

A Humean metaphysics of character

Ethan Chambers

Submitted for the degree of PhD

Department of Philosophy

University of Cardiff

2018

Abstract

The primary aim of this thesis is to analyse the tension between Humean epistemology and Humean ethics. The tension is identified as being due to the strict Humean empiricist epistemology, which severely limits the scope of ethical investigation. A key point of contrast, between Hume's bundle theory of the self and his commitment to long-term durable character traits, is analysed, to exemplify the difficulties. Humean empiricism leads directly to a bundle conception of the self, one that is comprised entirely of observable impressions and ideas, yet it is unclear how this conception is related to durable character traits, which are an integral part of Humean ethics.

The solution offered is that, to be understood coherently, Humean metaphysics should incorporate aspects of powers-based ontologies to replace summative models. Specifically, character traits such as virtues and vices should be understood as tending towards an outcome, rather than adhering to a standard stimulus-response dispositional model. It is argued that this reconceptualization can be achieved without contravening Humean epistemology in any significant way, and that it can be done while maintaining the basic Humean principle that all ideas must precede from a corresponding impression. This thesis will conclude that adopting the metaphysics of powers in a limited sphere of investigation can lead to a stronger Humean metaphysics of character and eliminate contradictions from Humean philosophy as a whole.

Contents

Introduction 7

Chapter 1: Investigating the tension 13

1.1: The problem of character traits 14

1.2: Hume's basic empiricism 19

1.3: The role of the passions 24

1.4: Hume's theory of belief 27

1.4.1: What is a belief? 27

1.4.2: How is a belief formed? 29

1.4.3: How is a belief justified? 32

1.5: General exceptions and induction 35

1.6: The laws of nature 40

1.7: Conclusion 44

Chapter 2: The mind-self distinction in Hume 46

2.1: The mind's tacit role in Humean philosophy 46

2.2: Hume's scepticism 50

2.3: The role of self and mind 52

2.4: The peculiarity of character traits 56

2.5: Conclusion 58

Chapter 3: Hume and virtue ethics 59

3.1: The Virtues 60

3.1.1: What is a Humean virtue? 61

3.1.2: The nature of Humean virtues 64

3.1.3: The metaphysical problem 66

3.2: The role of the passions in virtue ethics 68

3.3: Moral taste 70

3.3.1: Hume's theory of taste 71

3.3.2: Moral taste as a distinct faculty 74

3.3.3: The role of reason in ethical judgements 77

3.3.4: The mixed view 79

3.3.5: Conclusion	82
3.4: Sympathy	83
3.5: Conclusion	85
Chapter 4: Character	87
4.1: Hume's position on character	89
4.2: Characters and selves	92
4.3: Character as behaviour or reputation	94
4.3.1: Baier	95
4.3.2: Kupperman and patterns of action	98
4.4: Character as mental properties	101
4.5: Character and personality	104
4.6: Character and skills	105
4.7: Moral and intellectual traits	107
4.8: Conclusion	110
Chapter 5: Scepticism about character	112
5.1: Harman's scepticism	113
5.2: Harman's data	114
5.3: Doris' localised traits	123
5.4: Conclusion	128
Chapter 6: Powers and Humean metaphysics	129
6.1: Hume's account of dispositions	130
6.1.1: The causal basis for dispositions	132
6.1.2: The stimulus-response model	134
6.2: The neo-Humean account	138
6.3: Neo-Aristotelian powers	140
6.4: Capacities and skills	143
6.5: The empirical accessibility of powers	146
6.6: Powers and character traits	153
6.7: The SCA revisited	156
6.8: Conclusion	157

Chapter 7: Integrating powers into a Humean metaphysics of character and the self 159

