 2007, 'The Narrative Practice Hypothesis: Origins and Applications of Folk Psychology', in D.D. Hutto (ed.), *Narrative and Understanding Persons*, Royal Institute of Philosophy Supplements Series, Cambridge: Cambridge University Press, 43-68

Isen, A.M. and Levin, P.F., 1972, 'Effect of Feeling Good on Helping: Cookies and Kindness', *Journal of Personality and Social Psychology*, Vol. 21, No. 3, 384-388, Study II

Johnston, M, 1992, 'How to Speak of the Colours', *Philosophical Studies* 68, 221-63

Kamtekar, R., 2004, 'Situationism and Virtue Ethics on the Content of Our Character', *Ethics*, Vol. 114, No. 3, 458-491

Kant, I., 1797, *The Metaphysics of Morals*, translated by M. Gregor, 1991, Cambridge: Cambridge University Press

Keneally, T., 1982, *Schindler's Ark*, London: Hodder and Stoughton

Keneally, T., 2009, *Searching for Schindler*, London: Hodder and Stoughton

Kieran, M., 2003, 'In Search of a Narrative' in M. Kieran and D. M. Lopes (eds.), *Imagination, Philosophy and the Arts*, London: Routledge, 69-87

Kim, J., 1998, 'Events as Property Exemplifications' in S. Laurence and C. Macdonald (eds.) *Contemporary Readings in the Foundations of Metaphysics*, Oxford, Blackwell, 1998, 310-26

Kirkham, R.L., 2001, *Theories of Truth: A Critical Introduction*, Fifth printing, Cambridge, Massachusetts: MIT Press

Kohlberg, L., 1981, *Essays on Moral Development: Volume One. The Philosophy of Moral Development: Moral Stages and the Idea of Justice*, New York: Harper and Row

Kosman, L.A., 1980, 'Being Properly Affected: Virtues and Feelings in Aristotle's Ethics', in A.O. Rorty (ed.) *Essays on Aristotle's Ethics*, Berkeley: University of California Press, 103-116

Lamarque, P., 2004, 'On Not Expecting Too Much from Narrative', *Mind and Language*, Vol. 19, No. 4, 393-408

Latané, B. and Darley, J.M., 1970, *The Unresponsive Bystander: Why Doesn't He Help?* New York: Appleton-Century-Crofts

Latané, B. and Rodin, J., 1969, 'A Lady in Distress: Inhibiting Effects of Friends and Strangers on Bystander Intervention', *Journal of Experimental Social Psychology* 5, 189-202

Lewis, C.S., 2001, *The Complete Chronicles of Narnia*, New York: HarperCollins

Lewis, D., 1987, 'Causal Explanation', in his *Philosophical Papers Volume II*, New York: Oxford University Press, 214-240

Lewis, D., 1997, 'Finkish Dispositions', *The Philosophical Quarterly*, Vol. 47, 143–158

Livingston, P., 2009, 'Narrativity and Knowledge' in N. Carroll (ed.) *The Poetics, Aesthetics, and Philosophy of Narrative*, Oxford: Blackwell, The American Society for Aesthetics, 2009, 25-36

Lombard, L., 1998, 'Ontologies of Events' in S. Laurence and C. Macdonald (eds.) *Contemporary Readings in the Foundations of Metaphysics*, Oxford, Blackwell, 1998, 277-294

MacIntyre, A., 1984, *After Virtue*, 2nd edition, Notre Dame: University of Notre Dame Press

Manley, D. and Wasserman, R., 2008, 'On Linking Dispositions and Conditionals', *Mind*, Vol. 117, No. 465, 59-84

Martin, C.B., 1994, 'Dispositions and Conditionals', *Philosophical Quarterly*, Vol. 44, 1-8

McDowell, J., 1979, 'Virtue and Reason', *Monist* 62, 331-50 reprinted in McDowell's *Mind, Value, and Reality*, 1998, Cambridge, Massachusetts: Harvard University Press, 50-76

McAdams, D.P., 2002, *The Person: An Integrated Introduction to Personality Psychology*, Third Edition, New York: Wiley

Memento, 2000, film, directed by Christopher Nolan, USA: Newmarket Capital Group

Merritt, M, 2000, 'Virtue Ethics and Situationist Personality Psychology', *Ethical Theory and Moral Practice*, Vol. 3, 365-383

Milgram, S., 1974, *Obedience to Authority*, London: Tavistock

Mill, J.S., 1859, *On Liberty*, E. Rapaport (ed.), 1978, Indianapolis: Hackett

Miller, C., 2003, 'Social Psychology and Virtue Ethics', *The Journal of Ethics*, Vol. 7, Issue 4, 365-392

Misak, C., 2008, 'Experience, Narrative, and Ethical Deliberation', *Ethics*, Vol. 118, No. 614-632

Mischel, W., 1999, 'Personality Coherence and Dispositions in a Cognitive-Affective Personality System (CAPS) Approach' in D. Cervone and Y. Shoda (eds.), *The Coherence of Personality: Social-Cognitive Bases of Consistency, Variability and Organisation*, New York and London: Guilford Press

Mischel, W. and Shoda, Y., 1995, 'A Cognitive-Affective System Theory of Personality: Reconceptualising Situations, Dispositions, Dynamics, and Invariance in Personality Structure', *Psychological Review* 102, 246-68

Mulholland Drive, 2001, film, directed by David Lynch, France and USA: Les Films Alain Sarde

Nemirovsky, I., 2007, *Suite Francaise*, translated by S. Smith, London: Vintage

Newcomb, T.M., 1929, *The Consistency of Certain Extrovert-Introvert Behaviour Patterns in 51 Problem Boys*, New York: Columbia University, Teachers College, Bureau of Publications

