

**REASONABLE DISPOSITIONS: A RESPONSE TO THE
SITUATIONIST CRITIQUE OF VIRTUE ETHICS**

BY

TOBIAS VANDENBERG

A thesis

Submitted to the Victoria University of Wellington

In fulfilment of the requirements for the degree of

Master of Philosophy

Victoria University of Wellington

2018

Acknowledgements

I would like to thank Ramon for his advice and assistance through all the years I've studied under him; my parents for their unwavering support; Jo for always asking the difficult questions and never letting me off with an easy answer; Courtney for her talent for ferreting out contractions; Matt for his boundless enthusiasm for all things philosophical; Paul for putting up with me; and everyone else who has made this possible.

Abstract

In *The Nicomachean Ethics* Aristotle presents a theory of ethics as the development of character, specifically the cultivation of particular character traits lying on the mean between two vices which he calls the virtues. Since the revival of virtue ethics in the mid-20th century, the virtues have often been interpreted as stable, broad-based dispositions to act in virtue appropriate fashion in response to eliciting conditions. In more recent years virtue ethics has come under a sustained attack known as the ‘situationist critique’, which argues that experiments in social psychology show no evidence of broad-based dispositions in the general population. If there are no broad-based dispositions, there are no virtues – virtue ethics is therefore empirically inadequate. My analysis of the evidence the situationists present in favour of their critique will show that it fails to unambiguously support the conclusion that virtue ethics is empirically inadequate. An investigation into the causes of the ethical failings identified in those experiments allowed me to propose remedies that are perfectly consonant with Aristotelian virtue ethics, requiring dedicated and disciplined reflection on past actions with an ongoing commitment to bringing our thoughts and emotions into harmony with our values. In the process I show that the situationists’ evidence supports the “cognitive-affective” theory of virtue over the broad-based dispositional theory, that this cognitive-affective theory is also supported by a careful reading of Aristotle, and therefore that the situationist critique simply helps to replace an inadequate virtue theory with a more robust one.

Introduction

What we might call the rediscovery of Aristotelian ethics in the 20th century has undoubtedly invigorated the ethical debate, and in particular has brought notion of character back into focus. Aristotle's *Nicomachean Ethics* was the inspiration for a new class of ethical theories that became known as 'virtue ethics' in reference to the virtues, the key components of character outlined in that famous work. To its supporters, virtue ethics offers insight into and guidance on important aspects of human life that were ignored or side-lined in the more venerable deontological and utilitarian ethical theories (c.f. Hursthouse 2001, chapter 1 for discussion). In particular, virtue ethics was seen to offer a holistic view of ethical action, bringing the emotionality of human life out from behind the shadow of rationality, placing the two side-by-side and showing how these two aspects the ethical agent can work in harmony. Rather than deriving a set of moral laws, virtue ethics recommends that we develop virtues, conceived as character traits that lie on a mean between two vices and which are (at least typically) beneficial both to the possessor and to society as a whole. Because it takes the whole human being into view and offers a unified theory of human agency virtue ethics is often prized for its naturalism, and indeed this is what attracts many of its adherents. However, this naturalism also makes virtue ethics vulnerable in ways that less overtly naturalistic ethical theories (such as deontology) manage to avoid; virtue ethics is vulnerable to a charge of 'empirical inadequacy.'

In this paper I will discuss a critique of virtue ethics that pursues this very line. The 'situationist critique' of virtue ethics makes the case that virtue ethics is empirically inadequate in the sense that it is unsupported by the available scientific evidence. As Harman, one of the most prominent and vocal situationists describes it, 'Empirical studies designed to test whether people behave differently in ways that might reflect their having different character traits have failed to find relevant differences...Since it is possible to explain our ordinary belief in character traits as deriving from certain illusions, we must conclude that there is no empirical basis for the existence of character traits' (Harman 1999, p.316). Given this lack of an empirical basis for the existence of character traits, the

situationists would argue that our commitment to naturalism should lead us to disavow virtue theory.

I intend to show that the situationist critique fails. In particular, I will show that the critique rests on a narrow interpretation of the concept of a virtuous disposition and the relation between such dispositions and virtuous character traits. The evidence that the situationists appeal to not only admits of explanations that are consonant with virtue ethics, but actually supports an alternative understanding of the nature of virtuous character traits that has a firmer basis in Aristotle's own work. Moreover, far from undermining the importance of the development of character in ethics, I will show the evidence actually reinforces its importance.

I will proceed as follows: In chapter one I present a summary of Aristotle's ethical theory and in particular his theory on the nature virtues as character traits, followed by a brief discussion of more contemporary formulations. In chapter two I start by outlining the situationists' understanding of the virtues as broad-based, robust dispositional traits before presenting a summary of the experiments and historical case studies the situationists appeal to as evidence for their critique, including a situationist interpretation following each. In chapter three I proceed with a critical evaluation of the evidence, discussing a number of objections that have arisen in the literature, and while concluding these objections do not invalidate the situationist critique I derive some questions that a viable virtue theory would need to address in order to counter said critique. Finally, in chapter four I show that virtue ethics can answer those questions, that it does offer the tools we need to mitigate the pernicious influence of situational factors provided we disciplined enough to practice them, and arguing that the situationist evidence in fact supports a "cognitive-affective" theory of virtue over the "dispositional" theory.

On notion: Throughout this paper I use the terms "situationism" and "situationist critique" interchangeably. This was a deliberate choice to enhance readability. Strictly speaking "situationism" is a paradigm in social psychology rather than a critique of virtue ethics, but given that "situationism" isn't a prevalent term in ethical discussions I considered there was no danger of ambiguity.

Chapter 1 – Virtue Ethics: Past and Present

The rebirth of virtue ethics in the mid-20th century followed a period in which the philosophical debate was dominated by just two classes of theory: deontological and utilitarian. While there has been a great deal of development and ‘cross-pollination’ in the decades since, virtue ethics was originally conceived as ‘an approach in normative ethics which emphasizes the virtues, or moral character, in contrast to an approach which emphasizes duties or rules (deontology) or one which emphasizes the consequences of actions (utilitarianism)’ (Hursthouse 2001, p. 1). The feeling among the early advocates of virtue ethics was that the overwhelming focus of normative ethics on identifying the correct ‘principle of right action’ is disastrously one-sided, the ethical debate had become so obsessed with ethical action that it had lost sight of the ethical actor. Virtue ethics can be seen as an attempt to restore the balance: rather than beginning with a set of rules or imperatives and classifying the ethical agent as one who complies with such rules, virtue ethics begins with a discussion of the nature of the ethical agent and only then seeks to derive principles of action. Indeed, even this is saying too much, as one of the distinctive characteristics of virtue ethics (as opposed to deontology and utilitarianism) is that there is (typically) no sharp act/actor division; to correctly interpret the act we must also interpret the actor, and vice versa. This holistic view of ethical action, unifying individual psychology and action, offers a sharp contrast with deontological and utilitarian approaches in which the actor often appears to be nothing more than a cypher, an X in the ethical equation, lacking any characteristics other than the capacity to comply or fail to comply with the rule in question.

While this commitment to taking the real, living individual into account is part of the appeal of virtue ethics, it does open it up to class of critiques to which deontology and utilitarianism are largely immune. Virtue ethics has rather a lot to say about human psychology; in particular, it is committed to the claim that human beings have, or at the very least are able to develop, certain character traits (on which more below). If it can be shown that humans do not have and cannot develop such traits, then virtue ethics automatically fails. Normative ethics is fundamentally a practical discipline, and if it is impossible to *be* as

virtue ethics says one should be, then it is hard to see what value it could have. The situationist critique is of precisely this kind; it states that virtue ethics is founded on a false theory of human psychology. Situationism is ‘an error theory. It claims that people are systematically mistaken in attributing virtues’ (Alfano 2011, p. 123).

Before turning to discuss the situationist critique in detail however I will need to narrow down the target. So far I have talked rather loosely about ‘virtue ethics’ as though it were a unitary concept, but of course there is no single virtue ethical theory. While there is some broad agreement (at a minimum they all speak of virtues or virtuous acts), there are also some critical differences that will determine whether the situationist critique is even a starter. Virtue ethics in its modern form was essentially a rediscovery of Aristotle’s ethical theories, and even now his *Nicomachean Ethics* is the closest thing we have to an authoritative text on the subject, so I will begin by presenting an outline of his views.

1.1. Aristotelian virtue: general aspects

Aristotle makes it clear at the outset of *The Nicomachean Ethics* that his goal is to determine what is ‘good’, not in some absolute or transcendent sense but in the very specific and practical sense of determining what constitutes a good human life. After considering and discarding various possibilities he settles on *eudaimonia* (variously translated as ‘happiness’, ‘well-being’ or ‘flourishing’) as ‘the good’, that is, the one thing that a human can possess that is good in the unqualified sense. By way of illustration, it is worth considering an analogy that Aristotle himself uses a number of times, the nature of health. Just as physical health is good in an unqualified sense (the very notion that it might be preferable to be unhealthy strikes one as absurd, and one would immediately suspect someone who claimed otherwise was being disingenuous), so *eudaimonia* is good in an unqualified sense, and indeed it would seem that there is good reason to think that Aristotle views *eudaimonia* as a kind of ‘health of the soul’ which parallels physical health, for example when he says: ‘By human virtue we mean not that of the body but that of the soul; and happiness also we call an activity of the soul’ (NE 1102a: 14-16).¹ *Eudaimonia* is the

¹ For further examples of the health analogy see NE: 1097a: 5-15 & 1104a: 4-10, among others. It is also worth noting in this context that the word translated as ‘soul’ is the Greek word *psyche*, which could be translated as

central concept of Aristotle's ethics and the whole point of his work is to assist his readers in achieving it, as he makes quite clear when he says 'the present inquiry does not aim at theoretical knowledge...(for we are inquiring not in order to know what virtue is, but in order to *become* good, since otherwise our inquiry would have been of no use)' (NE: 1103b: 25-30, my italics). It is worth drawing attention to this point because it is often lost; Aristotle's one and only goal is to assist us in achieving *eudaimonia*. The virtues can be conceived as properties of the person who possesses *eudaimonia*, and acquiring the virtues as the means towards achieving *eudaimonia*. But what are virtues?

First and foremost, virtues (and vices) are character traits, a notion that will be critical to the situationist critique and which will be explored in more detail presently. As such they are properties of the individual person (in Aristotle's terms: properties of the soul, in ours: psychological properties), and, crucially, they are causative: 'for as a *result* of virtue men tend to do noble deeds' (NE 1101b: 30-31, my italics). This causative element distinguishes Aristotelian virtue ethics from theories that speak of virtuous acts without ascribing the source of the act to an underlying trait in the actor. In these latter cases virtue terms (or 'aretaic' terms as they are sometimes called) are employed in a purely descriptive sense, rather than in an explanatory sense. Even more crucial for the discussion to come is the fact that the virtues do not arise in us 'by nature'; they are not traits we *have* but traits we *acquire*. 'Neither by nature...nor contrary to nature do the virtues arise in us; rather we are adapted by nature to receive them, and are made perfect by habit' (NE 1103a: 23-25) as Aristotle says, in characteristically dense but also highly suggestive fashion.

The above passage raises some interesting possibilities, which on reflection fit well with another of Aristotle's famous analogies, which compares the virtues with the arts. In fact, it's worth quoting him at some length here as this is a comparison he returns to time and again:

'the virtues we get by first exercising them, as also happens in the case of the arts as well. For the things we have to learn before we can do them, we learn by doing them...

'mind' in other contexts and is the source of our word 'psychology'. I will return to this in more detail in the final chapter.

This, then, is the case with the virtues also; by doing the acts that we do in our transactions with other men we become just or unjust, and by doing the acts that we do in the presence of danger, and by being habituated to feel fear or confidence, we become brave or cowardly...Thus, in one word, states of character arise out of like activities. This is why the activities we exhibit must be of a certain kind; it is because the states of character correspond to the differences between these. It makes no small difference, then, whether we form habits of one kind or another from our very youth; it makes a very great difference, or rather *all* the difference.' (NE 1103a:30-1103b:25)

In other words, we acquire virtues by practicing them, just as when we learn to paint, or to play a musical instrument. No amount of acquired musical theory will turn even the most innately talented would-be pianist into a master in the absence of practice, and likewise, no amount of acquired ethical theory will make even the most innately talented would-be virtuous person virtuous. We can even push this analogy a little further and say that the fully virtuous person is to ethics as the virtuoso is to music. In order to become a virtuoso one requires natural talent *and* continual practice, and the same one suspects would be true with respect to virtue. To choose an extreme example, we know that there are individuals who are incapable of feeling any concern for others due to a neurological defect; they thereby lack any innate talent for virtue. It should not be surprising if others are particularly well-endowed 'by nature' with the appropriate qualities to develop virtue to the maximal extent (though exactly what these qualities may be we are in no position to say).

This analogy with the arts also allows us to see why it makes such a great difference to form the right habits 'from our very youth'. While there is no doubt that one can start learning to play the piano when one is, say, 30 years old, no one could expect to reach the same level of skill as they would have if they'd started practicing when they were a child. This does not mean that one cannot learn and improve at that age, and in the case of virtue we would certainly expect that an individual who realised the importance of virtue even at a late stage would still practice it because no matter one's current state it is always better to become more virtuous, but if they are realistic they will also realise that they cannot expect to master it, to acquire full virtue. Everyone can practice virtue, just as everyone can practice an art, and everyone can improve through practicing virtue in the same way; but

most people no matter how hard they practice will never reach the level of consummate and seemingly effortless mastery of an art, and the same goes for virtue. The fully virtuous person (if such a person exists) would be much like an artistic genius, a 'Da Vinci of virtue' one might say.

We can now see why Aristotle says in the passage quoted above that the virtues do not arise in us either by nature or contrary to it, that 'we are adapted by nature to receive them, and are made perfect by habit.' For just as we aren't by nature endowed with ability to paint (well) or play the piano (at all), so it is that nature doesn't endow us with true virtue. But just as we *are* endowed with the capacities that allow us to *learn* to paint (better) or play the piano (to some degree), so nature has endowed us with the capacities we need to learn virtue.² It just so happens that nature's endowments are showered on some people rather more liberally than on others. No one is born virtuous as Aristotle says, but as he also says, 'all who are not maimed as regards their potentiality of virtue may win it by a certain kind of study and care.' (NE 1099b) The virtues are 'made perfect by habit' in just the same way other skills are; with sufficient practice one reaches the state where they can be performed effortlessly, 'without thinking'. The better one is at something (anything, just pick a skill) the less one has to consciously think about what one is doing as one does it. At the highest peaks of excellence one no longer has to consciously think at all, one just knows what needs to be done and one does it (indeed this knowing and doing happen together, there is no gap between them). This suggests the modern psychological concept of 'flow' (also known as 'the zone'), the state of full immersion in an activity and the accompanied enjoyment in the process.

Finally, this analogy with the arts shows that virtue can only be acquired through the deliberate pursuit of virtue. One can never learn to paint well or play the piano 'by accident', or without intending to; one must choose to learn, and deliberately practice at it. Indeed, one can only learn by continually choosing to learn,

² This raises an interesting question: whether virtue is a human *invention* and, if so, what implications this has for virtue ethics. It is clear that the arts are human inventions (which is not to say they were deliberate inventions) that presumably use capacities that evolved for other purposes in order to create works of art; I suspect we can say the same of virtue. And this leads to another interesting thought: the purpose and value of art isn't the same for the artist as it is for the art lover, the one who creates does so for different reasons than the one who 'consumes'. We may well wonder, then, whether this difference of perspective can be fruitfully applied to virtue as well.

day after day; it requires commitment. One can practice at virtue from day one, but one may not acquire it until one has done a lot of practice; just as one can practice playing the piano from day one, but may not be able to play even the simplest tune recognizably for some time. This idea that virtue can only be acquired, even in part or to a limited extent, as the result of a deliberate commitment to practicing virtue will be crucial to responding to some aspects of the situationist critique.

Before moving on to discuss which traits are virtues, there is an important dis-analogy between virtue and artistic skill, which can fortunately be dealt with quickly. We have said that the virtues are character traits which are acquired in a similar way to artistic skill, yet we would hardly be inclined to say that a pianist possesses a piano-playing trait, nor, if asked what kind of person they were (as opposed to what they *do*), reply that they're a pianist. This contrasts with the virtues where it seems perfect natural to say of an individual that they possess courage, or they're a courageous person; so why the difference? In my view it comes down to the simple fact that character assessment is a social concern; when we ask about a person's character, what type of person he is, we want to know how he is likely to behave *because it affects us*. If we are satisfied that someone is a kind person this allows us to have some confidence in how they will behave towards us, if we are sure they are cruel we know we should avoid them if we can. The fact that someone can play the piano, while it may be pleasant, tells us nothing about how they are likely to engage with us. In fact, in an important sense, piano playing *in itself* isn't a social activity even if there are others present. This contrasts sharply with the virtues, which typically imply certain patterns of social engagement.³ One can even suggest the following: the virtues are habitual patterns of engaging well with others.

³ While it may seem that one can, for example, be kind or generous to oneself, to the extent this can actually possible I believe it is also a social act in that it requires us to think of ourselves 'as if' we were someone else, it requires a certain 'doubling' that lies at the very heart of reflexive thinking. Now, this reflexivity is critical to ethical development as I will discuss in the final chapter, but for now it suffices to note that there is no parallel with artistic skill.