7.1: The nisus 159

7.2: The epistemology of powers as character traits 168

7.3: The copy principle 171

7.4: Powers in Humean ethics 174

7.4.1: Nativism in Humean character traits 174

7.4.2: Where do powers fit in Humean moral psychology? 176

7.4.3: Powers and causation 179

7.5: Conclusion 184

Chapter 8: Resolving the tension 185

8.1: Psychological determinism 185

8.1.1: Hume's argument against the liberty of indifference 187

8.1.2: The dispositional modality 192

8.1.3: Two models of causation 194

8.1.4: Necessity and blame 197

8.1.5: Conclusion 199

8.2: Humean moral psychology and powers 199

8.2.1: The will 203

8.2.2: Passions 206

8.2.3: Taste and sympathy 207

8.2.4: Personality traits 209

8.2.5: The copy principle and other mental operations 209

8.2.6: Belief 211

8.3: Solving the problem of character 213

8.4: Conclusion 214

Bibliography 216

Acknowledgements

I would like to thank my parents, Julia and Gregory, for their unwavering support throughout the writing of this thesis, and throughout my entire life, without which this would not have been possible.

My deepest appreciation also goes to Jonathan Webber, whose keen insight, wealth of knowledge and warm encouragement have benefitted me immeasurably during my academic life. Feedback and support from Richard Gray and Chris Norris has also been indispensable.

Finally I would like to give special thanks to Rhian Rattray, who has been a constant source of reassurance and assistance in her role as postgraduate manager at Cardiff University.

Introduction

The goal of this thesis is to identify a tension in Hume's philosophy, between his metaphysics and virtue ethics, to diagnose its roots in his epistemology and to begin to resolve that tension by appealing to contemporary accounts of the metaphysics of powers. It will be shown that doing so can result in a more coherent Humean philosophy, one that does not contradict any of Hume's key principles.

Hume's epistemology is grounded in his theory of impressions and ideas, which he contends form the basis of all knowledge (1978, p.4). Every simple idea we possess must derive from a corresponding simple impression if it is supposed to represent anything at all. This is Hume's copy principle, commitment to which is a central factor in any Humean philosophy.

Complex ideas, such as an idea of Europe, can be built from multiple simple or complex impressions, but must necessarily be grounded in impressions to have content. A belief, for Hume, is justified to the extent that experience supports it, while true knowledge is a matter of relations between ideas – chiefly mathematical truths.

Following this epistemology led Hume to argue for another cornerstone of his philosophy: the bundle theory of the self. In the *Treatise* (1978, p.252), Hume applies his empiricism to the idea of the self, introspecting upon his own mental state to find an impression of that subject on which to base the idea. Instead, Hume found “nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in perpetual flux and movement”. Hume thus rejects that there is any

substantial self beyond this bundle of impressions and ideas, since such an idea is unintelligible without an impression to ground it.

However this rigid epistemology is apparently contradicted by Hume in his most famous works, the *Treatise* (1978) and both *Enquiries* (1975), by his commitment to virtue ethics. Hume's ethical theory is focused on the evaluation of character traits; virtues and vices which together constitute ones' moral character. These traits are described by Hume as long-lasting and 'durable' (1978, p.575), in contrast to the fluctuating impressions and ideas of Hume's earlier passage.

However it is unclear that one can gain an impression that would ground the idea of durable character traits, or of moral character. Traits do not seem available to introspection, nor is there a reliable way of gaining an impression of them through a third-person perspective. Here the tension is most explicitly demonstrated: Hume's ethics rely on long-term character traits, which themselves are at odds with Hume's strict epistemology. We may have complex ideas representing character traits, derived from myriad simple impressions; if this is the case some work will need to be done to show how belief in such an idea is justified.

A related issue is that Hume believed human behaviour to be fairly regular and predictable, as evidenced by his writings on 'moral evidence' (1978, p.404) and the example of a wallet left on Charing Cross Station that is sure to be taken (1975, p.91). Hume here illustrates his reliance on the regularity of human behaviour, one that is rather mysterious if all there is to agents is a bundle of frequently changing elements.