Niffenegger, A., 2005, *The Time Traveller's Wife*, London: Vintage

Nussbaum, M., 1999, 'Virtue Ethics: A Misleading Category?', *The Journal of Ethics*, Vol. 3, Issue 3, 163-201

Pettit, P. and Smith, M., 1990, 'Backgrounding Desires', *Philosophical Review*, Vol. 99, No. 4, 565-592

Piller, C., 2003, 'Critical Notice: Jonathan Dancy, Practical Reality', *Philosophy*, Vol. 78, Issue 3, 414-425

Pincoffs, E., 1971, 'Quandary Ethics', *Mind*, Vol. 80, No. 320, 552-571

Prinz, J., forthcoming, 'The Normativity Challenge: Why Empirically Real Traits Won't Save Virtue Ethics' in *The Journal of Ethics*, http://subcortex.com/NormativityCharacterPrinz.pdf (last accessed 29th July 2010)

Quine, W.V., 1960, *Word and Object*, Cambridge, Massachusetts: MIT Press

Ramsey, F.P., 1927, 'Facts and Propositions', *Proceedings of the Aristotelian Society, Supplementary Volumes*, Vol.7, 153-70

Rawls, J., 1971, *A Theory of Justice*, Cambridge, Massachusetts: Harvard University Press

Ross, L. And Nisbett, R., 1991, *The Person and the Situation: Perspectives of Social Psychology*, New York: McGraw-Hill

Ross, W.D., 1930, *The Right and The Good*, Oxford: Clarendon Press

Ross, W.D., 1939, *Foundations of Ethics*, Oxford: Clarendon Press

Roth, A., 1999, 'Reasons Explanations of Actions: Causal, Singular and Situational', *Philosophy and Phenomenological Research*, Vol. 59, No. 4, 839-874

Schueler, G.F., 2009, 'The Humean Theory of Motivation Rejected', *Philosophy and Phenomenological Research*, Vol. 78, No. 1, 103-122

Sehon, S.R., 1997, 'Deviant Causal Chains and the Irreducibility of Teleological Explanation', *Pacific Philosophical Quarterly* 78, 195-213

Shoda, Y., Mischel, W. and Wright, J.C., 1994, 'Intraindividual Stability in the Organisation and Patterning of Behaviour: Incorporating Psychological Situations into the Idiographic Analysis of Personality', *Journal of Personality and Social Psychology* 67, 674-87

Sidgwick, H., 1966, *The Methods of Ethics*, seventh edition, New York: Dover

Slote, M., 1997, 'Virtue Ethics' in M.W. Baron, P. Pettit, and M, Slote, *Three Methods of Ethics*, Oxford: Blackwell, 175-238

Smith, M, 1994, *The Moral Problem*, Oxford: Blackwell

Sorabji, R., 1980, 'Aristotle on the Role of Intellect in Virtue' in A.O. Rorty (ed.) *Essays on Aristotle's Ethics*, Berkeley: University of California Press, 201-219

Sreenivasan, G., 2002, 'Errors about Errors: Virtue Theory and Trait Attribution', *Mind*, Vol. 111, No. 411, 47-68

Strawson, G., 2004, 'Against Narrativity', *Ratio*, Vol. 17, Issue 4, 428-452

The Italian Job, 1969, film, directed by Peter Collinson, UK: Oakhurst Productions

Thomson, J.J., 1997, 'The Right and the Good', *The Journal of Philosophy*, Vol. 94, No. 6, 273-298

Trudeau, N.A., 2002, *Gettysburg: a Testing of Courage*, New York: Harper Collins

Tversky, B., 2004, 'Narratives of Space, Time, and Life', *Mind and Language*, Vol. 19, Issue 4, 380-91

Velleman, J.D., 2003, 'Narrative Explanation', *The Philosophical Review*, Vol. 112, No. 1, 1-25

Vranas, P.B.M., 2005, 'The Indeterminacy Paradox: Character Evaluations and Human Psychology', *Noûs*, Vol. 39, No. 1, 1-42

Wallace, R.J., 1990, 'How to Argue About Practical Reason', *Mind*, Vol. 99, No. 395, 355-85

Walsh, W.H., 1970, *An Introduction to the Philosophy of History*, London: Hutchinson (originally published 1951)

Webber, J., 2006a, 'Virtue, Character and Situation', *Journal of Moral Philosophy*, Vol. 3, No. 2, 193-213

Webber, J., 2006b, 'Character, Consistency and Classification', *Mind*, Vol. 115, No. 459, 651-658

White, H., 1973, *Metahistory*, Baltimore: Johns Hopkins Press

White, H, 1978, 'The historical text as literary artefact', in his *Tropics of Discourse*, Baltimore: Johns Hopkins Press, 81-100

Wiggins, D., 1998, *Needs, Values, Truth*, Oxford: Clarendon

Williams, B.A.O., 1981, 'Internal and External Reasons', *Moral Luck*, Cambridge: Cambridge University Press, 101-113

Williams, B.A.O., 1985, *Ethics and the Limits of Philosophy*, London: Fontana

Williams, B.A.O., 2001, 'Postscript: Some Further Notes on Internal and External Reasons' in Millgram, E. (ed.), *Varieties of Practical Reasoning*, Cambridge, Massachusetts: MIT Press, 91-97

Wilson, G.M., 1989, *The Intentionality of Human Action*, Stanford, California: Stanford University Press

Wollheim, R., 1984, *The Thread of Life*, New York: Cambridge University Press

Wright, G. H. von, 1963, *The Varieties of Goodness*, London: Routledge and Kegan Paul