1.2. Aristotelian virtue: specific aspects

So far we've established that the Aristotelian virtues are character traits which are: causative (in that their possessors 'tend to do noble deeds'); contribute to human flourishing or *eudaimonia*; and acquired by practicing 'like activities' (as one acquires the skill of piano playing by playing the piano). This is all very well so far as it goes, but it doesn't get us any closer to actually identifying any virtuous traits. Finding a means for identifying which traits are virtues has been a bugbear of virtue ethics since its rebirth, though there is some broad agreement. Courage and justice, for example, appear as virtues both in Aristotle's list and in those of the majority of contemporary virtue ethicists. On the other hand, Aristotle considers pride a virtue (true pride in his view being a combination of the virtuous person's belief that they are worthy of great deeds with the fact that they *are* worthy of them, as opposed to the vices of undue humility and vanity - *NE* 1123a: 35) to the quiet disdain of modern commentators (see for example Hursthouse 2001, chapter 1); while compassion or benevolence is the foremost virtue for most contemporary virtue ethicists (as well as some of their most vocal critics as we shall see), a trait notable in its absence from Aristotle's canon. Fortunately, it will not be necessary to come up with a complete list of virtues for the purpose of our present discussion as the situationist critique doesn't target individual traits, but all 'virtue-like' traits. However, it will be useful to see how Aristotle classifies the virtues, to see whether the traits the situationists do use to illustrate their point fit the mould.

Aristotle is (refreshingly) honest in repeatedly stating that it is difficult to determine in general terms precisely what is required of 'good conduct' (where 'good' should be interpreted along the lines of 'conducive to *eudaimonia*'), and that it is ultimately up to the individual to determine in each occasion what is required of them. However, employing the analogy with health once more he suggests that 'it is the nature of such things to be destroyed by defect and excess, as we see in the case of strength and health...So too is it, then, in the case of temperance and courage and the other virtues. For the man who flies from and fears everything and does not stand his ground against anything becomes a coward, and the man who fears nothing at all but goes out to meet every danger becomes rash' (*NE* 1104a:10-25, and note again that courage is on *everyone's* list of virtues, though they may differ in how much significance to attach to it). From this Aristotle concludes that

a virtue 'is a mean between two vices, that which depends on excess and that which depends on defect' (NE 1106b: 36 – 1107a:3, this is famously referred to as the 'golden mean'), and just as it is often easier to tell what ill-health consists in as opposed to health, so it may be easier to identify virtues by first examining vices (vices being considered as harmful character traits, deleterious behavioural habits acquired through practice).

So for example, one might note that alcoholism is a harmful habit, the result (at least initially, before it becomes a habit) of an inability to resist the pleasures of drink, and that it seems similar to drug-abuse and gluttony, so we can group them all together under the term 'self-indulgence' and classify it as a vice. Now, there is nothing wrong taking some occasional pleasure in drinking alcohol in most people's view, and certainly not with eating now and then (drug use though...?), so clearly never indulging could be a vice though it is not clear what to call it (Aristotle considers an individual who avoids all pleasures to be a boor, so boorishness will do). The middle ground between these two vices of self-indulgence and boorishness is the virtue of temperance, which is nothing more or less than the ability to take pleasure in food and drink and sex and the like without indulging to the point it does one harm.

The simple fact that one's behaviour appears to lie on the mean between two vices is not sufficient for one to count as virtuous in that respect, at least in Aristotle's eyes. 'Praise is appropriate to virtue' he says (NE 1101b:30), implying an individual deserves credit or recognition for their virtue, yet it makes no sense to praise a person who unwittingly performs an act appropriate to virtue, or who does so for the wrong reasons. Here, then, is another dis-analogy between virtue and the arts, for 'the products of the arts have their goodness in themselves', regardless of how they came into existence, while for an act to count as virtuous 'The agent...must be in a certain condition when he does them; in the first place he must have knowledge, secondly he must choose the acts, and choose them for their own sakes, and thirdly his action must proceed from a *firm and unchangeable character*' (NE 1105a: 25-35, my italics). It isn't enough to perform the 'correct' act in order to count as virtuous - and here virtue ethics reveals the importance it attaches to the internal state of the actor – one must choose the correct act knowingly, finding the required mean point, and perform it without any ulterior motive. Finally, and this is the critical point when we come to consider the situationist critique, the action must proceed from a 'firm

and unchangeable character' (though it may well be that this follows to some extent from the previous two points. Can one select and perform the correct act with knowledge and for its own sake 'on a whim'?). It would be well to keep this passage in mind throughout the proceeding discussion, for in order to correctly interpret (from a virtue ethical standpoint) the actions of the subjects of the experiments we will consider it will be necessary to consider not only *what* they did but also *why* they did it. If, for example, I was to donate money to charity so that others will like me then, regardless of any good that donation may do, I will not count as virtuous in that respect in Aristotle's eyes. Indeed, if anything I would be exhibiting the vice of vanity (and possibly others besides) in that I'm trying to elicit praise and admiration that I do not deserve. This concern with the motives lying behind an individual's acts accords well with everyday character assessment. We're all familiar with people who deliberately cultivate a certain image of themselves in certain circumstances while putting the lie to it in others (such as the gossip who is all smiles to your face then slanders you behind your back). We have good reason to be concerned about the motives of precisely those who seem to be acting virtuously; at least an openly vicious act wears its viciousness on its sleeve.

This need to 'have knowledge' in order to act virtuously leads to a consideration of a rather special virtue, one which does not consist of a mean between two vices and is not properly a character trait but rather the capacity to discern what virtue requires in a particular situation, the ability to locate the mean; this is the intellectual virtue of *phronesis*, or 'practical wisdom'. To possess the virtue of liberality, for example, it is not sufficient to be disposed to dispense one's money for worthy ends (there are so many worthy), and certainly one would not count as liberal if one gave all one's money to the first person who asks for it. Rather, it is necessary to make the best use of it, to weigh up the pros and cons (what if one gives away all the money one can spare to a friend in need, and then encounters another friend in greater need?), to make an informed, reasoned *decision*, the best decision one can on the basis of the information at hand. The difficulty in determining the best course of action in light of imperfect knowledge *even when our motives are pure* is reason why Aristotle repeatedly emphasizes the rareness of virtue. It 'is no easy task to be good' he says, 'For in everything it is no easy task to find the middle...anyone can get angry – that is easy – or give or spend money; but to do this to the right person, to the right extent,

at the right time, with the right motive, and in the right way, *that* is not for everyone, nor is it easy; wherefore goodness is both rare and laudable and noble' (NE 1109a: 24-29).

This difficulty in locating the mean points to the need for practical wisdom. To return to the medical analogy it isn't enough to want to cure illness in order to become a good doctor, one must also be able to diagnose, to discern what the illness is in order to treat it. Practical wisdom is the capacity to diagnose a situation, to discern which virtue or virtues are relevant and which course of action is in accordance with them; 'virtue makes the goal correct, and practical wisdom makes what leads to it correct' (NE 1144a: 8-9), or to put it another way, virtue provides the end and practical wisdom discerns the means to that end.⁴ Again, as we move forward and away from Aristotle to consider contemporary virtue ethics and the situationist critique thereof it will be worth remembering how difficult Aristotle considered virtue to be. Even when people have the right motivations they may fail to behave virtuously in a reliable fashion because they lack practical wisdom. There are therefore two ways a person may fail to act virtuously: they may lack the requisite character virtue, i.e. they may fail to act compassionately because they aren't compassionate; or they may lack practical wisdom, i.e. they are compassionate but they fail to recognize the situation calls for compassion, or recognize what the compassionate act would be in this particular situation.

By the above account virtue in the Aristotelian sense seems eminently possible, or at least not obviously impossible, but this is a far cry from saying that ethics as such just is virtue ethics in this sense. Fortunately there is no need for me to address that issue, and there is also no need to come up with a definitive list of virtues (any such list would be risk courting controversy or merely repeating contemporary platitudes). For my purposes it will be sufficient to note that any character trait that can be conceived of as a mean between two vices can provisionally serve as a virtue; if the situationists can throw doubt on the existence of traits *of that kind* then this would deal a serious blow to any virtue ethicist who seeks to follow in Aristotle's footsteps, irrespective of whether the trait in question should

⁴ In *Practical Intelligence and the Virtues* (2009), Daniel Russell makes the argument that practical reason is required in specifying the end as well. The virtues provide an end in a general sense (be generous, be courageous, be just...), and engage in a dialectic with practical reason to determine the specific end the virtues require in each particular instance. I find Russell's argument compelling, but the simple gloss I've given above will be sufficient for the purposes of this paper.

actually be considered a virtue. With the Aristotelian background now established I will move on to an outline of the characterisation of virtue in contemporary virtue ethics.

1.3. Virtue here and now: contemporary formulations

One drawback to Aristotle's ethical writings, at least to modern eyes, is that it seems to lack a firm structure, a clear progression and formulation of ideas. He often repeats thoughts on the same topic in different sections of *The Nicomachean Ethics*, reformulates the same ideas in different ways, gets side-tracked (or so it seems to me), and introduces numerous caveats and qualifications. Now this may have been unavoidable in his view, given his repeated statements to the effect that ethical discussion is necessarily piecemeal, that it is difficult to make statements that have both a general application and are unreservedly true (see, for example *NE* 1107a: 30-32: 'among statements about conduct those which are general apply more widely, but those which are particular are more true'), but it does make it difficult to find a concise, coherent characterization of the virtues that can form a firm starting point for further debate. Fortunately, if there is one thing philosophers in the analytical tradition excel at it is distilling the essence of another philosopher's work into the most concise and logically coherent form possible. So let us see how they've done with their analysis of Aristotle.

The crucial concept from the point of view of the situationist critique is the concept of a character trait, at least as it is employed in virtue ethics. Fortunately, there is a broad agreement on how to interpret this critical concept, both within virtue ethics and by their situationist opponents. For example, in the words of Rosalind Hursthouse, probably the most prominent of contemporary virtue ethicists: 'The full Aristotelian concept of virtue is the concept of a complex character trait, that is, a disposition that is well entrenched in its possessor and, as we say, "goes all the way down" ... This is because your virtues (and your vices) are a matter of what sort of adult you are, and involve, most particularly, your values' (Hursthouse in Dreier 2006, p. 101-102). This idea that virtues are dispositions appears to go back to Aristotle (c.f. *NE* 1106a: 4-6: 'in respect of the passions we are said to be moved, but in respect of the virtues and vices we are said not to be moved by to be disposed in a

particular way), and is more or less universally accepted by both sides in the current debate, so it will be discussed in more depth presently.

The requirement that the disposition go “all the way down” relates to Aristotle’s statement that a virtuous act must proceed from a ‘firm and unchangeable character’. Virtue is a property of persons, not of acts, the descriptor ‘virtuous’ only applies to acts in a derivative sense. A virtuous act is simply an act performed *as the virtuous person would perform it* (c.f. *NE* 1105b). Suppose, for example, that I am in a situation where a virtuous person would speak the truth, that is, a situation that requires ‘honesty’.⁵ And suppose I do tell the truth, but I do so merely because I am afraid that I will be caught out if I lie. Then I haven’t performed the act for its own sake (I did it for myself) and I didn’t perform it from ‘a firm and unchangeable character’ either (I wanted to lie), so while in some sense I performed the ‘correct’ act it does not count as virtuous, and if anything it is vicious in the sense that I acted out of cowardice. With the virtuous person there is no conflict between what they want to do and what they choose to do, their desires and reasons for acting are integrated, and it is precisely this that ensures that a virtuous person can be relied on to act virtuously.

One of the great advantages of this account of virtuous action is that it resolves the problem of ethical motivation. While ‘some moral theories start by identifying a moral rule, such as “Help people in need” and then have to explain how and why a person can be motivated to follow the rule...virtue ethicists don’t need to give a separate account of moral motivation. A virtue is a disposition that expresses itself in acting, reasoning, and feeling in certain ways. A brave person is not someone who has learnt about bravery, decided he should be brave, and then needs to find a motivation to act bravely. Instead, he is someone whose character tendencies have been formed in such a way that he acts, reasons, and reacts bravely’ (van Zyl in Besser-Jones & Slote 2015, p. 190). This holistic picture of virtuous action where reason and emotion operate in an integrated fashion contrasts with the schizoid picture which treats reason and the emotions as necessarily separate and entirely

⁵ Honesty is considered a virtue by many contemporary virtue ethicists including Hursthouse, though interestingly it does not appear in Aristotle’s catalogue of virtues. It seems clear that it could be a virtue in his sense however, virtuous honesty being a mean between the vices of deceitfulness (deficiency) and, say, tactlessness? To *always* tell the truth regardless of the circumstances and the nature of the truth is harmful to both the teller and those around them. Striking the right balance is tricky and I suspect whole papers could be written on this, but then that is the nature of virtue in Aristotle’s eyes.

different in kind, a view prominent enough in the Age of Reason but still common today; we will see traces of this in the situationists too. So often it seems ethics is presented as a matter of doing ‘the right thing’ regardless of how we may feel about it; virtue ethics on the other hand would have us train our emotions so that we want to do what we should. But this is hard, it requires discipline, and it throws into doubt our familiar picture of human agency and responsibility, where as a result of our having ‘free will’ we are able to take full credit when we ‘choose’ to do good, and we can attribute blame to others when they ‘choose’ to do wrong (though when matters are reversed we are often more willing to find mitigating factors to explain away our own failings, and attribute the success of others to fortuitous external factors for which they deserve no credit).

Turning to the situationists now, we will see that their interpretation of the concept ‘character trait’ as employed in virtue ethics essentially coincides with those quoted above. Gilbert Harman, the original proponent of the situationist critique, describes character traits as ‘relatively long-term stable disposition[s] to act in distinctive ways...The relevant dispositions must involve habits and not just skills, including habits of desiring...A person with the relevant character traits has a long term stable disposition to use the relevant skills in the relevant way’ (Harman 1999, p. 317). Harman refers to the long-term stability of such character traits their “robustness”, an apt term that has become the standard. John Doris, who wrote the first full-length book dedicated to the situationist critique, prefers to refer to the target of his critique as ‘global’ character traits, which he characterises as possessing the following three properties:

- 1) *Consistency*. Character and personality traits are reliably manifested in trait-relevant behavior across a diversity of trait-relevant eliciting conditions that may vary widely in the conduciveness to the manifestation of the trait in question.
- 2) *Stability*. Character and personality traits are reliably manifested in trait-relevant behaviors over iterated trials of similar trait-relevant eliciting conditions
- 3) *Evaluative integration*. In a given character or personality the occurrence of a trait with a particular evaluative valence is probabilistically related to the occurrence of other traits with similar evaluative valences. (Doris 2002, p. 22-23)

When taken together, these points 'construe personality as more or less coherent and integrated with reliable, relatively situation-resistant, behavioral implications' (Ibid, p. 23). Consistency and stability are related once again to Aristotle's notion of virtuous action proceeding from 'a firm and unchangeable character', and so Doris' characterisation is in accord with Hursthouse's when she says, for example, that 'people's virtues and vices...are strongly entrenched, precisely because they involve so much more than mere tendencies to act in certain ways. A change in such character traits...can happen slowly, but on the rare occasions when it happens suddenly, the change calls for special explanations – religious conversion, an experience that changes the person's whole outlook on life, brain damage, or drugs' (Hursthouse 2001, p. 12). Now that the stage is set, and the positions of the players known, it is time, finally, to give a preliminary expression of the situationist critique.

Chapter 2 – Situationism: The Argument and the Evidence

As indicated at the beginning of the previous chapter, the situationist critique is a form of error theory with respect to virtue (and indeed of vice also) as it argues that we are systematically mistaken in attributing such traits, either because:

1. There are no such things as character traits - humans are not the kind of beings who can possess such traits (Harman 1999); or,
2. We have no good reason to believe in such things – we have no empirical evidence for their existence, and can explain our belief in such traits without invoking their existence (Harman 2009, Doris 2002); or,
3. The vast majority of human beings do not possess such traits, and it is unlikely or impossible that they will be able to develop them (and perhaps we wouldn't want them anyway) (Doris 2002).

In any case, the situationist critique is fundamentally a critique of the 'psychology of virtue'. It argues that social psychological studies show humans are not the way virtue ethics requires them to be. More specifically, these studies show that whatever traits humans do have are not 'robust', or in Doris' terms, they are neither consistent nor particularly stable. It follows that either virtue ethics is untrue or it is unscientific, and for an ethical theory that prides itself on its commitment to naturalism this latter would be a bitter pill indeed. So if the situationists are correct, virtue ethics fails.

2.1. The Situationist Argument

In order to assess the situationist critique we'll need to evaluate the evidence, but before considering the evidence it will be helpful to have in mind the basic outline of the argument they are making, so we can see how the evidence supports it. The most prominent situationist is Gilbert Harman, who pioneered the application of situationism as a social psychological paradigm to the analysis of the psychology of virtue ethics. Following Harman, John Doris produced the first book-length work dedicated to the situationist critique, providing a far more thorough argument and analysis of the available evidence. Since Harman and Doris agree on all the essential details I will employ Doris as the foremost

authority on the situationist critique, appealing to other situationists as necessary to demonstrate their agreement, or when they raise a relevant point that Doris doesn't cover.

As stated previously, the situationist critique is a critique of the 'moral psychology' of virtue ethics', where moral psychology is understood as the investigation of 'psychological properties of moral agents' (Doris 2002, p. 3. I will refer to the 'psychology of virtue ethics' rather than to its 'moral psychology' for reasons I will explain in the final chapter). Given the commitment of virtue ethicists to naturalism – where naturalism can be understood as a minimal requirement that an ethical theory not contradict our best scientific theories – any argument that successfully undermines the psychology of virtue ethics will be fatal. And as Doris rightly notes, the language of virtue ethics is replete with psychological terminology. Aristotle, for example, speaks not only of character traits, but also discusses emotions, desires, deliberation, the nature of choice; and the list goes on. So it is hard to argue with Doris' claim that 'Such talk easily admits of – indeed, cries out for – empirical evaluation... At this point it is enough to notice that central commitments of character ethics are very naturally understood descriptively; the empirical investigation I advocate is one the tradition invites' (Ibid, p.6). Given that we are now in a position to study such claims *scientifically* it would be inexcusable for virtue ethicists to simply ignore the implications psychological studies may have for virtue ethics, and if nothing else Doris deserves credit for throwing down the gauntlet. For the record, my commitment to naturalism in this context is absolute; *if* psychological studies falsify the psychology of virtue ethics, *then* I will consider virtue ethics as a whole to be refuted. But this is a big if, and the situationists are well aware that 'folk psychology' is committed to talk of character traits (and given that such talk isn't obviously false, the situationists require a lot of evidence to be convincing). Fortunately (for them) Doris rises to this challenge, and a great deal of the value of his 2002 book *Lack of Character* lies in that fact that he collects and evaluates all the major studies that situationists have appealed to in support of their critique.