This thesis aims to show that the tension can be resolved, but that doing so requires Hume's character traits to rest on a metaphysics that Hume does not provide. Hume advocates a stimulus-response model of dispositions, which has the benefit of being metaphysically uncomplicated and so congruous with Hume's epistemology. However this model is unsatisfactory when seeking to explain how character traits operate in Hume's ethics. Hume's description of character traits as durable properties of an agent differentiates them from physical properties in several crucial ways. Primarily, it is unclear how one is to understand the stimulus that is required for a character trait to manifest, or what the appropriate response is supposed to be.

More promising is the metaphysics of powers, which has seen a resurgence of popularity lately in neo-Aristotelian literature. Powers, as understood by Mumford (2013), cannot be understood by the stimulus-response model. Rather they are constantly tending towards their outcome, and do possess real causal efficacy. However Hume's epistemology, which coheres with a stimulus-response model of dispositions, is not obviously amenable to the idea of powers, the idea of which Hume notes is not rooted in an impression from the external world (1975, 63). As a result Hume and neo-Humeans are often seen as standing in direct contrast to powers-based neo-Aristotelian accounts.

This thesis will argue that Humeanism does not exclude the existence of real causal powers, and that powers-based metaphysics can be incorporated into Hume's philosophy in a certain sphere. To achieve this without contradicting Hume's epistemology, an impression of power must be

available to us. Hume (1975, 67n) does recognise that there is such an impression, but he takes it to be of little import and so only gives the matter a short mention in a footnote. This impression he refers to as the *nisus*, which is the feeling of straining one is aware of when physically exerting oneself. Mumford (2013) refers to this as proprioception, and uses it to form the basis of an empirical argument for the existence of powers. It will be argued here that the *nisus* is the impression that can give rise to the idea that character traits can be conceived as powers in a Humean ontology, as the pressure exerted by character traits is analogous to the *nisus*. This argument does not seek to establish the existence of powers in the external world, for the reasons Hume gives in the footnote. Certain mental, agential powers are empirically available to us in a way that the physical powers of objects cannot be, and so are compatible with Hume's epistemology.

It will further be argued that a mind, as distinct from the bundle of impressions and ideas that constitutes the self, is the bearer of these powers. Though Hume proposed no theory of the mind himself, one can find evidence, in Hume's work on character in particular, that Hume supposed the mind to play an important role in an agent's psychology. Memory, reason and the copy principle are all examples of concepts that would be difficult to make sense of without the idea that there is more to our mental landscapes than introspectable ideas and impressions. This does not contradict bundle theory – the self on this picture remains a bundle of impressions and ideas discoverable only through introspection – but is an addition to it. It will be concluded that the mind possesses powers,

understood as character traits, that shape our behaviour, explaining why our behaviour is as regular and predictable as Hume believed.

A potential objection to this argument is that other areas of Humean philosophy will not be compatible with this reworked understanding of Humean ethics, and that there is little point in showing that Hume's virtue ethics and metaphysics are compatible at the expense of problematizing another fundamental aspect of Humean philosophy. Hume's argument against freedom of the will, or liberty of indifference, may be thought in jeopardy owing to its reliance on psychological determinism, a thesis which is often thought to be opposed to Aristotelian powers. It will be argued that understanding character traits as mental powers does not entail a rejection of determinism, since real causal powers can be understood as deterministic properties. Thus one can hold a coherent Humean ethics, epistemology and metaphysics without contravening other significant areas of Hume's philosophy.

This thesis is not intended as a reading of Hume. His comments in the *nisus* footnote, for example, make it clear that he did not regard the *nisus* as sufficient to establish a broader idea of powers. Nor is my claim that Hume's philosophy entails the conclusion presented here. It is not argued that Humeans must accept this solution, or that this approach is the only way to properly conceive of character traits as Hume describes them. However the argument is intended to show that these tensions do exist in Hume's work, but that they are not fatal to the Humean project, since importing aspects of powers metaphysics is one way of overcoming them, allowing one to hold an

epistemology, ethics and metaphysics that reflects the fundamentals of Hume's work.