The reason the situationist critique is such a threat is that it targets the very heart of virtue ethics: the virtues themselves. Recall that the virtues are character traits, and character traits are almost universally interpreted in a dispositional sense. In Doris' words, 'to attribute a character or personality trait is to say, among other things, that someone is disposed to behave in a certain way in certain eliciting conditions' (Ibid, p. 15), so for

example, a compassionate person would be disposed to act compassionately in response to certain situations (say when they encounter a lost child).

Why do we say a compassionate person is one who is *disposed to act* compassionately rather than simply saying that a compassionate person is one who *does* act compassionately in certain situations? First, this allows for the possibility that a person might experience internal conflict, for example a person may be compassionate and just, and sometimes justice may require that they not act compassionately. Rather than say that this person is not compassionate because they didn't act compassionately, it seems better to say that they *are* compassionate (to some degree) *because* they were disposed to act compassionately, even though they didn't; at least so long as they had a virtuous reason not to act compassionately. Secondly, and relatedly, to speak of dispositions allows for the element of choice. The virtuous person is not a mere automaton who robotically responds to eliciting conditions with a corresponding action; rather they choose to act on certain dispositions and not on others.⁶ While there is an ongoing debate in the philosophical literature on exactly what dispositions are, one which shows no signs of reaching any sort of consensus, it suffices to note that dispositional language is widely used by virtue ethicists when describing the virtues, including Hursthouse and of course Aristotle himself. Moreover, there is nothing riding on the term 'disposition.' The crucial characteristic of virtuous traits for this discussion - their weak point from the point of view of the situationists - is their 'robustness'.

Robustness is related to Aristotle's requirement that virtuous action proceed from 'a firm and unchangeable character'. A person with a robust trait 'can be confidently expected to display trait-relevant behavior across a wide variety of trait-relevant situations, even where some or all of these situations are not optimally conducive to such behavior' (Ibid, p. 18). So an honest person, in the virtuous sense, can be relied on to be honest even when it is difficult to be so. A person who was only honest when it served their interests, or was honest with their friends but not with strangers, would not count as virtuously honest; indeed, practically everyone would be honest in those situations.

⁶ Aristotle makes it clear that the courageous person still feels fear, and so may feel a disposition to flee from danger for example. But if virtue requires it they will not act on this disposition. Courage is admittedly a rather unusual virtue in that it seems to consist precisely in this ability to overcome fear that may be intent on realising one of the other virtues, but the lesson seems to apply to the others too.

This compound notion of a character trait as ‘robust disposition’ immediately suggests a possible research project because the notion of a ‘metaphysically robust disposition underwrites the prediction and explanation of behavior’ (Alfano 2011, p. 122). Someone who has such a disposition can be relied upon to act on it in a broad range of trait-eliciting conditions, so we can surely test whether people do have such a trait by observing how they *do* act in a range of trait-eliciting conditions. What we will be looking for in particular are what Doris calls *diagnostic situations*, situations that are ‘unfavorable enough to trait-relevant behavior that such behavior seems better explained by reference to individual dispositions than by reference to situational facilitators’ (Doris 2002, p.23). While this puts the burden of proof on the virtue ethicists this seems fair in that they are the ones making a positive existence claim; the central situationist claim is a negative one and so they have parsimony on their side.

The situationists do not actually engage in this research project themselves because they believe that the evidence is already there, that it has already ‘been demonstrated through a host of celebrated laboratory and field studies...that manipulations of the immediate social situation can overwhelm the importance of the type of individual differences in personal traits or dispositions that people normally think of as being determinative of social behavior’ (Ross & Nisbett 1991, p. xiv), or, in Doris’ terminology, ‘*globalist conceptions of personality are empirically inadequate*’ (Doris 2002, p. 23). In order to assess this claim I will need to turn to the evidence, but before I do I’ll briefly discuss the purported character trait that is central to the two main experiments I’ll consider: compassion.

I noted in the first chapter that compassion appears in most contemporary virtue ethicist’s list of virtues (although Michael Slote bucks the trend by listing benevolence instead), whereas it is notably absent from Aristotle’s list. One could, if one were so minded, argue that compassion is not a virtue, it is only a feeling. In that case any experiments showing that compassion isn’t a robust trait shouldn’t come as a surprise, and they certainly wouldn’t have any implications for virtue theory. Now I do not think this is an adequate response, and not because it seems all too easy (sometimes there are easy answers, and the mere fact that Doris wrote a whole book on this subject does not preclude the possibility that he made a fundamental error at the very start), but because we can distinguish

between compassion as an emotion and virtuous compassion as a possible trait, in the same way we often distinguish between love as an emotion and a more robust notion of love that involves a deep commitment to a loved one that enables us to reliably look out for their best interests regardless of our current emotional state, which can only ever be transitory. It is at least feasible that compassion as a virtue can be conceived of as a mean between two vices, say callousness (the vice of defect, signifying a total lack of concern or feeling for others) and, for want of a better term, excessive compassion (vice of excess, where one feels compassion to an excessive degree, either to the point that one is so caught up in the needs and sufferings of others that one does not take care of oneself, or to the point that it interferes with the requirements of other virtues such as justice). So I am willing to concede that compassion in this sense may be considered a virtue, or at least a 'virtue-like' trait, and accept Doris' contention that 'compassion the character trait is a stable and consistent disposition to perform beneficent actions; failures to behave compassionately when doing so is appropriate and not unduly costly are evidence against attributing the trait' (Doris 2002, p. 29).

2.2. The Evidence

The evidence the situationists appeal to in support of their critique is of two kinds: experimental and historical. The experimental evidence consists of the results of a variety of social psychological experiments, including some of the most well-known and influential studies in that field, and it is the experimental evidence that situationists lean on most heavily for support. However, the historical evidence, and specifically Doris' investigation of the Holocaust, while generally considered of secondary importance is significant in its own right as it shows how the situationists apply their interpretation of the experimental evidence. Such real-world case studies also allow them to side-step some of the criticisms of the situationist critique we will consider in the next chapter. So I will consider the experimental evidence first and in greater depth, outlining the interpretation the situationists give of that evidence, and then briefly show how they apply that interpretation to the historical cases.

The experiments the situationists typically appeal to in support of their critique of virtue ethics (or ‘character ethics’ to employ Doris’ more inclusive term) are usefully summarised by Alzola as ‘a large collection of experiments in social psychology ranging from the infamous Milgram experiment on obedience to authority (1974) to the Stanford prison experiment by Zimbardo (1974), and including empirical research on mood effects (Isen and Levin, 1972) and group effects (Latané and Darley, 1970) and the Good Samaritans studies (Darley and Batson, 1973)’ (Alzola 2008, p. 343).

While the Stanford prison experiment is arguably the most well-known experiment in this collection the situationists rightly do not attach much weight to it, for although the results were certainly shocking and in line with their general thesis the experiment was never repeated (for good reason) and it is debatable how representative it could be given that all the subjects were young male college students. The experiments on mood and group effects are interesting in their own right, and I will discuss them briefly, but it is the Good Samaritan and especially the Milgram experiments which provide the most critical evidence for the situationist critique, as is clear from the even the briefest survey of the literature. So while Doris boldly claims that his ‘argument does not stand or fall with an experiment or two’ (Doris 2002, p.13) I hold that this is mere bluster, that the amount of weight he attaches to Milgram in particular is abundantly clear, and that if it could be shown that neither the Milgram nor the Good Samaritan experiments could be relied on to support the situationist critique then the argument would at the very least be brought to its knees. So I will devote much of the rest of this chapter to presenting those experiments in some detail, and discussing the inferences situationists draw from the results, before turning a critical eye to the experiments and interpretations in the next chapter.

2.2.1. Milgram

The Experiment

The Milgram experiments on ‘destructive obedience’ are continually referred to by the situationists as the best evidence in favour of their critique of virtue ethics, and there are a number of good reasons for this. First, the results seem sufficiently counter-intuitive to

raise serious questions about our understanding of human moral psychology in general, and it is notable in this respect that neither Milgram nor a survey of lay-people came close to predicting the actual results (c.f. Doris 2002, p. 49). Second, the experiments have been repeated numerous times with broadly similar outcomes, and have included both male and female subjects of a various ages and from a range of ethnic, cultural and socio-economic backgrounds; so it would seem the results are repeatable and representative of the population at large. This contrasts with the Stanford Prison experiment where both the repeatability and representative nature of the results can be called into question. Third, Milgram ran the experiment through a number of iterations, varying the experimental arrangement so that, while the subjects were required to perform essentially the same task in each iteration their situation varied. This last aspect of the Milgram experiments is absolutely critical to deriving the situationist conclusion.

In order to appreciate the situationist argument it will be necessary to consider the Milgram experiments in some detail. The following is a summary of the account Doris gives on pages 39 to 46 of *Lack of Character*:

There were three participants to the experiment: the subject, the experimenter, and a 'confederate'; that is, an individual who pretends to be a subject but is really 'in on it'. The subjects were all adults from a wide variety of backgrounds and professions. At the outset the subject is met by the experimenter who introduces them to the other 'subject' (i.e. the confederate) and explains the experiment is to test the effects of punishment on learning. They then draw lots to determine who will be the 'learner' and who will be the 'teacher', with the lot being fixed so the subject is always the 'teacher'. The learner, an 'affable middle-aged accountant' (it is obviously important that the learner be affable to avoid any suggestion teacher may be acting out of malice), is then strapped into a chair "to prevent excessive movement" and an electrode is attached to his wrist with an electrode paste "to avoid blisters and burns". The teacher is administered a mild shock to convince him of the scenario's authenticity; however, the 'shocks' that will be administered to the learner are fake.

The teacher and experimenter retire to an adjoining room and the teacher proceeds to administer a word-association test, asking questions in a pre-arranged sequence while

experimenter sits behind them and observes. Each time the learner gets a question wrong the experimenter instructs the teacher to administer a shock as punishment. With each wrong answer the voltage is increased, the voltages being marked on the dial at 15 volt increments with signs ranging from “Slight Shock” to “Danger: Severe Shock” and ultimately simply “XXX”. With each shock a pre-recorded response by the learner is played to convince the teacher that the effects are real; as the experiment progresses the learner progresses from calmly evincing mild pain, to alarm as the ‘shocks’ increase in strength to the point that the learner refuses to continue with the experiment when the level reaches 150 volts. If the teacher continues past this point (as they are instructed to) the learner begins writhing in pain, screaming and begging for the experiment to stop before ceasing to respond at 330 volts. The experiment doesn’t necessarily stop there however, as the teacher is instructed to treat no response as an incorrect answer and to continue to administer shocks and increase the voltage up to maximum of 450 volts. If the subject expresses concern about the procedure at any stage (and, unsurprisingly, most did) the experimenter replies with a standard set of responses: (1) “Please continue,” (2) “The experiment requires that you continue,” (3) “It is absolutely essential that you continue,” (4) “You have no choice, you *must* go on.” At all times, according to Milgram, the experimenter’s tone of voice is “firm, but not impolite.” If the subject still objects after number 4 they are termed ‘disobedient’ and the experiment is terminated; if the subject continues administering shocks to the maximum of 450 volts they are termed ‘obedient’.

So, that is the Milgram experiment. As for the results, the table below shows the outcomes from a standard Milgram experiment involving 40 test subjects:

Voltage level	Number disobedient
Slight Shock (15-60V)	0
Moderate Shock (75-120V)	1
Strong Shock (135-180V)	7 (6 at 150V when learner expressly refuses to continue)
Very Strong Shock (195-240V)	0
Intense Shock (255-300V)	3
Extreme intensity Shock (315-360V)	2 (No learner response after 330V)
Danger: Severe Shock (375-420V)	1
XXX	0

Total disobedient	14
Total obedient	26
Percentage obedient	65%

Rather than the 1-2% obedience rate predicted by participants in Milgram's post-experiment surveys, a full 65% of subjects in this experiment were fully obedient. Not blindly obedient by any means, rather what was notable was how conflicted the subjects were. Obedient subjects were often observed to "sweat, tremble, stutter, bite their lips, groan, and dig their fingernails into their flesh" (according to Milgram, 1963); but they still obeyed. The results appear to show a disturbing lack of compassion among a majority of the test subjects, and not just among the 65% who were fully obedient, for let us not forget that a further 18% of the subjects proceeded to administer shocks past the point where the learner expressly refused to continue. At first glance this certainly seems to call into question any character-based ethical theory; 'can we really attribute a 2 to 1 majority response to a character defect?' Harman asks rhetorically (1999, p. 322).

Of course, these results don't demonstrate the cross-situational inconsistency situationists rely upon to draw their conclusion. To show this they appeal to other iterations of the experiment that vary the set-up, the results of a few of which are summarised below:

- When subjects were free to choose the shock levels to administer to the victim, only 3 percent delivered the maximum shock (Milgram 1974: 61).
- When the experimenter was physically absent and gave his orders by phone, obedience was 21 percent (Milgram 1974: 60).
- In a "touch-proximity" condition where the subject was instructed to press the victim's hand onto a "shock plate" to administer the punishment, obedience was 30 percent ((Milgram 1974: 35).
- When a confederate "peer" administered the shock while the subject performed only subsidiary tasks such as administering the test, obedience was 93 percent ((Milgram 1974: 119).

There can be no denying that these results demonstrate considerable variation in the proportion of people willing to engage in the same task (i.e. apparently torturing an innocent, affable stranger to the point of death) depending on such situational differences as whether the person giving the instructions was in the room with them or not. Especially when we consider as a baseline that only 3% of subjects chose the maximum voltage when

they were free to do so (and indeed, one might suspect that in this case the subject knew the experiment was a fake) the fact that the physical presence of the experimenter made so much difference is certainly disturbing; but does it really support the situationist claim that it should undermine our belief in robust character traits?

Interpretation

The situationist interpretation of the Milgram experiments is usefully summarised by Doris as follows: ‘Milgram’s experiments show how apparently noncoercive situational factors may induce destructive behavior despite the apparent presence of contrary evaluative and dispositional structures. Furthermore, personality research has failed to find a convincing explanation of the Milgram results that references individual differences. Accordingly, Milgram gives us reason to doubt the robustness of dispositions implicated in compassion-relevant moral behavior; his experiments are powerful evidence for situationism’ (Doris 2002, p. 39). The inference is clear, and given the strong results of the experiments it certainly seems plausible on the face of it: if compassion were a robust trait we would expect the subjects to act compassionately regardless of apparently minor details as whether the experimenter who was instructing them to administer the electric shocks was in the room with them, rather than issuing instructions over the phone. Yet, to contrast the results of the original experiment with the results of one of the variants, over 40% of experimental subjects could be expected not to act compassionately for precisely this reason. The results are even more striking when we compare the baseline of 3% with the obedience rate of 93% when the subject ‘merely’ administered the test and instructed another confederate to administer the shock. Surely if one is not willing to freely administer a 450 volt shock to an innocent stranger one should not instruct another person to do so simply because one is asked; yet 90% of subjects were apparently prepared to do just that.

Ross and Nisbett provide the following summary of some of the factors they believe led to the striking results: ‘the stepwise character of the shift from the relatively unobjectionable behaviour to complicity in a pointless, cruel, and dangerous ordeal’, ‘the difficulty in moving from the intention to discontinue to the actual termination of their participation’ and the fact that ‘the events that unfolded did not “make sense” or “add

up”...The subjects’ task was that of administering severe electric shocks to a learner who was no longer attempting to learn anything’ (Ross & Nisbett 1991, p. 56-58).

We should be careful, therefore, in how we construe the behaviour of the obedient subjects. Doris’ gloss is that ‘the majority of the subjects were willing to torture another individual to what seemed the door of death without any more direct pressure than the polite insistence of the experimenter’ (Ibid, p.42), but he gives the lie to this himself when he goes on to speak in the same paragraph of the ‘horribly conflicted’ nature of the subjects’ obedience. The subjects (at least the vast majority of subjects) were by no means *willing* to obey; they clearly would’ve preferred not to, *yet they obeyed anyway*.⁷ It is this obedience in spite of a clear disposition not to obey that requires explanation, and part of this explanation will have to address the uncertainty the subjects faced, as evinced by the Ross and Nisbett passage quoted above. The situationist explanation is that the subjects’ obedience demonstrates that compassion is not a robust trait; we shall see how their opponents respond in the next chapter.

2.2.2. Good Samaritans

Turning now to the Good Samaritan experiment we will see that it reinforces the situationists’ argument by providing more evidence of significant situational variation, while also differing from the Milgram experiments in a couple of key respects that will help to buttress the situationist case against some of the criticisms we will consider presently. The following summary of the experiment and its results comes once again from Doris.

The Experiment

In 1973 John Darley and Daniel Batson invited students at the Princeton Theological Seminary to participate in ‘a study of “religious education and vocations.”’ The students were first asked to fill out questionnaires in one building before reporting to a nearby

⁷ Doris rightly notes, referencing Milgram, that this conflicted obedience is the best evidence for the ‘experimental realism of his paradigm’ (Doris 2002, p.43). Those who wish to claim that the majority of Milgram’s subjects must have known the shocks were fake will have a hard time explaining away these physical signs of internal conflict.

building where they were required to give a short verbal presentation. Upon completion of the questionnaire, subjects 'were told either that they were running late ("high hurry" condition), were right on time ("medium hurry" condition), or were a little early ("low hurry" condition): thus the conditions exerted a different degree of time pressure on the subjects. The behavior of interest occurred on the walk between the two sites, when each seminarian passed an experimental confederate slumped in a doorway, apparently in some sort of distress' (Doris 2002, p. 33-34).

In the original experiment the subjects were assigned a score on a scale from 0 to 5 (0 indicating a total failure to notice the person in distress, 5 indicating the subject provided assistance and refused to leave the distress person until further help arrive), but for our purposes it is sufficient to consider simply the percentages of subjects that provided some assistance. The results are summarised in the table below:

	Degree of Hurry		
	Low	Medium	High
Percentage helping	63	45	10

Interpretation

Once again we see a high degree of situational variance, especially for those subjects who thought they were running late. But what is notable in this case, in contrast to the Milgram experiments, is that the subjects were not acting under orders not to assist. This was not a case of 'destructive obedience' and cannot therefore be explained by some kind of innate susceptibility to following orders given by an authority figure. Moreover, the subjects were completely unaware that they were participating in an experiment, again contrasting with the Milgram case where the subjects were well aware they were taking part in an experiment (albeit not of the kind they anticipated). These differences will support the situationist case against suggestions that the Milgram experiments are anomalous and not representative of typical human behaviour.

Employing the results of the questionnaire Darley and Butson found no correlation between individual subjects' religious and moral beliefs and their willingness to help; the only variable that was predictive was the degree of hurry (c.f. Harman 1999, p. 323). Of

course, it is no surprise that, all things being equal, people in a hurry are less likely to provide assistance. What is surprising is 'the apparent disproportion between the seriousness of the situational pressures and the seriousness of the omission... the demands of punctuality seem rather slight compared with the ethical demand to at least check on the condition of the confederate' (Doris 2002, p. 33-34). It is this disproportionate response to 'situational pressures' that is so striking, and that forms the core of the evidentiary case for the situationist critique. Neither situationists nor virtue ethicists expect people to be totally unresponsive to situational variables, in all these cases (and in the vast majority of the real-life and hypothetical scenarios ethicists typically considered) the agent is responding to a situation and should be sensitive to the details of the situation if they are to act correctly. But situationists and virtue ethicists agree that behaviour should, insofar as it is ethical, respond only to ethically relevant situational factors, and it is hard to deny that one's degree of hurry does not seem especially important when compared to a stranger's health.

2.2.3. Mood and Group Effects

I will only briefly discuss some of the other experiments situationists regularly appeal to as they play a more marginal role, their main contribution being to provide some insight into the mechanisms that may explain situational variance. These consist of a number of experiments on mood and group effects.

In a 1972 study Isen and Levin discovered that subjects who found a dime in a pay-phone coin return slot were far more likely to engage in 'prosocial behaviour' (which in this case involved helping a confederate pick up some papers she'd dropped) than subjects who didn't; overwhelming so in fact: 14 of the 16 subjects who found a dime helped the confederate, while only 1 of 25 subjects who didn't find a dime helped her. Isen and Levin's explanation was that the 'small bit of good fortune elevates mood, and "feeling good leads to helping"' (Doris 2002, p. 30). Now this seems plausible enough, and the results are certainly striking, so we would be well advised to consider mood effects as a possible explanation for situational variance in observed behaviours. Yet what is not clear is that this has any relevance to a critique of virtue ethics. Virtue ethics does not require that *all* behaviour proceed from 'a firm and unchangeable character', only that virtuous behaviour

must; and it is not clear that virtue is called for here, since the 'prosocial behaviour' consisted merely in helping a confederate pick up a folder of papers. So I will watch for the possible influence of mood effects, but I do not consider this experiment provides any evidence either for or against the existence of compassion as a robust character trait.

The experiments on the group effect on the other hand have more relevance, group effect being a more general term for what became known as the 'bystander effect' after the famous Genovese case where a young woman was stabbed to death in New York City on 13 March 1964, reportedly with 38 witnesses who either saw or heard the attack yet failed to intervene or call the police (this account was published in *The New York Times*, and has since been called into question). The effect itself is quite real however and demonstrated in a number of experiments, a simple description would be that individuals are more likely to act on an environmental cue if they are alone than if they are with a group of other individuals who remain passive. The usual interpretation is that if an individual is in a group they will typically look to other members of the group for cues, if the other members of the group remain passive this will increase the chance they will remain passive also. Given that the group effect isn't counter-intuitive I will only present one experiment relating to it, as the explanation the experimenters give for the observed behaviour will be relevant when it comes to evaluating the situationist critique.

In 1970 Latané and Darley 'asked students to participate in a group discussion of the problems faced by college students in an urban environment. The ostensible "discussion" proceeded by intercom with the experimenter absent and the subject isolated in a cubicle, ostensibly to preserve anonymity; in fact, the other "participants" were tape recordings, and the situation was designed to address a variant of the group effect. One tape-recorded participant described his difficulty with seizures; he later gave an arresting impression of someone suffering a seizure...100% of subjects believing themselves alone with the seizure victim intervened, while only 62 percent of subjects in a "group" consisting of subject, victim, and five more tape-recorded participants did so' (Doris 2002, p.33). The experimenters determined that the 'inhibiting mechanism consisted at least partly in a "diffusion of responsibility": The presence of others meant that no individual was forced to bear full responsibility for intervention. When the experimenter terminated each trial after 6 minutes, unresponsive subjects in group conditions appeared aroused and conflicted'

(Ibid). This notion of 'diffused responsibility' is an important one and I will return to it in due course, aside from this the experiment provides more evidence for the power of situational variables on patterns of individual behaviour.

2.2.4. Historical studies

While the social psychological experiments discussed above undoubtedly form the evidential core of and primary motivation for the situationist critique, both Harman and Doris appeal to real-world historical cases to support their positions. Once again it is Doris who does so in greater depth, so I will begin with his study on the Holocaust and then refer to Harman to demonstrate the parallels between the two. I will merely present a summary of the cases they make and return to critically examine them in the next chapter.

The Holocaust

Doris begins his discussion on the Holocaust by acknowledging the difficulties in trying to analyse such a complex historical event in detail. Such an inquiry, he says, 'can be undertaken only with trepidation: The lens of history grows cloudy with time, and human beings have limited capacities by which to fathom unfathomable evil' (Doris 2002, p. 53-54). But all misgivings aside he proceeds to apply a situationist analysis of some of the main participants in the Holocaust: the SS doctors at Auschwitz extermination camp.

One of the major duties for the SS camp doctors was perform 'selections', that is, they would meet arriving prison trains and separate the new arrivals into two groups: those who were fit for work (generally healthy adolescents and relatively young adults), and those who were not (children, the elderly, expectant mothers or mothers with infant children, the sick and infirm). Those who were judged unfit were sent immediately to the gas chambers, the others were kept alive so they could be used for slave labour. These 'selections' were considered something of a rite of passage for new camp doctors, and they were generally required to perform their first within days of arrival (for details refer to Robert Lifton's monumental 1986 study *The Nazi Doctors*, a source Doris relies on heavily and on which I

will also rely in my response). Some resisted for a while before eventually taking part, yet while there were no explicit penalties for not performing selections in the entire time the camp was in operation only a single doctor is known to have refused to participate. 'Given the institutional pressures at work in Auschwitz, and under the Third Reich more generally, it's tempting to explain such refusals by appeal to moral character...Yet "virtually all" Auschwitz doctors performed selections; did only men of bad character find their way into the camp?' (Doris 2002, p. 54). Doris clearly thinks not: 'A plausible conjecture, just as with Milgram's obedients or the Stanford guards, is that a very substantial percentage of perpetrators in the Holocaust had previously led lives characterized by ordinary levels of compassion' (Ibid).

Doris does suggest that robust dispositional structures may have been present in some of these war criminals, but only in those who 'perpetrated cruelties with more energy than required by even the most morally depraved Nazi job descriptions' in whose case 'the evil had to come from within, not without' (Ibid). But while situationism 'does not deny the existence of monsters', it does 'deny that the explanation of their behavior will be applicable to the generality of cases' (Ibid, p. 55). For the others, the majority of the doctors, Doris appeals to the abundant evidence of internal conflict to argue that the determinative factors were situational rather than personal, in this way drawing an explicit parallel with the Milgram experiment (which, incidentally, was specifically conceived with the events of the Holocaust in mind). In his view the likely explanation for this conflict is 'that the subjects had previously internalized ordinary canons of decency, or to put it another way, they possessed an ordinary complement of compassion.' But during 'an intensive socialization process' with veteran camp doctors, heavy drinking and openly expressed reservations about the camps would eventually result in 'group rationalizations' which were 'a means for the doctors to establish consensual validation for behaviors that were strongly dissonant with precamp values' (Ibid, p. 56). This socialization process would enable the new doctors to complete their first selection, and as they adjusted to life in the camp selections became easier until "it became almost routine" (Auschwitz doctor quoted in Doris 2002, p. 55). This socialization process, the wider institutional context, and the feeling that they were engaged in a war for the very existence of the German race were the determinative factors that caused ordinary men to commit such heinous crimes; or so Doris argues. The case of

the Auschwitz doctors therefore provides both an illustration of how the situationist thesis applies in practice, and evidence for its validity.

The Balkans

In his papers of 1999 and 2009 Harman briefly discusses the application of situationism to the interpretation of the ethnic conflicts that followed the break-up of Yugoslavia in 1992. In his view, the relatively rapid escalation of inter-ethnic violence following a long period where ethnic tensions were, if not absent, at least kept under control cannot adequately be explained by appealing to the character traits of the individual actors. Instead he asks us to consider how in a time of chaos and struggle over limited resources, coalitions might form around along ethnic and religious lines as these make it relatively easy to differentiate friend from foe. This 'us and them' mentality is no doubt solidified by considering historical injustices perpetrated by 'them' against 'us'.⁸ The formation of such a coalition tends, by oppressing the opposing ethnic group, to catalyse the formation of an opposing coalition along similar lines, which then carries out atrocities in response, leading to further escalation (Harman 1999, p. 329).

Harman notes that these conflicts are often attributed to historical 'ethnic hatreds', and as a result we often doubt there is anything we can do to prevent them. How does one even begin to go about resolving ethnic tensions that have gone on for generations, even centuries? On the other hand, 'If we understand the way violence arises from the situation, we may see more opportunities to end the conflict' (Ibid). Not only does situationism provide an explanation for the ethnic violence, it also tells us where to look for a solution; at least, so the argument goes. It is important to note that the situationists aren't merely offering a critique of virtue ethics; they also consider that the acceptance of situationism will lead to more ethical outcomes. But this is the least we should expect, it would be reckless to discard a functional if flawed ethical theory without something constructive to

⁸ Whatever the final verdict on situationism, it can be of no small importance when trying to understand how so many ethnic Serbs were able to carry out such atrocities during the wars following the break-up of the former Yugoslavia to note how brutally Serbs were oppressed by the Croatian Ustaše regime during the Second World War.

put in its place, at the very least one must be able to argue that the alternative constitutes an improvement.

2.3 Conclusions

The situationists speak with one voice regarding the implications of the evidence listed above; it is utterly damning for virtue ethics, and indeed any other theories, ethical or otherwise, that presuppose the existence of robust character traits. Harman, for example, concludes that ‘there is no empirical basis for the existence of character traits’ (Harman 1999, p. 316), and suggests that we ‘simply replace thought and talk about virtuous character traits with thought and talk about virtuous acts...because ordinary thinking about character traits is such a mess’ (Harman 2009, p. 241). Indeed, he goes so far as to say that ‘ordinary thinking about character traits has deplorable results, leading to massive misunderstanding of other people, promoting unnecessary hostility between individuals and groups, distorting discussions of law and public policy, and preventing the implementation of situational changes that could have useful results’ (1999, p. 330). The implication is that those who maintain a belief in such traits are not only mistaken, they are actually *bad*, in that they are obstructing ethical progress.

Likewise, Doris holds that the evidence refutes any personality theory that presupposes the existence of robust character traits (or, as he calls them, global traits), and instead contends that personality ‘should be conceived of as *fragmented*: an evaluatively disintegrated association of local traits’ (Doris 2002, p. 64), where local traits are narrow, situationally specific regularities of behaviour (for example: courage in battle, courage-with-sharks, courage-in-public-speaking...), as opposed to the broad, robust traits required for virtue ethics (see also: Harman 1999, p.326).

If the situationists are correct, virtue ethics is a dead-end theory, at least for all committed ethical naturalists. In the next chapter we will consider some objections to the situationist interpretations of the evidence presented above; then, in the final chapter we will consider the more fundamental question of the nature of virtue and whether the situationists (and indeed, many of their opponents) have misconceived it.

Chapter 3 – Evaluating the Evidence

In the previous chapter I presented the evidence for the situationist critique, in this chapter I will critically examine this evidence. I will introduce and assess some of the most common objections to that evidence; in particular the objections to the use the situationists make of experiments they admit were never designed to investigate the existence of robust character traits. Some of these objections are rejected; some will meet with partial acceptance. None of them, whether alone or in combination are sufficient to completely undermine the situationist position; the evidence for widespread situational variance in behaviour is simply too compelling. However, my goal is not to attack situationism but to defend virtue ethics; and so I intend to engage with the evidence, to interpret it, to draw out the implications and ultimately to assess whether the evidence supports the situationist critique of virtue ethics. For the situationists are taking a very strong line in concluding that the moral psychology of virtue ethics is empirically inadequate, which in effect is the claim that it is inconsistent with, or unsupported by, the available evidence. I will be satisfied if I can show that the evidence is compatible with virtue ethics; whether it is also compatible with situationism is a lesser concern. Indeed, I do not deny the widespread existence of situational effects and their ethical significance, I simply intend to explore the possibility that these effects are not a revelation and that virtue ethics does provide guidance on how individuals can recognize the effects of situations on themselves and to overcome them, at least within reasonable limits; this is the task I will take up in my final chapter.

3.1. Masking

One response to the situationists appeals to *masking* dispositions, a familiar concept drawn from the philosophical literature on dispositions.⁹ Masking occurs when two or more opposing dispositions exist within the same individual; for example, when a schoolboy confronted by a bully is torn between the disposition to stand his ground and the disposition to flee. As it is not possible to *both* stand one's ground *and* flee, only one of those

⁹ As just one example, see Jan Hauska: *Dispositions Unmasked - Theoria* 75 (4):304-335, 2009

dispositions can be effectively expressed in action; the other is therefore said to be 'masked', i.e. it is present but not expressed.

How could this line of reasoning be applied to the evidence discussed in chapter Two? Well, in the case of the Milgram experiments, one might argue the results support the existence of two opposing dispositions: the dispositions of compassion and obedience. We've all been trained since infancy to obey authority figures, so it's at least plausible that there exists something like a disposition to obey those recognized as being in authority. And there can be little doubt that the experimenter is the one in charge of an experiment. In the normal course of events it is of the utmost importance that an experimental subject obeys the experimenter's commands; failure to do so could invalidate the results. So perhaps Milgram's obedient subjects weren't lacking compassion, perhaps we have a case where 'both the disposition of compassion and the disposition of obedience were elicited...but since the subjects' behavior could not be both obedient and compliant, their behavior would have appeared inconsistent with *some* disposition' (Russell 2009, p.283). Rather than concluding that the obedient subjects were lacking compassion, couldn't we conclude that they were *both* obedient *and* compassionate? If one pursued this argument to its logical conclusion, one could even suggest that some obedient subjects could have a stronger disposition to compassion than some disobedient subjects; it's just that their compassion was masked by an even stronger disposition to obey.

One advantage of this argument is that it offers an explanation for the conflicted nature of the obedience Milgram observed. It would be difficult to explain the intense distress the obedient subjects typically displayed unless one assumes they felt *some* compassion for the learner. But whatever merit there may be to this argument, it fails to undermine the situationist critique. The problem, as Russell rightly notes, is that 'the question was never whether the subjects faced conflicting pressures – the experiment was designed to *ensure* that they did – but whether their supposed behavioral dispositions would be elicited by forces that were intuitively *stronger* than the opposing forces' (Ibid, p. 286). The situationist case is not that the experiments discussed should undermine our belief in compassionate dispositions; rather the argument is that those experiments should undermine our belief that compassion exists as a *robust character trait*, and the masking argument does not touch on the issue of robustness. While masking may assist in our

understanding the Milgram experiments, would-be opponents of situationism will need to look elsewhere.

3.2. Remote Scenarios

Similar considerations tell against another objection to situationism, namely that the Milgram experiments are ‘remote’ or ‘improbable’ scenarios that tell us little or nothing about the existence of character traits. But despite the fact the objection ultimately fails it still raises an important issue that will need to be addressed by any virtue theory that wishes to successfully *overcome* the situationist critique, rather than merely evade it.¹⁰

Of course, there is a sense in which any psychological experiment is a ‘remote scenario’, in that participation in any such experiment is a rare occurrence in most people’s lives; but this is not what we mean by remote in this instance. An experimental scenario can be considered remote if it is intended to replicate an experience the subject would be unlikely to encounter in real life. In this sense the Good Samaritan experiments aren’t remote scenarios, as it is very likely the average person will encounter distressed individuals in need of medical assistance at some point during their life. The Milgram experiments present a different story however; it is *highly* unlikely any one of us will ever be asked to electrocute another person in any circumstance, so it is safe to say the scenario presented by the Milgram experiments is a remote one.