Chapter 1: Investigating the tension

This chapter shall give a brief account of Hume's bundle, or causal, theory of the self, paying close attention to any aspect of it that clashes with virtue ethics, specifically the idea of enduring character traits. This chapter is not intended to resolve the tension, just to detail the primary problems that this thesis will attempt to solve.

The difficulty for the Humean will be identified as the problem of character traits; of how we are supposed to obtain an intelligible idea of the constituents of moral character when there is no corresponding impression of them, exemplifying the general tension between virtue ethics and metaphysics, rooted in epistemology, that motivates this thesis.

This chapter will also analyse the building blocks of Hume's epistemology; beginning with basic Humean epistemological principles and covering impressions, ideas and the passions, questioning whether virtues and vices can be reduced to any of these in a way that would resolve the disparity between bundle selves and ethical character. It will be concluded that they cannot, so must exist outside of the bundle self. Hume's theory of belief will also be analysed to clarify what a 'justified belief' could be in a Humean framework, and to better answer the question of whether a Humean can justifiably believe in real causal powers later on. The role of induction will then be addressed to give a full scope of Hume's epistemological methods. Finally this chapter shall draw attention to instances in which Hume appeared to provide exceptions to his own epistemology, focusing primarily on the copy principle and his reliance on the laws of nature in 'Of Miracles'.

1.1: The problem of character traits

This section will highlight the fact that Hume's theory of the self is at odds with his virtue ethics, and that the root cause of this is epistemological.

Hume's central epistemic precepts entail the bundle theory of the self, a conception of the self which does not have room for long-lasting character traits.

Hume's causal theory of the self was a radical departure from the essentialist thought that was popular at the time and which still holds intuitive appeal today. Roughly, essentialist theories posit that every individual has one self; a certain essence that begins when we are born and ends when we die. This identity withstands all the physical and mental changes in between, and consists of a core of essential properties, to which non-essential, or accidental, properties are added and subtracted. These accidental properties, such as a fondness for wine, a brief bad mood, or fair hair may come and go without any real change in identity. However essential properties, the criteria for which differ across accounts, cannot alter without constituting a change in identity (Matthews, 1990 p.251).

Aristotelian essentialism espouses something similar to this; he writes that there is a primary substance to all beings that denotes 'what a thing is' (Warrington, 1956, p.167), and which provides it with its essential properties. Virtue ethics seems to couple nicely with this view of the self, giving Aristotle a metaphysical grounding for his ethics, which Hume appears to lack. Character traits, for an Aristotelian, exist as aspects of the substance of a subject, and are knowable through their regular manifestation. However, for Hume, operating with a strict empiricist

epistemology, to have a coherent idea of a character trait, there must be a corresponding impression from which it is copied. It is not clear that this is the case; if our idea of traits are simple ideas, then the corresponding simple impression is not obvious. If we conceive of durable traits as complex ideas, derived from numerous and varied impressions, then the question of whether we *should* believe in them remains open, and one defending such a position will have to justify their inference from various impressions of behaviour and feelings to belief in a durable causal mental entity.

Hume's thorough empiricism led him to reject any view that held that there was an essential core to our being, famously writing that:

For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never can catch *myself* at any time without a perception, and never can observe any thing but the perception. (1978, p.252)

In other words Hume could never sense an essential core that grounded his properties; instead he sensed only the properties themselves; impressions, ideas, emotions and other mental features, all of which are originally derived from our impressions of the external world. Lacking any direct primary impression of the essence of a self, Hume could not form an idea of one without breaking his central epistemic principle, hence he defined a 'self' as nothing more than the conjunction of certain mental features; the summation of all an individuals' thoughts and feelings at any one time.