With this in mind we can now state the remote scenario objection: no inference should be drawn about an individual’s character based on their behaviour in a remote scenario (such as the Milgram experiments). As Oakley notes, ‘we sometimes...describe behavior observed in remote and improbable scenarios as ‘not in character’ for the agent in question, and we may even subsequently exclude such behavior in extraordinary circumstances from the evidence we take as instructive for informing warranted character-trait attributions to the agent in question’ (Oakley in Birondo & Braun 2017, p. 32). Rather

¹⁰ It should be clear by now that it is always possible to evade an argument in philosophy, if only by cleverly redefining the terms of the debate. If an argument has any force then it has some value even if it is ultimately mistaken. While I consider the situationist critique to be misguided, a judicious analysis of the reasons for its failure allows us to replace the flawed modern virtue theories that invited it with something better.

than providing evidence that compassion is not a robust trait, it could simply be the case that Milgram contrived a situation so unlike any in his subjects' past experience that they could find no precedent on which to act. Unable to decide for themselves how to proceed, it is only natural that many subjects would resort to following the instructions of an expert who surely 'knows best'.

There is much to be said for this objection, and I will return to it when I consider the question of interpretation below. But in itself the remote scenario objection can hardly be fatal to the situationist cause: first, because of all the experiments the situationists appeal to only the Milgram experiments can plausibly be considered to present the subjects with a remote scenario; and secondly, because it offers no explanation for the observed situational variance. Recall that the situationist critique is a critique of the psychological underpinnings of virtue ethics. If no psychological explanation consonant with virtue ethics can be provided for the observed situational variance the situationist critique retains its force. It cannot be denied that Milgram's obedient subjects failed an ethical test; any plausible virtue ethics must provide an explanation for those failures and in doing so it will inevitably have to account for the situational variance.

3.3. Inconclusive results

Perhaps this is a little hasty though, perhaps there is another way out for the dedicated virtue ethicist. The situationist critique relies on the results of a set of social psychological experiments that demonstrated significant situational variance in behaviour, but the experiments employed represent a small fraction of the total number of experiments that may be relevant. This may be unavoidable, a truly comprehensive review of all the experiments that *might* be relevant would be an enormous undertaking, possibly doomed to failure at the outset (though I have no wish to deter anyone from making the attempt). But we can reasonably expect the situationists to present a representative or balanced sample of the relevant experiments; any suggestion that the situationists have only considered those experiments that strongly support their critique while avoiding (consciously or otherwise) discussion of experiments which provide little or no evidence of significant situational variance would seriously undermine the situationist project, while

potentially opening them to a charge they're often so willing to lay on their opponents, that of *confirmation bias*.¹¹

According to Alzola this is precisely what has happened. He argues that the situationist critique is based on a skewed sample of the available experimental evidence and 'subsequent variants of some of the experiments show less convincing results than those presented by Situationism' (Alzola 2008, p. 348). To support this claim he refers to a 1974 repeat of Isen and Levin's 'coin slot' experiment by Blevins and Murphy which 'found quite different results: 43% of the subjects helped in spite of not finding any coin and 40% of the subjects who did find a coin in the return slot did not help the confederate' (Ibid); and 'a second version of the Good Samaritans experiments (Batson et al., 1979) in which the importance of what the subject was hurrying for did make a significant difference to the results. Batson and collaborators told half of the subjects that their performance on the task awaiting them in the second building was 'of vital importance' to the experimenter and the other half were told that theirs was 'not essential'. (While in the original experiment only 10% of the subjects in the rushed condition were helpers, 70% stopped to help when they were in a hurry for something of little importance' (Ibid)).

While these results do present a notable contrast to the original experiments and provide a useful reminder of the hazards of relying too heavily on the results of individual experiments, I believe it would be premature to suggest this invalidates the situationist critique. I've already indicated in Chapter Two that the Isen and Levin experiments contribute little to the situationist critique on their own, so it would be hypocritical of me to claim the Blevins and Murphy result referenced above as evidence against situationism. As for the results of Batson et al, contra Alzola I believe one could interpret the results as further evidence *for* situationism, rather than against it. Under the influence of a mere verbal cue, on being told that a task 'of vital importance' awaits, more than half of the experimental subjects apparently ignored a real living person in distress; what could possibly be more vital than that? It is not the mere fact of situational variance that grounds the situationist critique, it is the disproportion between the situational cues and their effect on individual behaviour, and if anything this experiment seems to demonstrate just such a

¹¹ Simply put, 'confirmation bias' is the tendency to seek and absorb evidence that confirms a belief one already holds, and to filter out or avoid contrary evidence.

disproportionate response. Suffice to say, I do not consider that there is good reason to conclude that the results of the experiments Alzola references are at odds with the results of the experiments surveyed in Chapter Two. Short of engaging in a comprehensive survey of all those experiments that may be relevant I believe we should take the results of the experiments we've considered at face value; which does not mean we must accept the situationist interpretation of them.

3.4. Experiments not fit for purpose

A more sophisticated response to the experiment is to question their relevance to the question at hand. Granted that the results are striking, even alarming at times, do they really provide evidence against the existence of robust character traits? The Milgram experiments, for example, were carried out 'to test whether the excuse that one was only obeying orders is one which had any force, i.e. would most people follow the dictates of morality or would they follow orders to the contrary'. However 'there is nothing...that indicates that the Milgram experiments were either designed to prove or did prove that there are no such things as character traits' (Athanassoulis 2000, p. 216). Much the same could be said for the other experiments the situationists appeal to; not a single one of them was designed for the express purpose of determining whether people possess robust character traits, they are all being employed by the situationists for a purpose for which they were neither designed nor intended. Experimental design requires great care, the aim being to isolate only the variable or variables one wishes to study while keeping everything else fixed. So the fact that the experiments we have considered were not designed to confirm or deny the existence of robust character traits does not bode well for the situationist. We must not be hasty however; it *could* be that the experiments are fit for purpose despite the fact that they were not designed to be. If we want to convince a situationist that the experiments they rely on fail to provide evidence either for or against the existence of robust character traits we should be able to specify how they fail, and by implication what design features experiments might need to have in order to detect such traits if they do exist. I will not attempt to provide an exhaustive list of the aspects of the experiments that detract from their applicability to the question of the existence of robust

character traits, but instead focus on the three aspects I consider to be most critical: 1) The limitations of one-shot studies, 2) Ecological validity or lack thereof; and, 3) Failure to consider ‘construal’, or how the subjects’ perception of the experimental situation impacted on their behaviour.

3.4.1. Limitations of one-shot studies

The experiments we considered in Chapter Two are supposed to provide evidence against the widespread existence of ‘robust’ or ‘global’ traits by demonstrating (in Doris’ terminology) that they are neither *consistent* nor *stable* (though Doris does allow stability of narrow, situationally specific, ‘fragmented’ traits). *Consistency* is the requirement that traits be ‘reliably manifested in trait-relevant behavior across a diversity of trait-relevant eliciting conditions’, while *stability* requires that traits be ‘reliably manifested in trait-relevant behaviors over iterated trials of similar trait-relevant eliciting conditions’ (Doris 2002, p. 22-23). When creating an experiment to test for consistency, then, it seems we’d need to arrange for the *same* individuals to experience multiple situations that are designed to activate a particular trait ‘across a diversity of trait-eliciting conditions.’ The trait in a particular subject would be consistent to the extent that they exhibit the same trait in response to the various eliciting conditions for that trait. Likewise, when testing for stability, one would think we’d need to arrange for the *same* individuals to experience ‘similar trait-relevant eliciting conditions’ across a number of occasions, presumably extended over a considerable period of time. In this case the trait would be stable to the extent that the same trait relevant behaviour would be exhibited in response to the eliciting conditions.

When we consider the situationist experiments with these considerations in mind it is strikingly obvious that they ‘did not track the behavior of particular individuals across situations on multiple occasions. Experimenters typically observed any given individual only on one occasion in a particular situation’ (Alzola 2008, p. 349). It follows that these experiments can tell us nothing about the stability or consistency of any character traits an individual participant may possess; after all, ‘what can be concluded about the consistency of the subjects’ behavior on the basis of a single observation?’ (Ibid). It would appear that the situationist experiments are incapable of supporting the situationist cause; their very

design seems to preclude any valid inferences regarding the stability or consistency of the character traits of their participants. If this is so (and it certainly seems so) there would be nothing more to say; all the experiments the situationists rely on fail to support the situationist critique, hence the critique would fail *on its own terms*.

In reality the implications aren't so dire for situationism (otherwise I would have just started with this objection and brought the matter to an early end), and this challenge would hardly come as a surprise to its adherents. Doris was well aware that the experiments 'can tell us nothing directly about the consistency of the subjects' (Doris 2002, p. 38), such direct evidence would require 'systematic observation of that individual's behavioural patterns. To gather this sort of evidence, one requires longitudinal studies that observe individuals over a period of many years in numerous and diverse situations' (Ibid). What the experiments do provide, in his view, is 'a powerful indirect argument against the existence of widespread consistency in helping behavior' (Ibid). While he does not go on to set out this 'powerful indirect argument' I believe we can do the work for him.

For the moment I will just consider the Milgram experiments, what follows will be more-or-less generalizable to the other experiments. Moreover, to keep things simple, let us focus on just two of the variants: the original experiment where the experimenter sat in the room with the 'teacher' when ordering them to administer shocks, and the variant in which the experimenter was absent and gave orders by telephone. According to the figures quoted in Doris (pages 42 & 46), in a typical experiment of the 'experimenter in the room' type the incidence of full obedience was 65%, while in a typical experiment of the 'experimenter by phone' type the incidence of full obedience was a mere 21%. Let us grant, as seems plausible, that the difference between receiving orders by phone and receiving them in person is too minor to warrant varying one's behaviour between the two cases, thereby implying that for an individual to be fully obedient in the one case but not in the other would be compelling evidence of inconsistency on their part. Let us also grant that the two groups of experimental subjects were more or less 'representative', which is just to say that we have no reason to think that one group was more likely to be obedient overall than the other. Then as Doris suggested we have some powerful evidence for 'widespread inconsistency', in that the incidence rate of full obedience in the 'experimenter in the room' experiment is triple that of the 'experimenter by phone' experiment. It is all very well to

suggest that the situationist experiments were not designed to investigate the consistency of individuals, but failing to address this inconsistency between the group results would be a failure to engage with the situationists on their own terms.

I would like to suggest that the true implications of the inconsistency cited above lies somewhere between the extreme views of the situationists and their more virulent opponents. There is widespread inconsistency, this should simply be accepted as an established fact; but we should be careful not to place too much weight on such a vague term as 'widespread'. In absolute terms, granting the assumptions above that the two groups of subjects are more-or-less equivalent, we can conclude that 44% of the subjects in the 'experimenter by phone' group wouldn't have been fully obedient in that experiment, yet would have been fully obedient if they had been in the 'experimenter in the room' experiment. Even granting the assumptions that have to be made to reach this figure (assumptions that, if granted, buttress the situationist cause if anything) can the inconsistency of 44% of subjects really be considered evidence against the widespread existence of robust character traits? Even if we appeal to other variations of the Milgram experiment to expand the percentage inconsistency, the question will still remain: how widespread does the inconsistency need to be to throw doubt of the very possibility that human beings can, in general, *develop* robust character traits, including the virtues? For this is all that the virtue ethicist requires from the naturalistic point of view, that human nature be such that the virtues *can* be developed (at least 'for all who are not maimed as regards their potentiality of virtue' – NE 1109b:10), and not that most people already have them. Indeed, Aristotle seems to explicitly deny that *anyone* just has them (c.f. NE 1109b: 5 – 1110a:5); virtue where it is present is always acquired. The experiments simply do not allow us to determine whether the widespread absence of robust traits is unavoidable, the situationists seem on the whole to be satisfied with demonstrating the actual widespread absence of robust traits without questioning whether there might be contingent reasons for this, without wondering whether and how they might be produced. It is worth remembering that situationists don't (typically) claim that no one possesses robust traits, they just claim that where individuals do these are exceptional cases that are not generalizable. Consider, for example, what Doris says regarding those who risked their lives to save Jews during the Holocaust: 'even if rescuers exhibit a consistency of behavior suggesting highly robust

dispositions to compassion, this is something I can grant, because situationism does not preclude the existence of a few saints, just as it does not preclude the existence of a few monsters. But these “tails of the bell curve,” the situationist claims, are the exceptions that prove the rule: “Altruistic personalities” with consistent behavioural implications, if they exist, are remarkable precisely because they are rare’ (Doris 2002, p. 60). If this is to be more than a token admission of the natural limits of human inquiry, of the impossibility of demonstrating the *universal* absence of robust traits, then it should be accompanied by an inquiry into the reasons for these acknowledged exceptions. How did some people become ‘saints’, and are others able to follow in their footsteps? This inquiry is sadly lacking in the situationist canon.

3.4.2. Ecological Validity

Another challenge to the claim that the situationist experiments show that humans generally do not possess robust character traits appeals to the requirement that social psychological experiments be ‘ecologically valid’ if any valid inferences are to be drawn from their results. In Alzola’s useful summation, ‘Experiments are carried out to make inferences to other non-experimental situations. The experimenter observes events in a standard situation while holding constant everything other than the particular independent variable under investigation. For this technique to allow valid inferences, it is crucial that the experimental situation adequately reflects the process under investigation’ (Alzola 2008, p.348). If the experimental situation does not adequately model the ‘non-experimental’ situation it is intended to model, the experiment lacks ecological validity, and since we are presumably only concerned with non-experimental implications the experiment may be of no use at all. Alzola goes so far as to suggest that ‘situational factors are less powerful in natural contexts than they are in experimental contexts’ and that ‘Experimental conditions may weaken dispositional traits’¹² (Ibid). I am not prepared to claim that the relationship between the experimental context and the power of situational conditions is so simple, and in any case this is not necessary; if an experiment can be shown to lack ecological validity then its results are likely to be distorted and little weight should be attached to it.

¹² Note these are *not* equivalent, though they may lead to equivalent results.

The question immediately arises: were the situationist experiments ecologically valid? For the most part I'm inclined to say they were; the bulk of the experiments model relatively simple situations and, crucially, the experimental subjects were unaware that they were participating in an experiment. As they didn't realise they were in an experiment we may reasonably assume the Good Samaritan subjects behaved as they really would if they encountered a person in distress while in a hurry, that Isen and Levin's subjects really would be more likely to help pick up papers after finding a dime, that individuals *generally* would be less likely to offer assistance to another person when accompanied by inactive peers. Moreover, the situations aren't far-fetched, we could all encounter similar situations in our day-to-day lives, most of us probably already have. Without any evidence that a significant number of subjects behaved otherwise than they would have done in similar non-experimental situations (and I am not aware of any such evidence) I believe we should concede that these experiments are ecologically valid.

It should have been noted that there is one significant situationist experiment missing from my list in the above paragraph; the Milgram experiments. Unlike the others I believe there is good reason to believe that the Milgram experiments fail to be ecologically valid, for two key reasons: 1) the subjects knew they were in an experiment; and 2) the extent of the deception involved necessarily distorted the subjects' perception of their situation (and indeed this distortion is absolutely fundamental to the design of the experiment). In combination these two factors ensure that the Milgram experiments fail to be an adequate model of a non-experimental scenario.

With regards to the first reason, this has been considered as an objection in and of itself, what we might call the 'objection from artificiality'. The idea is that 'the very artificiality of the experiment context itself – the fact that subjects knew they are 'in a test', and in a contrived and unnatural environment – mitigates Milgram's findings' (Russell 2009, p. 281). Russell is sceptical of this argument, and I'm inclined to agree that appealing to artificiality will not allow us to 'explain away' the situational variance observed in the Milgram experiments, though I'm less inclined to accept his intuition that 'such a factor as artificiality does not exert a greater compassion-opposing force than the hysterical screams of an innocent person exert compassion-eliciting force' (Ibid). The problem is we do not *know* what effect the knowledge they were in an experiment would have on the subjects,

and in all likelihood the effect would vary from person to person. It is not enough to know what the subjects did; we also need to know why they did it in order to adjudge the evidence as being in favour of either situationism or virtue ethics (or both, or neither). If they simply reacted to the verbal cues we might think the results favour situationism, if they continued to administer 'shocks' out of fear of the experimenter who was sitting just behind them asking them in a "firm, but not impolite" tone to torture the affable stranger they had just met we might conclude that they were lacking courage. The problem is the artificiality of the situation, or rather the subjects' awareness of it (unlike the other experiments, which were artificial but without the subjects being aware), casts a little seed of doubt into any interpretation of the results. Granted it is not sufficient in itself to invalidate the situationist interpretation, but in conjunction with other considerations one starts to realise that interpreting the Milgram experiments is far from being a simple matter. In particular the experiment fails to isolate the situational influences that were presumably (for all intents and purposes) identical for each subject from the influence of each subject's individual psychological response to the situation they perceived themselves to be in.

The second reason is linked to the first, and together they throw even more doubt on the possibility of any interpretation of the results that 'factors out' the subjective experience of the individual subjects. The Milgram experiments stand out from the others in the level of deception that is employed, the experiment is founded on the complicity of the experimenter and the confederate in deceiving the experimental subjects. Some amount of deception was necessary for Milgram's purposes; obviously the subjects couldn't be allowed to know that they weren't really administering electric shocks, this would have invalidated the entire experiment. The problem is we have a situation in which the participants knew they were in an experiment, thought they knew what the experiment would involve, gradually became aware that their presuppositions were false and were left to their own devices to try to make sense of the situation they found themselves in. The subjects went from a very stable situation, confident in the knowledge of what was to come, a confidence that was deliberately fostered by the polite experimenter and the affable confederate, only to eventually find themselves in a situation they surely could not have foreseen, for all intents and purposes being instructed to torture an innocent stranger. What situation did the subjects think they were in after all of this? I doubt many of them were sure. In any

case, the idea that they were being offered a simple choice between torturing and not torturing the learner seems a little naïve, if not disingenuous. In order to properly interpret why individual subjects acted the way they did we need to take into account the uncertainty they must have experienced, an uncertainty that must have thrown into doubt many things they thought they knew (such as the belief, unthought perhaps but no doubt present in most of the subjects' world-view, that such things don't happen in a 'free country'), an uncertainty that would have taken great presence of mind to overcome and which was fostered if not created by the deceit that was part-and-parcel of the experiment.

Given these two factors - the artificiality of the situation resulting from the fact that the Milgram subjects knew they taking part in an experiment, and the deception that was practiced on them – let us turn again to the question of ecological validity. What situation is the Milgram experiment designed to model? If it is simply a matter of determining whether an 'ordinary person' would follow criminal orders to torture or kill another person then we must conclude it is ill-designed for that purpose, the extent of the deception involved is incompatible with the situations in which an individual might be *ordered* to commit crime. An order is an instruction to act given by a person with authority to demand and if necessary *compel* obedience. There is no place for deception of the one ordered and no need in any case; if an order isn't followed you don't trick the person into following your order, you take disciplinary action and in the meantime find someone who will follow orders. On the other hand, the Milgram experiment could be designed to model a situation in which an individual volunteers for an apparently simple task, only to discover part way through that they are faced with the choice between carrying out a criminal act or standing up to the very person who is calmly asking them to carry it out (to say nothing of the fact that the 'victim' played a part in the overall deception); but what 'non-experimental' situations could this possibly be intended to model? Either way you look at it, the Milgram experiments fail to be ecologically valid: either they fail to model the situation they are intended to, or they model a situation no one actually faces. This is not to say the results are insignificant; the situational variances still need to be explained. But it is to say that we should be wary about drawing any wide-ranging conclusions from them, in particular whether the observed variance is primarily the result of external situational factors or internal subjective ones. In order to get a handle on this problem we need to consider a factor the situationists typically either avoid or dismiss:

the power of construal, that is, how the individual subject's perception and interpretation of their situation influences their behaviour.

3.4.3. Construal

In this section I will focus my attentions on the Milgram experiments, as they are put forward by the situationists as the best evidence for their critique of virtue ethics and the implications will be generalizable to other experiments. In order to interpret individual human *action* (as opposed to impulsive reaction) it is necessary to consider not only the situation they were in but, more critically, the situation they *believed* they were in. The obedient Milgram subjects had time to think over their situation, to make decisions, to consider possible courses of action and possible consequences thereof, and still continued to 'shock' the learner. The question we should ask is: why? The answer will naturally vary from case to case, but we can simplify matters by focusing on a specific case. Remember that the evidence cited in favour of the situationist critique is the significant variance in behaviour under the influence of apparently insignificant situational factors. One such situational factor, already cited above, was the fact that the experimenter sat in the room with the subject when instructing them to continue with the test, rather than delivering instructions by phone. As a result of this a significant proportion of the subjects were fully obedient when the experimenter was in the room who would not have been if the instructions had been delivered by phone (I am willing to accept this, any evidence to the contrary would only be unfavourable to the situationist). Now we simply need to ask: why did the fact of the experimenter being in the room make such a difference to so many subjects? It is striking that the situationists seem to feel no need to ask this question, they are satisfied to merely observe that 'seemingly insubstantial situational factors have substantial effects on what people do' (Doris 2002, p. 28), and to draw situationist conclusions from them. But if we're going to decide what counts as a 'seemingly insubstantial situational factor' we'd better not be relying on 'intuition', which spectacularly failed to predict the results of the Milgram experiments in the first place. On the other hand, it is hard to know what other basis situationists could have for making claim that the experimenter being in the room was an insubstantial situational factor (the *evidence* seems

to show that a significant number of people found it quite substantial), so perhaps we should make an attempt to understand the situation from the subject's point of view.

The problem of construal arises because situations are open to interpretation, and it is the situation as the individual sees it that motivates their behaviour, not the situation itself (though obviously, the two will be related in all but the most psychotic of individuals). Alfano characterises issue of construal as follows: 'Ambiguous environmental cues require interpretation...trait-eliciting conditions should be divided into objective and subjective components. Someone may be in a situation where helping is the appropriate response but not see it that way; conversely, someone may be in a situation where helping would be inappropriate but believe she ought to help' (Alfano 2011, p. 127). But he then goes on to essentially deny that there is any issue here for the situationist, since 'Introducing the intervening variable of construal between objective stimulus and behavior just gives a fuller account of how people can fail to react virtuously; it does not save virtues from empirical critique. Thus, the ultimate objects of correlation remain objective conditions and behavior, but subjective construals form part of the theory connecting the two' (Ibid, p. 128).

It is worth noting at this point, that the Milgram experiments are valid only to the extent that the subjects misconstrued their situation; if they'd known what was really going on the experiments would have no value whatsoever. So with respect to each of the Milgram subjects there are three situations that could be considered: 1) the situation as it was, 2) the situation as they were supposed to see it, and 3) the situation as they actually saw it. The problem for Alfano is that correlating between the 'objective conditions and behaviour' could lead to the conclusion that the subjects *did nothing wrong*, at least on a consequentialist account, because the objective conditions were that no shock was being administered (on this account it's only the experimenters who were behaving unethically, essentially tormenting their unwitting subject). And if we don't apply consequentialism we must take into account the subject's intentions, which means we also have to consider why they did what they did, and we can only understand this 'why' if we consider the situation from the subject's point of view. The 'objective conditions' are of no use to the situationist; they actually *require* that the subject misconstrue their situation, only in a particular way. Returning to the example above, the situationist requires that the difference between having the experimenter in the room rather than giving instructions by telephone be

relatively innocuous. And indeed it appears to be, when you see it in writing; but what about when you are actually sitting there, with the experimenter sitting behind you, calmly instructing you to torture another person? Is it really so hard to believe that subjects might find this rather intimidating? Threatening even? But once you consider this the experiment appears in a whole new light, and the difficulties for the subject wanting to extricate themselves from this strange and uncertain situation multiply with each step. The subject cannot know that they merely have to ask four times to leave and they'll be allowed to do so, they have no way of knowing what the sadist (for how else can you describe the experimenter from the subject's point of view?) behind them will do if they insist on stopping the experiment. Just ask yourself: would you be more comfortable being in the same room as a sadistic killer, or speaking to him over the phone? Would you think just as clearly in either case?

I won't pursue this story any further, it is inevitably speculative and besides, my point is already made. Once we recognize the importance of construal we are in a position to account for an aspect of the experiment that the situationists note, without providing any explanation for: the conflicted nature of the obedience. As Doris notes, 'obedients saw themselves as at least partly responsible actors in proceedings that they believed to be – as their manifest anxiety attests – morally objectionable' (2002, p. 50). If the obedients recognized they were doing something wrong, why didn't they stop? It is interesting that Doris never felt the need to answer this question, perhaps he knew that the answer, whatever it was, would be of no assistance to him. Aristotle, on the other hand, in a highly suggestive passage seems to hit the nail on the head: 'moral virtue is concerned with pleasures and pains; it is on account of the pleasure that we do bad things, and on account of pain that we abstain from noble ones. Hence we ought to have been brought up in a particular way from our very youth, as Plato says, so as both to delight in and to be pained by the things that we ought; this is the right education' (*NE* 1104b:5-13). The simplest explanation for the obedient subjects continuing to commit an act *they* believed was wrong is that they were afraid of the consequences if they refused; after all, fear in its broadest sense is simply a natural reaction to anticipated pain. Once it is recognized that the situationist is not, cannot, be requiring the subjects to respond to the 'objective conditions' this immediately throws open a wide range of subjective conditions that need to be taken

into account when interpreting the behaviour of the obedient subjects. To be able to sift the essential elements, to think clearly and to make the best choice in an uncertain situation, regardless of the possible danger to oneself, requires a certain amount of mental fortitude and *courage*.

Consideration of the subjective is also critical to distinguishing between virtue and continence, i.e. between acting 'as the virtuous person does' and merely acting 'in accordance with virtue'. Continence is much more common than virtue; it is the norm, it is compliance with the requirements of social convention and the law, without any inner commitment to them. Any act that hits the mark of virtue exists on a continuum between pure continence and pure virtue (assuming it is a deliberate act of course). The merely continent person is liable to miss the mark of virtue when he encounters obstacles or temptations, so one could conceivably determine where a person sits on this continuum by placing them in situations where virtue requires the same act, yet which present more or less in the way of obstacles; this is precisely the format of the various Milgram experiments. If we want to explain the variance in behaviour between the different versions of the experiment virtue ethics provides a relatively simple explanation: many of the disobedient subjects in the less onerous experiments were merely continent, when the demands of virtue became harder they failed to meet the test.

The importance of construal is immediately apparent in other situationist experiments as well. Interpreting the implications of the Good Samaritan experiment, for example, becomes much more complicated when one reads that 'Darley and Batson reported that subsequent interviews with subjects revealed not that hurried subjects simply noticed but ignored the victim's distress, but that many of them 'did not perceive the scene...as an occasion for ethical decision' (Russell 2009, p.274). What is the appropriate censure for a subject who simply didn't perceive the situation 'correctly' due to the fact they were in a hurry? Granted they may not qualify as virtuous (perhaps virtue requires some kind of 'situational awareness'), but surely they cannot be placed in the same category as an individual who saw a person in distress yet refused to help? The experiments provide bountiful evidence that 'subjects struggle to know how to 'encode' such highly 'ambiguous' situations' (ibid, p. 277), and observers likewise. *Phronesis*, or practical wisdom, must in large part consist in the ability to correctly 'encode' ambiguous situations, and full virtue

presumably entails the ability to choose the best act, taking into account what is known and accepting the unknown. It is rare to know everything there is to know about a situation, life just isn't like that. To recognize that, accept the ambiguity and to act as best one can with confidence and courage in spite of uncertainty; that is surely admirable from any perspective and worthy to be called virtue.

3.5. Assessment of Historical Case Studies

The historical studies discussed in chapter Two were introduced by Doris and Harman as something of a side-note, an illustration of how situationism can be applied to explain racially motivated violence and even genocide, rather than as evidence in their own right. But clearly, to the extent that situationism successfully explains ethnic cleansing in the Balkans and the actions of the Auschwitz doctors, these case studies will support the situationist position. It is therefore the applicability of situationism to these historical events rather than the events themselves that would constitute evidence in favour of situationism; it is time to put this claim to the test. We must be careful however, as any study of a complex historical event provides a wealth of details that may support multiple interpretations, and therefore there is plenty of scope for confirmation bias. It is worth recalling Doris here: 'all parties should agree that behavioral outcomes are inevitably a function of a complex interaction between organism and environment' (Doris 2002, p. 25); the question is the relative weight to be attached to situational versus characterological factors. The situationists argue their historical examples demonstrate the preponderant influence of situational factors, I will merely argue that any explanation of these events must appeal to situational and characterological factors and that we have no grounds for saying either one predominates. The reason this will support virtue ethics over the situationist critique is that our character is something we have some control over and responsibility for, while we often have no control over or responsibility for the situations we find ourselves in. The cultivation of virtue is in part a preparation for future adverse situations that may be outside of our control. Virtue allows us to 'make the best of a bad situation' while situationism offers precisely no assistance in this regard. I will say more about this in chapter Four.

In his discussion of the Auschwitz doctors Doris makes heavy use of Robert Lifton's *The Nazi Doctors: Medical Killing and the Psychology of Genocide*, a monumental study utilising interviews Lifton personally conducted with former Nazi doctors (including Auschwitz doctors) and former death camp prisoners together with primary source materials and eye-witness accounts of the actions of doctors he was unable to witness (whether because they were unwilling or already deceased). Doris uses Lifton's work as a source of first-hand accounts of the actions of Auschwitz doctors which he considers support his argument for the primacy of situational factors in explaining their genocidal actions. I saw no evidence he had taken the time to read Lifton's own psychological analysis of the doctors. This is rather surprising as Lifton, being a psychiatrist, should be considered something of an expert on human psychology, certainly compared to an academic philosopher. If his analysis supported Doris' case one would expect him to make use of it, if it does not this should have given Doris pause and at the very least (in the interest of intellectual integrity) he should have registered this disagreement and attempted a response.

The key concept for understanding Lifton's analysis of the psychology of genocide is the concept of 'doubling'. Doubling, as he describes it, 'was the psychological vehicle for the Nazi doctor's Faustian bargain with the diabolical environment in exchange for his contribution to the killing; he was offered various psychological material benefits on behalf of privileged adaptation' (Lifton 2017, p.418). Doubling is a kind of splitting of the self, not in the pathological way typical of schizoid conditions, but as a defence mechanism, a means for the psyche to adapt to unpleasant circumstances and in particular to enable an individual to commit acts that appear to violate the ethical standards they identify with. Lifton makes it clear that the potential for doubling lies in each of us, one can infer that it is adaptive to the extent that it may allow individuals to survive circumstances that would otherwise break their spirit, but it is also 'the psychological means by which one invokes [the] evil potential of the self. The evil is neither inherent in the self nor foreign to it...in the process of doubling, in fact, lies an overall key to human evil' (Ibid, p. 423-424).

There can be no denying that some of Lifton's analysis seems to support situationist conclusions, for example when he discusses the nature of 'Auschwitz as an *institution* – as an atrocity-producing situation...An atrocity-producing situation is one so structured

externally (in this case, institutionally) that the average person entering it (in this case, as part of the German authority) will commit or become associated with atrocities' (Ibid, p. 425). Yet on closer inspection the implications are less clear. It should not be forgotten that the 'average person' entering the institution of Auschwitz as a part of the German authority was a member of the SS, a group that typically took pride in being without pity and had sworn "obedience unto death" to their Fuhrer and their superiors (Ibid, p. 435). The doctors were not normal people who suddenly found themselves in a death camp and committed atrocities due to situational factors, rather Auschwitz was the end of a long and slippery slope during which they had already compromised themselves many times over. 'Much feeling had been blunted by his earlier involvement with Nazi medicine, including its elimination of Jews and use of terror, as well as by his participation in forced sterilization, his knowledge of or relationship to direct medical killing ("euthanasia"), and the information he knew at some level of consciousness about concentration camps and medical experiments held there if not about death camps such as Auschwitz' (Ibid, p. 442-443). Of course, one could attempt to argue that the earlier decisions that made the doctors complicit in Nazi crimes were the result of situational factors themselves, but this is one thing situationists cannot appeal to. To adduce an action or event as evidence for the primacy of situational influences it must be the aspects of the situation the actor *is then in* that influence the behaviour to be explained. Any suggestion that past behaviour or patterns of behaviour may be the primary factor is evidence *against* situationism in that it appeals to the importance of personal factors over situational ones. Each of the Auschwitz doctors followed individual paths that led them to Auschwitz, each compromised again and again with a criminal organization, whether out of true belief in the cause, or simply from 'careerism'. There can be no doubt that there is a slippery slope (and here we see the link with Milgram), one compromises on small things, commits little crimes that lead inexorably to greater and still greater crimes, and the further one goes down this slope the harder it is to climb back because that would require one to admit that one is a criminal. The best thing, naturally, is to never start down this path, to never compromise on one's principles, even in small things; situationism has nothing to say on this matter, virtue ethics certainly does.

It is worth discussing Lifton's analysis of doubling in a little more detail, as we will see that it helps to bridge the gap between psychology and virtue ethics, in particular how

the cultivation of certain habits can help individuals to resist the temptation to make 'Faustian bargains' of their own. Lifton identifies five characteristics of doubling, I will focus on two of them: First, there is 'a dialectic between two selves in terms of autonomy and connection. The individual Nazi doctor needed his Auschwitz self to function psychologically in an environment so antithetical to his previous ethical standards. At the same time, he needed his prior self in order to continue to see himself as a humane physician, husband, father. The Auschwitz self had to be both autonomous and connected to the prior self that gave rise to it.' And, second, 'a major function of doubling, as in Auschwitz, is likely to be the avoidance of guilt: the second self tends to be the one performing the "dirty work"' (Ibid, p. 419). Lifton is in agreement with Doris in the view that the typical Auschwitz doctor was no 'monster', took no pleasure in the murderous task with which they were assigned and knew that what they were doing was in breach of their previous ethical standards. Auschwitz facilitated doubling in the sense that most of the doctors entering that institution had to double in order to survive there, but Lifton is clear that there is more at play than situational factors. 'The doubling is adaptive to the extreme conditions created by the subculture, but *additional influences, some which began early in life, always contribute to the process*' (Ibid, p. 423-424, my italics). Lifton couldn't be clearer; it requires interplay between the situation and individual factors to facilitate the doubling that allowed relatively 'normal' people to perform such horrific crimes. The lesson we should take away is that if we wish to prepare people (say, our children) so that they will not submit to temptation to make Faustian bargains then we have to start young. And here Lifton is in complete accord with Aristotle when he says 'It makes no small difference, then, whether we form habits of one kind or another from our very youth; it makes a very great difference, or rather *all* the difference' (1103b:23:25). Lifton's analysis of the psychology of the Auschwitz doctors doesn't contradict Aristotelian virtue ethics, it *reinforces* it.

This analysis is easily generalizable to other atrocities such as the ethnic violence in the Balkans that Harman discusses. Ultimately the events are too complex to invite a simple conclusion that either situational or characterological factors are primary. Instead what we see is an interaction between situational influences and the character of individual participants, a destructive reinforcement, a 'positive feedback loop.' It is all very well to take

the situationist line and advocate that we prevent such situations from arising in the first place, but then no serious virtue ethicist would have suggested otherwise. The problem is that advice is of no use to someone who finds themselves in circumstances where situational influences encourage them to commit atrocious acts. Unless we are to say that nothing can be done to prepare oneself (surely too radical a claim even for the situationists), then the question I ask is: Given that something can be done, does this something fit within the framework of virtue ethics?

3.6. Lessons

After reviewing the evidence it is clear that there can be no question as to whether there is significant, widespread, disproportional situational variance in behaviour; the evidence is still very strong that there is. However, under critical analysis it is not clear that the evidence supports the situationist critique of virtue ethics either. Rather, the results provide some clues that will allow us to determine what characteristics an individual would need to have in order to gain some mastery over situational influences, to learn to become more aware of them and to distinguish and respond to appropriate situational cues while resisting others. I will summarise the lessons below, the question for the final chapter will be whether virtue ethics can provide the necessary tools. If it can, then the situationist critique will fail, even though situationist evidence remains largely unchallenged. It will turn out that the situationists simply over-reached in their interpretation.