This has led to the theory being referred to as 'bundle theory', as it depicts our self as nothing more than a bundle of different impressions, but

also because it makes a good analogy. A bundle is no more than its constituent parts; there is no bundle over and above its contents and so it is with a self. That Hume's epistemology apparently leaves no room for anything beyond this leads to the problem of character: durable character traits are not included in this bundle, as they are neither an impression nor an idea, yet they play an important role in Hume's wider philosophy.

It is problematic for the Humean that these entities do not seem to be empirically available to us while Humean ethics relies so heavily on them. Standard virtue theories hold that if a person shows habitual bravery, say they are a fire-fighter and regularly and knowingly risk their lives to save others, then they must possess the trait of bravery, meaning that their character can be said to be brave at all times, not just when they are exhibiting bravery. But can the fire-fighter gain an impression of their own bravery when they are at home eating dinner? We can say that they have certain thoughts and feelings at certain times that could amount to bravery, they have memories of being brave, the intent to be brave in the future, knowledge of how to behave under pressure etc., but none of these features have the stability requisite for a true character trait. Nor is it plausible that an observer could have sensory access to a trait like bravery while it lies dormant in their subject.

Although he does not explicitly rely on essentialist metaphysics, Alfano's 'hard core' (2013, p.24) of virtue ethics, a list of features a theory must possess to be considered a virtue theory, highlights the differences between Hume's epistemology and standard contemporary theories. Alfano argues that any virtue ethicist must hold that virtues possess, among other things,

stability, meaning that, *ceteris paribus*, they will endure over time. This is built into the definition of a virtue – if one were to lose the trait arbitrarily after possessing it, it could not have been a true trait. If a person has one spate of generosity while being miserly at all other times, then we cannot claim that they have the virtue of generosity. Their ‘trait’ was a mere flash in the pan, likely motivated by something other than the otherwise concealed generosity at their core and perhaps brought about instead by some trivial feature of their environment or by another, less obviously related trait. Yet Hume’s epistemology leaves one unable to claim that a trait can endure over time, as they are only sensible through fleeting impressions. If there is a single durable entity causing these distinct impressions, it is unknowable through introspection and therefore not part of the bundle self.

Another essential feature of a virtue is consistency; their ability to determine behaviour across contexts. Explanatory and predictive power are also included; virtues will be useful in explaining and predicting an agent’s physical and mental behaviour. In examples to be raised subsequently in this thesis Hume also indicates that he thinks of virtues and vices as having these features. However, it is difficult to see how Hume would understand ‘consistency’ – if ‘consistency’ of generosity is supposed to mean that one repeatedly performs generous actions across different types of situations, then Hume will need some account of what makes disparate actions all fall under the heading of ‘generous’. The problem here is that a variety of disparate actions, with no one common feature, are all taken to be manifestations of the same disposition - a claim which may be difficult to support using Hume’s strict epistemology.

The different approaches of Humean and neo-Aristotelian virtue ethics lead to conflicting conclusions. Alfano can quote Aristotle as saying that ‘the virtuous person is still virtuous when asleep’ (2013, p.26), but Hume did not even believe that the self existed when asleep (1978, p.252). These contrary conclusions arise again from contrary methodologies. Hume’s sceptical empiricism would have left him doubtful that a person could be courageous when asleep, as courage belongs to the self and the self is constituted entirely by perceptions, divided into impressions and ideas (1978, p.1). If we can sense no courage, then we are not courageous. Hume’s epistemology may be too rigid to allow an observer to discover qualities of consistency and stability, making it difficult to distinguish true virtues, or indeed non-moral personality or intellectual traits, from non-durable traits or passing whims, distinctions that are crucial to Hume’s virtue ethics.