The discussion of masking raises the question of whether people can become aware of the fact that they are experiencing two (or more) conflicting dispositions, and if so whether this awareness can tip the balance in favour of the disposition (or dispositions) that accord with their values. For example, if I feel that I should help someone, but I am afraid because doing so would put my life in danger, then I may become 'trapped' in an ambivalent state where I don't make a firm decision either way. If, on the other hand, I can acknowledge the competing dispositions, allow myself to feel them while assessing the situation in light of the facts and my values and as a result decide that I should help in spite of my fear, then it seems plausible that this could help tip the scales against fear and in favour of assisting.

The discussion of remote scenarios and the limitations of one-shot studies should lead to questions about the effect of past ethical failings on future behaviour. It seems clear from the evidence that the obedient Milgram subjects in particular felt they failed on their own terms, committed acts they knew they shouldn't. It can often be difficult in the moment to decide how one should act, particularly in the face of uncertainty; but if on reflection after the fact it is possible to *analyse* one's behaviour, assess the cause or causes of one's acknowledge failure and *resolve* to do better in future, then it could just be that such failings act as a spur to future ethical progress. The key question here one of attitude and commitment: are you ready to face your past mistakes and learn from them?

Connected to this, but worthy of separate consideration, is the question of how people cope with the need to act in the face of uncertainty. When discussing ecological validity and construal we noted that the Milgram subjects in particular were unlikely to have a stable, clear view of their situation. It is much simpler to act when one is sure of the situation one is in (even if it should turn out afterward that one was mistaken); it is much harder when one is uncertain, particularly when the situation requires one to make serious ethical decisions, as the Milgram situation would have appeared to the subjects. Even the wisest person, a *phronimos*, cannot be expected to correctly read every situation; what guidance can virtue ethics offer for situations like these?

Finally, when discussing the historical case studies and in particular Lifton's analysis of the Auschwitz doctors we noted the danger in knowingly compromising on one's own ethical values, on making the 'Faustian bargain' in order to gain power, status, wealth or whatever else is on offer. To compromise in this way without repudiating one's values promotes 'doubling', and doubling can enable people to commit terrible acts while deflecting guilt and a sense of responsibility. One might even suggest that it would be better to openly, consciously repudiate one's prior values than to knowingly deceive oneself in this way. We will see in the final chapter that doubling, or any splitting of the self is anathema to virtue ethics, that the 'unity of the virtues' is not a dubious and dispensable postulate of classical virtue ethics, but rather constitutes its very essence.

Far from being clear-cut evidence that virtue ethics is empirically inadequate, the various situationist experiments instead provide some clues as to the characteristics

individuals would need to cultivate in order to resist the potentially baleful influence of situational influences. The questions then, are: What are these characteristics? Can an average person acquire them through practice? Do they fit within a virtue ethical framework? Do they lead to the development of a firm and 'unshakeable character'? These are the questions to be answered in the final chapter.

Chapter 4 – Return to Virtue

Now that we have evaluated the evidence for situationism and thereby identified a number of critical problems for a viable virtue ethics it is time for the final part of this enterprise: the resolution of those problems within a virtue ethical framework. Recall that all of the evidence cited in favour of situationism involves cases of apparent widespread ethical failure on the part a number of individuals which the situationists claim are not explicable within the moral psychology of virtue ethics. The argument is that the results of the various experiments and historical case studies demonstrate that virtue ethics is ‘empirically inadequate’, and therefore those who retain belief in virtue ethics despite the evidence are essentially turning their backs on naturalism; in other words, if the situationist critique is valid one cannot combine virtue ethics with naturalism. Given its purported naturalism is one of the qualities of virtue ethics that attracts so many of its adherents this critique, if upheld, would potentially be fatal to the whole virtue ethical enterprise.

It must be clear by now that I do not believe that the situationist critique is well founded. While I can agree with the situationists that the evidence shows that people in general are often susceptible to major changes in behaviour under the influence of a minor variance in situational factors, and in particular that people in general may often commit acts *they would themselves* consider unethical under the influence of such factors, I disagree with the situationist conclusion that the evidence shows that there is nothing that could be done by the average person to limit the more pernicious instances of situational variance. My position is founded on two considerations: 1) There were no examples presented by the situationists of universal ethical failure, i.e. in each case some individuals behaved ethically by the situationists’ own standards; and, 2) The investigation into the nature of the failures highlighted by my evaluation of the evidence will show that there are methods that are within the capacities of ordinary people which would enable them to avoid such failures.¹³ The first point is immediately clear from the evidence cited by the situationists, which makes it all the more damning that they attempt no detailed investigation into the reasons

¹³ Or, to be more precise, to ‘raise the bar’ for failure, since it is accepted by both sides that there are circumstances in which no one could conceivably behave ethically. See for example Hursthouse’s discussion of tragic dilemmas in Hursthouse, 2001.

why some subjects succeeded where others failed. The second point is the subject of this chapter.

I will proceed by considering once again the ‘lessons’ cited at the end of chapter 3, and investigate what habits of thought and action might be developed in order to counter deleterious situational influences. To simplify matters I will focus on techniques that may assist individuals to act in accordance with their own expressed values, as it is clear that in many of the examples cited the subjects acted contrary to their own values (and suffered as a result). Conversely, if an individual places no value whatsoever on the well-being of other people their failure to act compassionately (in the Milgram experiment for example) couldn’t be put down to situational factors. In the discussion that follows I will be guided by the following principle from Epictetus¹⁴: ‘Happiness and freedom begin with a clear understanding of one principle: Some things are within our control, and some things are not. It is only after you have faced up to this fundamental rule and learned to distinguish between what you can and can’t control that inner tranquillity and outer effectiveness become possible’ (Epictetus, 1995, p.3). It is by focusing our attention in a disciplined way on what is within our power and disregarding those things that are not that we are able to obtain a certain amount of independence from the vicissitudes of those situations we may find ourselves in and become more effective at realising our own values, whatever they may be.

4.1. Self-awareness and the resolution of internal conflict

As a first step I’d like to return to the question raised in response to the pervasive signs of internal conflict exhibited by many of the subjects who behaved unethically in the various experiments and historical studies reviewed in chapter Two, and in particular the obedient subjects in the Milgram experiments. I raised the possibility that this conflict was the result of conflicting dispositions within the obedient subjects and that it was the fact that this conflict was unresolved that lead to their ultimate failure to resist the demands of

¹⁴ Introduced here for the first time, I will appeal to Epictetus regularly in what follows. Though I do not agree with his Stoic metaphysics, after Aristotle I find him the most insightful commentator on virtue, whether ancient or modern.

the experimenter. Indeed, it is difficult to see how the conflict could be explained without appealing to conflicting dispositions. If the subjects *only* felt a disposition to help, or at least not to harm the learner (what Doris and Harman refer to as ‘compassion’), then they would have stopped administering shocks the moment the disposition arose. On the other hand, if they subjects didn’t feel any disposition not to harm the learner then it would come as no surprise that they continued to administer shocks and there would have been no cause for internal conflict. The only sensible conclusion is that those obedient subjects who were conflicted in their obedience were experiencing two or more dispositions that were in conflict and that this conflict remained unresolved.¹⁵

As I indicated previously, the most plausible conflicting dispositions these subjects were experiencing were a disposition to help, or at least not to harm; and a disposition to obey, or at least not to disobey. Note that each of these dispositions will be accompanied by an emotion or by a cluster of related emotions; but for the sake of the analysis let us assume there is only one emotion associated with each. In the case of the disposition to help or not to harm, the associated emotion is likely to be compassion, pity, sympathy or something similar; let us call this the compassionate disposition. In the case of the disposition to obey or not to disobey, the associated emotion is most likely fear (either of the experimenter specifically, or of the uncertain consequences of a refusal to continue); let us call this the fearful disposition. The question then arises: can virtue ethics provide a means to resolve internal conflict of this sort, or at least to mitigate its effects? Given this is a general question to which I will give a general solution, I will take the compassionate and fearful dispositions as a case study, without assuming these were the operative dispositions experienced by the Milgram subjects.

The emotions accompanying the compassionate and fearful dispositions are not mere accompaniments; rather they form an integral part of those dispositions. Without the emotion of compassion for example there could be no compassionate disposition, and the same goes for all the dispositions that are of interest to virtue ethicists and situationists. Our emotional response to a situation registers that something in that situation is of significance

¹⁵ While we have figures on the proportion of obedient subjects who exhibited conflict, it is worth pointing out that the situationists rely on this conflict as evidence that these subjects weren’t vicious. If they were able to torture another person without compunction this would be evidence of a total lack of compassion and would not therefore support the situationist critique.

to us, but it is up to us to determine precisely what is significant, and also whether our emotional response is appropriate.¹⁶ According to Aristotle 'the emotions are not irrational and unpredictable feelings that sometimes overcome us. Instead, they are distinctively human ways of registering the importance of things; to have an emotion is for something to strike one as mattering in some way. The virtuous person has the appropriate feelings because he has the right attitude to things. The courageous person is able to overcome his fear because he values human life; he sees that saving a life, though dangerous, is the right thing to do' (van Zyl in Besser-Jones & Slote 2015 p. 187-188). But no one is innately virtuous (Aristotle is perfectly clear on that point), while we are innately emotional beings, so it must be that a virtuous person learns to have the right attitude to things as they learn to have the appropriate feelings, via a dialectical interaction between emotion and reason. I will argue that it is precisely by reflecting on situations when we experience emotional and dispositional conflict that we can strengthen the connection between our emotions and right reason (by which I mean those reasons that accord with our personal conception of what is right or good), and so increase the chances that we will have the appropriate emotional and dispositional response to situations we face in the future.

In principle there is nothing difficult about what I am proposing, it simply requires us to reflect on what we are feeling when we experience internal emotional conflict. If I am in conflict and experience a disposition to avoid something or someone I can categorize the disposition and accompanying emotion as (broadly) fearful. I can then ask myself: Do I have reason to be afraid? If so, what are those reasons? Is my response proportionate? Are there reasons I should act in spite of my fear, stand and fight say rather than run away? Answering these questions will help to determine whether I should act on my fear or not. Granted, I even after answering them and deciding that I shouldn't succumb to fear I may still run away; I am only human after all. But recognizing what I am feeling and asking questions like these is a step towards gaining some mastery over my emotions and dispositions, towards making them responsive to reasons. If we are committed to this goal of emotional training it will not be easy, not because it requires special talents that only a few people possess, but because it requires *discipline*. One has to be willing to following Epictetus when he advises that 'Every difficulty in life presents us with an opportunity to turn inward and to invoke our

¹⁶ Fear brought on by a phobia would be one example of an inappropriate response.

own submerged inner resources. The trials we endure can and should introduce us to our strengths...On the occasion of an accidental event, don't just reach in a haphazard fashion: Remember to turn inward and ask what resources you have for dealing with it' (Epictetus 1995 p. 17). I would suggest that it is a lack of disciplined commitment to self-improvement of this kind that is largely responsible for the conspicuous absence of virtue in the modern world.

So how does this apply to the Milgram subjects? Well, if they followed the advice above they would first have become aware of their internal conflict and the opposing dispositions. They are still acting in the face of uncertainty as they do not know what will happen if they refuse, but at least they can be sure about what they're feeling. If they'd read Epictetus they might reflect that the actions of the experimenter are outside of their control and therefore not their concern, their only concern should be with what lies within their power. They have the power to continue to shock the learner, or to refuse to continue; and they are responsible for the choice they make. So long as the obedient Milgram subjects were unconscious of what they were feeling they were unable to effectively *act*, in the sense of making a real *decision*. In order to make a decision one must know the reasons why one is acting (the motive) and the goal to be achieved (the aim). If, on reflection, the obedient Milgram subjects *decided* that the correct action was to continue to shock the learner then their failure (if it was a failure) was in their reasoning and not in the fact that they lacked a certain disposition. But one gets the impression that most of them didn't truly act in this sense at all.

The lesson can be summarised: Emotions are important because they alert us to the presence of situational factors that are significant, yet untrained emotions are unreliable and can lead us astray. By reflecting on our emotional and dispositional responses to situations we can analyse these to see whether the actions they dispose us towards are, on balance, in accordance with our own values ('right reason'). If they are then there is no problem, if they are not then we are in a position to bring all the tools of the conscious mind into play in order to tip the balance in favour of the action that is in accordance with our values (including memories of past failings, anticipated regret if we knowingly commit a shameful act, etc.). Over time this process gets easier, becomes habitual, and eventually (if Aristotle and Epictetus are correct) we may reach the stage where our emotions and our reason (in

particular, our values) are in complete accord. Given that the situationists rely on the fact that the failing of the obedient Milgram subjects was not the result of flawed values, the process outlined above demonstrates that there are means within the power of the average person to overcome a 'Milgram situation'; it simply requires a great deal of disciplined preparation ahead of time.

4.2. Reflective thought and ethical development

The analysis above points towards the importance of reflective thought in virtue ethics. Turning 'the mind's eye' inward, assessing and evaluating thoughts, feelings and actions, taking responsibility for personal development; this is fundamentally what virtue ethics is about. The irony is that ethical failings like those revealed by the Milgram experiments don't so much demonstrate the absence of virtue as demonstrate the *need* for it. All the situationist experiments provide compelling evidence that it is not sufficient just to believe that certain acts are right and others are wrong; one must also be capable of performing the act that one believes to be right, and in difficult circumstances this requires more than mere 'willpower'. Our ethical failings, and *our* acceptance of them *as* failings should motivate us to reflect on the reasons why we failed and what might be done to avoid such failures in the future. If we are genuinely committed to our values then we must be committed to acting in accordance with them, if we regularly fail to act in accordance with our values and we do nothing to remedy these failures we should question whether they are really our values at all. Perhaps we just profess them because it is expected of us, or to make us feel like we are decent people, or from habit; these are ignoble reasons to profess values. So I would like to say a little more on the reflective nature of virtue ethics, as it relates to ethical failures.

Questions arise when something goes wrong. So long as all our actions succeed we have no reason to reflect on the nature of our actions and by extension on our own nature. In this way we can see that the Milgram experiments could have been precisely the spur that some of the subjects needed to critically reflect on their own nature and to start down the path of virtue. As Aristotle made abundantly clear, 'none of the moral virtues arise in us by nature' (*NE* 1103a: 25), we can only acquire them if we intentionally develop them. Why

would anyone spend time and effort acquiring virtues? I would like to suggest the following: because they recognize that there is a gap between the person they feel they 'should' be, and the person they are revealed to be through their actions. I say 'should' because I do not wish to imply I'm referring to some categorical imperative, there is an imperative but it may be strictly subjective. Regardless of your position on the moral spectrum, *if* you have values that should be realisable (however acquired), and *if* you regularly fail to act in accordance with those values then you have four options: 1) continue to live in a discordant state, 2) repress the unpleasant knowledge of your failings, 3) renounce those values you cannot live up to; or 4) *become* someone who can live in accordance with those values. This is a personal choice; I would not want to make it for anyone else even if I could, and anyone who chose options 1 through 3 would find no value in what I am about to say (or indeed, in this paper as a whole).

Anyone who chooses option 4 is committing to achieving personal freedom, where freedom is understood not in the sense of 'free will' (an unintelligible concept in any case), or as 'the right or ability to do whatever you please. Freedom comes from understanding the limits of your own power' (Epictetus 1995, p. 21); it comes from developing your own resources and making the most of your natural capacities. Freedom in this sense comes in degrees and is attained to the extent that you accept your natural limitations (hence you are not entirely free if you believe in 'free will', any more than if you wish to attain immortality) while developing your natural abilities to the fullest extent possible. When understood in this sense I wholeheartedly agree with Epictetus' statement that freedom 'is the only worthy goal of life. It is won by disregarding things that lie beyond our control. We cannot have a light heart if our minds are a woeful cauldron of fear and ambition' (Ibid, p. 26). Of course, it is no mean feat to determine what things are within our control, but the first step is to *reflect*.

We come to know ourselves the way we come to know others, over time. If we never reflect on our past actions we will never have a chance of distinguishing good from bad habits, and so we will never have a chance of reinforcing the former and weakening the latter. A habit cannot be changed in the moment, but with discipline and conscious effort all but the most deeply ingrained habits can be overcome, and newer, healthier habits adopted. It is only 'By identifying one's own temperamental and behavioral tendencies, one

is able to gain a measure of potential control over them...This kind of self-knowledge is not knowledge of an unchanging essence or a metaphysical fact. Rather, it is knowledge of the impermanent, varying, and otherwise obscured beliefs and desires that shape one's agency' (O'Hagan in Birondo & Braun 2017, p. 107). So long as we allow ourselves to be the victims of our unexamined beliefs and desires we limit our agency, in particular our ability to act according to our values. The path to freedom will be different for each individual according to their natural abilities and acquired habits, but the process will be much the same: continual reflection on past actions and in particular the examination of motives to ensure they accord with our values. The aim is continual improvement, even if it occurs in small increments. As Epictetus says, 'Do not measure yourself against others or even against your ideal self. Human betterment is a gradual, two-steps-forward, one-step-back effort' (1995, p. 99). Our point of comparison should be our prior self. So long as we are improving we are moving in the right direction, focusing on an ideal state that is not currently achievable is at best a distraction.

It is worth emphasising the importance of reflective thought and 'tracking one's self-conception and the reasons for which one is inclined to act' (O'Hagan in Birondo & Braun 2017, p. 107), even (perhaps especially) in those instances where we perform the 'correct' act. If, for example, I return a lost wallet, this is (presumably) the correct act. But if I do so for the sake of a reward then I've performed the correct act for the wrong reasons, and on reflection I should note this and resolve to disavow those reasons in future. In this case it may seem no harm was done, but *allowing* oneself to act on the wrong reasons, to develop a *habit* of it, may have insidious long-term effects. Perhaps there was no harm in this case because greed happened to motivate the same act that virtue would, but there is no guarantee that this will always be the case. It is better, therefore, to be aware when there are multiple reasons for performing the same act and to identify only with good reasons. Any serious consideration of virtue ethics is likely to raise the suspicion that people are no better in our day than they were in the past, despite the fact that they are more 'moral' (at least according to Sam Harris and others of his ilk). The pressures to conformity must be so much greater now than they once were, the average person is far more aware of what is expected of them and the means for surveillance both by the state and our peers are far more comprehensive. But to perform the right acts merely to conform to expectations is

not virtue, and the moment expectations change the true self is likely to emerge. The readiness, even enthusiasm, with which so many 'civilised' people perform monstrous acts during wartime seems ample evidence of the dangers of relying on conformity rather than virtue; a virtuous person's actions depend on their own evaluations and do not vary according to the expectations of others.

4.4. *The unity of the virtues*

I want to turn now briefly to discuss the unity of the virtues thesis before undertaking a re-examination of the nature of virtuous dispositions. The 'unity of the virtues' thesis has its origin in Aristotle, and is a natural consequence of the fact that the virtues, at least according to Aristotle, are habits that produce *eudaimonia*. If *eudaimonia* is to be a unitary state then, and all of the virtues are necessary to produce it, then it stands to reason that the virtues must form a unitary whole. This notion has fallen out of favour with many contemporary ethicists however, and some moreover consider that it is entirely dispensable; they claim that virtue ethics is perfectly viable without it. Now, I won't dispute that you can describe a theory based on character traits that you call 'virtues' that are not integrated and thereby create a 'virtue ethics' that does not entail the unity of the virtues, but I would argue that in doing so you would be discarding one of the most distinctive properties of classical virtue ethics, and one of the guiding principles that enables us to identify which traits are virtues in a non-arbitrary fashion. I will present a couple of examples to illustrate my point before explaining why the virtues must be united.

First, consider Doris' claim that 'contemporary writers like McDowell have endorsed a unity thesis, where the apparently discrete virtues turn out to be different manifestations of a "single complex sensitivity". Such claims have struck many commentators as badly contrary to fact. It is easy to imagine a person who is, say, courageous and intemperate; indeed, it is tempting to think that such a person is courageous in part *because* she is intemperate' (Doris, 2002, p. 20-21). This is a simple misunderstanding on Doris' part, mistaking an act in conformity with courage from a truly courageous act (apart from the problem of whether my being able to 'imagine' something counts as evidence for it). Granted that a person who does not have full virtue may perform a courageous act, and at

another point perform an intemperate act, it is clear enough that they cannot perform an act that is *both* courageous *and* intemperate. To perform a truly courageous act is to recognize that virtue requires you to perform some action the prospect of which causes you fear, and then to choose to perform that act in spite of this. What Doris seems to have in mind is a 'rash' act, which would be intemperate but which Aristotle explicitly contrasts with true courage (c.f. *NE*, 1104a: 10-25).

A second objection comes from Christine Swanton, a proponent of 'pluralistic' virtue ethics. Her view is that 'Most contemporary versions [of virtue ethics] have been developments of predominately the Aristotelian conception of living well...By contrast a pluralistic virtue ethics believes that a plurality of conceptions of living well have strengths, and should take their place in an adequate comprehensive form of virtue ethics' (Swanton in Besser-Jones & Slote 2015 p. 210). A pluralistic virtue ethics 'will deny the 'Unity of the Virtues' thesis, but I argue in this section that such an ethics can nevertheless accept that the virtues are to some extent integrated: they are not 'fragmented.' What is the unity of the virtue thesis denied by pluralistic virtue ethics? Well here is one classic formulation: Aristotle's claim that to have one virtue you must have them all' (Ibid, p. 211). Now it seems to me that Swanton is also mistaken, though not to the extent we saw with Doris. Aristotle's virtue ethics is predicated on a very specific notion of what it means to 'live well'; his unity of the virtues thesis *derives* from that conception. If you introduce a new conception of 'living well' that depends on other character traits then that is perfectly fine, I just suspect that Aristotle would say they are not virtues *by definition*. Swanton seems to hold that there is some middle-ground between states that are character traits that are 'fragmented' and those that are integrated; I am going to argue she is wrong.

It is interesting to note that Swanton uses the word fragmented, since this is the very word Doris uses to describe the nature of the only types of human character traits he is willing to accede to. For example, he says 'I...contend personality should be conceived of as *fragmented*: an evaluatively disintegrated association of local traits' (Doris, 2002, p. 64). I am willing to agree with Doris to some degree; fragmented character traits in this sense are the norm. But this only because evaluatively integrated character traits is the goal of virtue ethics, rather than the starting point. The whole point of the virtues, what defines them *as* virtues is, in my contention, the fact that they are traits that are capable of being integrated;

it is this fact that distinguishes them from vices for example. One can be courageous and temperate and liberal (in Aristotle's sense) and just because all these traits lie on the mean, which allow each to influence the virtuous person's action while leaving the final decision on how to act to *phronesis*, 'which balances different ethical concerns with one another: sometimes courage involves standing and fighting, and sometimes it involves stepping down; one can know the difference in a particular circumstance only by understanding what is worth fighting for, and at what cost, and this means having an overall conception of the good (Russell 2009, p. 22). The vices, on the other hand, are in perpetual discord because they are extreme states. Vices conflict with each other; it is not possible to be consistently avaricious *and* intemperate because these two traits interfere with each other. The vices interfere with each other, and destroy the very means of their satisfaction, i.e. the possessor of vice, hence Aristotle says 'evil destroys even itself, and if it is complete becomes unbearable' (NE 1126a: 12). The virtues are just those traits that can be integrated and the development of which lead to integration of the self; the vices are just those traits that cannot be integrated and the development of which lead to the fragmentation of the self. This has direct implications for the situationist critique, as it is only the unitary self that is robust. This conception of virtue also links virtue ethics with those (such as Carl Jung) who see healthy psychological development as being the integration of the parts of an originally fragmented self into a true individual (c.f. Jung, 1956, among others). It also presents a remedy to the 'doubling' that Lifton identified as providing 'an overall key to human evil'; the more integrated the self the less vulnerable it will be to the temptations of that 'Faustian bargain'. Evaluative integration is no mere appendage to Aristotelian virtue ethics, it is the goal itself.

4.5. Virtuous dispositions reconsidered

We are now in a position to return to the nature of virtuous dispositions in view of the discussion above. It will become clear that the situationists (among others) held an erroneous view of the dispositional nature of virtuous character traits; once this error is corrected it we will see why the various situationist experiments were ill-suited to the task

of identifying virtues even if they had been present, and also why we'd expect them to be rare.

Aristotle repeatedly relates the virtues to habit; 'moral virtue', he says, 'comes about as a result of habit, whence also its name (*ēthikē*) is one that is formed by a slight variation from the word *ethos* (habit)' (NE 1103a:14-16). He also speaks of virtues (and vices) in dispositional terms, for example in the following passage where he contrasts them with the passions: 'in respect of the passions we are said to be moved, but in respect of the virtues and the vices we are said not to be moved but to be disposed in a particular way' (Ibid, 1106a:4-6). So perhaps it isn't surprising that the situationists seem to consider that the virtues are certain habitual ways of acting in response to situational cues. A habitual action is a stereotype, there is a simple cause-and-effect relationship between stimulus and action; I feel anxious so I bite my nails (at least I used to). The problem is that habitual action isn't intelligent, it is the very opposite. There is no place here for *phronesis*, instead we just have a knee-jerk response to a certain cue. It seems clear that the situationists, when analysing the Milgram experiments for example, would only accept identical *actions* across different situations as evidence of the presence of virtue. While certain actions may be indicative of virtue (and by no means do I wish to imply that the obedient subjects were *behaving* virtuously), they are only part of the story.

The first point to note is that Aristotle never identifies virtues *as* habits, nor does he identify them *as* dispositions. For example, if we consider that quotation in the first line of the above paragraph we can see that moral virtue comes about 'as a result of habit'. Likewise, virtues cause a virtuous person to be 'disposed in a particular way', rather than being the disposition itself. We should therefore consider the possibility that the habitual aspect of virtue obtains to something other than the actions consequent of it, and that one may have a disposition to act in the same way a virtuous person would without possessing the requisite virtue; i.e. the virtue is sufficient for the disposition, but not necessary.

Now the way towards a virtue ethical resolution of the apparent dilemma posed by the situationists becomes clear: the situationists believed that there was an inconsistency between the behaviours of the Milgram subjects in the 'experimenter on the phone' version as opposed to the 'experimenter in the room' version that wasn't explicable in terms of the

underlying psychology of virtue ethics. The assumption is that the subjects in the former version showed significantly greater compassion than those in the latter version, and the evidence for this is the fact that far more of the subjects in the former stopped before reaching the full shock level. From this it is inferred that more of the subjects in the former had a disposition to act compassionately and therefore that they *were* compassionate. But as we've just seen, while in certain situations we may infer that a virtuous person will feel a particular disposition, inference in the other direction is not possible without further information that the Milgram experiments, and the situationist experiments in general, are unable to provide.

The issues with the situationist experiments, and the reason why they ultimately fail to support the situationist critique, is that they are unable to distinguish between virtue proper, and mere continence. It is striking that the situationists do not address this problem, given the pains that Aristotle takes to distinguish action that is merely in accordance with virtue from true virtue. When he speaks of courage, for example, he states that 'The man...who faces and who fears the right things and from the right *motive*, in the right way and at the right time, and who feels confidence under the corresponding conditions, is brave' (*NE* 1115b: 17-20, my italics); but the situationists have no way of determining the motive of the disobedients in the Milgram experiments. Part of the problem lies in the fact that 'many actions that look dissimilar to an outside observer may in fact be regarded as instances of the same thing from the agent's point of view' (Russell 2009, p. 260). The habitual aspect of virtue, the aspect that guarantees its robustness, lies neither in the actions themselves, nor in dispositions per se, but in ways of thinking and feeling. As we saw earlier in this chapter, the virtuous person will continually reflect on their own actions, examine their motives, recognize that they often fail to meet the standards they have set for themselves and strive to do better in future. Virtue is the very opposite of habitual response to certain behavioural cues, it results from and entails a dynamic interaction between the individual and the situations the encounter. This is the "cognitive-affective" picture of virtue favoured by Russell (2009) and Hursthouse (2006), and it contrasts favourably with the dispositional view that is the target of the situationist critique. Under this interpretation, 'A virtue, unlike a mere habit, is a disposition to act *for reasons*, and so a disposition which is exercised through the agent's practical reasoning; it is built up by

making choices and exercised in the making of further choices' (Annas in Copp 2006, p. 516). It is this responsiveness to reasons that ensures that virtue is sensitive to situational changes, within the necessary limits of human fallibility. Dispositions, rather than being components of the individual's personality, arise as required in response to situations that call them forth; the dispositions of the virtuous are just as flexible as reason itself.

4.6. *The purpose of virtue and the meaning of ethics*

I have demonstrated that the situationist critique fails to undermine the viability of Aristotelian virtue ethics. At best it undermines the dispositional theory of virtue, which sees the virtues as broad-based habitual responses to situational factors. On the other hand, the evidence the situationists present favours the "cognitive-affective" theory of virtue which see the virtues as complex habitual modes of thinking and feeling which ensures that the virtues are responsive to reasons. Given that a careful reading of Aristotle indicates the cognitive-affective view accords with his own, it would seem the situationists have merely given us more reason to marvel at Aristotle's insight in developing an ethical theory that still remains viable and excites passionate debate to this day.

With the situationist threat allayed, I now want to ask a question that might have seemed more at place at the beginning of this paper: is virtue ethics a moral theory? Perhaps that seems too simple, or too obvious a question to need an answer, but perceptive readers may have noticed that I've been careful to speak of ethics rather than morality throughout this paper. The terms 'morality' and 'ethics' are often employed interchangeably, but for some time I've believed there is value in using these terms to refer to separate, though often interrelated enterprises; and writing this paper has only strengthened this belief. If we can consider Kantian deontology and classical utilitarianism as representative moral theories it seems perfectly normal for a moral law to require our compliance whether we like it or not. On the other hand, the idea that we should be virtuous whether we like it or not is incomprehensible; in order to *be* virtuous one must *want* to be virtuous because virtuosity is a skill that requires discipline and constant practice. Moreover, according to both Aristotle and Epictetus, it is irrational not to want to be virtuous because virtue *is good for us*. Epictetus, for example, insists that 'The virtue that

leads to enduring happiness is not a *quid pro quo* goodness. (I'll be good "in order to" get something.) Goodness in and of itself *is* the practice *and* the reward' (Epictetus 1995, p. 103). As I hinted earlier, I prefer to view virtue ethics as a form of character training, a way of integrating our thoughts, feelings and inclinations so that we can more effectively act in accordance with our values. By helping to reduce the sources of internal conflict and by encouraging us to make our own decisions and learn from them, virtue ethics helps us to develop into genuine *agents*; by which I mean, beings capable of making decisions and acting on them. Ethics conceived in this sense could be used in conjunction with a moral theory, which would provide the values to be acted on while ethics ensures the agent is capable of compliance. However, I prefer to view ethics as an end in itself, because one who develops in virtue and wisdom invariably realises that there is nothing to be gained through theft, cruelty, avarice, and indeed any 'evil' act. As Epictetus says, 'Authentic happiness is always independent of external conditions' (1995, p. 26).

Conclusion

In the preceding discussion I have demonstrated that the situationist critique ultimately fails to undermine virtue ethics. The initial plausibility of the critique stems from a misunderstanding of the nature of virtuous traits, identifying them *as* dispositions to act in certain distinctly virtuous ways, rather than recognizing that dispositions to act virtuously flow from the virtues. This misunderstanding is one which unfortunately is often fostered by the defenders of virtue, such as when Hursthouse states that the ‘full Aristotelian concept of virtue is the concept of a complex character trait, that is, a disposition that is well entrenched in its possessor’ (Hursthouse in Dreier 2006, p. 101-102). Identifying virtues *as* habits rather than being *acquired* by habituation, and identifying them *as* dispositions rather than *entailing* certain dispositions leads to a picture of the virtues as inflexible, knee-jerk responses to certain situational stimuli. The situationists, with some help from virtue theorists, mistook this inflexibility for the ‘firm and unchangeable character’ that Aristotle famously referred to. When their analysis of the situationist experiments failed to find widespread evidence of such traits they concluded that virtue ethics must be false; but in fact they were looking for the wrong thing.

True virtues are integrated habits of thought and feeling, consciously developed over time and typically the outgrowth of seeds that were planted in our youth before we had any conception of them. They can only come about as a result of a deep commitment to our own personal development, in particular to maximizing our agency by reducing sources of internal conflict so that we may act in accordance with our values. They are habits that direct our attention to certain concerns (for example: to help others in need, to develop our abilities, to never allow fear to prevent us from doing what we feel we should) so that we are perceptive to opportunities to be virtuous. Because our reason and emotions are integrated, virtue in this sense is amenable to reason. At the highest levels of virtue our emotions and our reason are in harmony, practical reason interprets the situations we are in, determines which virtues apply, balances their demands and makes a decision, from which a disposition to act follows. This “cognitive-affective” theory of virtue delivers the flexibility that the dispositional view lacks; it guarantees that the virtuous person can tailor

their response to the situation. The complexity involved, the immense discipline required to bring reason and emotion into harmony, and the fortuitous upbringing necessary to set the virtuous person on the right path from youth are reasons enough to expect that virtue will be extremely rare in an age that has no time for such quaint notions. But even if virtue was common, even if it had been wide-spread among the subjects of the situationist experiments, they still would not have seen it; for they were looking in the wrong place.

Bibliography

Alfano, M: *Explaining Away Intuitions About Traits: Why Virtue Ethics Seems Plausible (Even if it Isn't)*, Review of Philosophy and Psychology, 2001 2: 121-136

Alzola, M: *Character and Environment: The Status of Virtues in Organizations*, Journal of Business Ethics, Vol. 78, No. 3, 14th International Symposium on Ethics, Business and Society (Mar., 2008)

Aristotle: *The Nicomachean Ethics*, Oxford University Press, 2009

Athanassoulis, N: *A Response to Harman: Virtue Ethics and Character Traits*, Proceedings of the Aristotelian Society, Vol. 100 (2000)

Besser-Jones, L. & Slote, M. eds.: *The Routledge Companion to Virtue Ethics*, Routledge, 2015

Birondo, N. & Braun, S. eds.: *Virtue's Reasons: New Essays on Virtue, Character and Reasons*, Routledge, 2017

Copp, D. ed: *The Oxford Handbook of Ethical Theory*, Oxford University Press, 2006

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

Dreier, J. ed.: *Contemporary Debates in Moral Theory*, Blackwell Publishing, 2006

Epictetus, *The Art of Living*, Harper One, 1995

Harman, G: *Moral philosophy meets social psychology: Virtue ethics and the fundamental attribution error*, Proceedings of the Aristotelian Society, New Series, 119: 316-331, 1999

Harman, G: *Skepticism about Character Traits*, The Journal of Ethics, Vol. 13, No. 2/3, 2009

Hursthouse, R: *On Virtue Ethics*, Oxford University Press, 2001

Jung, C.G: *Symbols of Transformation, Volume 5 of the Collected Works of C. G. Jung*, Princeton University Press, 1956

Lifton, R: *The Nazi Doctors: Medical Killing and the Psychology of Genocide*, Basic Books New York, 2017

Ross, L & Nisbett, R.E: *The person and the situation*, Cambridge University Press, 1991

Russell, D: *Practical Intelligence and the Virtues*, Oxford University Press, 2009