Ultimately, it is unclear how Hume moves from his basic epistemological principle to stable character traits that are supposed to exist even when they are insensible. Given that he cannot rely upon substantialism as Aristotle did, it is also unclear how virtues and vices are meant to differ from lesser traits. Did Hume simply adopt his ethical theory without configuring it with his brand of empiricism? Is ‘Humean virtue ethics’ a contradiction in terms? This thesis will argue that it is not, but first to address another key problem in Hume’s account, one that again highlights how Hume’s epistemology problematizes his virtue ethics.

1.2: Hume's basic empiricism

Both of the above problems can be better understood by getting to grips with the foundations of Hume's epistemology: the basic principles of his empiricism and the conceptual tools he uses; impressions and ideas. This section will define these terms to clarify their use throughout this thesis. It will also question whether or not virtues and vices can be understood as equivalent to impressions or ideas, concluding that they cannot, entailing that character traits must be aspects of Humean psychology that exist in part outside of the bundle theory of the self. This will necessitate a more thorough analysis of Hume to address how they are to be understood in a Humean framework.

The most important factor in Hume's epistemology, one which should ground any Humean theory, is that all knowledge comes from experience. Hume writes that 'reason alone cannot give rise to any original idea' (1978, p.207), distancing himself from the rationalism of Descartes as well as classical Aristotelianism. We begin life as a blank slate, unencumbered by innate knowledge or rational principles and gain knowledge by perceiving the world around us. Perception for Hume is comprised of impressions and ideas. Simple impressions are created when a thinking being perceives an external object and serve to represent the object in the subject's mind. Impressions are formed by the senses, and as such are vivid, clear and simple. Ideas are then formed from these impressions as per the copy principle. Compared to an impression of an object an idea will have less vivacity, in that it will appear fainter to us than a sensory perception, and will strike us with less 'force and violence' (1978, p.1). They also become

more malleable, in that one can combine them, alter them and delineate their constituent parts by using one's imagination (1978, p.629). This is in contrast to impressions which must correspond to the world around us. Impressions and ideas are distinguished by their source as well as their vivacity; impressions are formed as a being perceives an object, while ideas are more reflective, and are born of the being examining and sorting their impressions.

Hume's adherence to this copy principle, that 'all our simple ideas in their first appearance are deriv'd from simple impressions, which are correspondent to them, and which they exactly represent' (1978, p.4), means that ideas must be preceded by a corresponding impression in Humean epistemology. If I see a specific lamp, then I form one idea of that one lamp as well as a more general idea of lamps as objects that resemble one another in terms of appearance and function. Perceptions of both kinds are further subdivided into simple and complex (1978, p.2). A simple impression or idea concerns a unity, for example the impression of heat or the idea of red. Complex perceptions concern objects of more than one unity, like the impression of a room, which concerns its walls, floors and adornments, or the idea of the British parliament, which is a compound of complex ideas about individuals, places, processes and so on. How we delineate complex perceptions, which sets of impressions we unite into an idea, shapes how we view the world around us. As Deleuze put it in his reading of Hume, a self is 'not a quality, but the qualification of ideas' (1991, p.64).

Once perceptions have been divided into impressions and ideas, Hume proposes further subdivisions. Impressions are divided into original and

secondary (1978, p.276), or those which are directly concerned with the external world and those which are reflections upon one's own mental state. To exemplify the difference, imagine listening to a favourite piece of music. The sound of the music, as heard by the listener, is one impression; a direct, original impression of what the music sounds like, unfiltered by our mood or how we feel about the music. A further, secondary, impression then represents our joy at hearing the music. This pleasure cannot be called a primary impression, as it represents nothing we have sensed in the outside world, however it is not quite an idea because it is an immediate, vivid feeling to us and it represents only one simple object: our happiness. Below, when further examining how Hume's theory fits with ideas of character, this thesis shall look at how secondary impressions are more intimately linked with what we call character. It appears that, upon a group of people hearing a song, all will have the same original impressions, the sound, but each will experience dissimilar secondary impressions of how they feel about the song; the former can tell us what the agent is doing, the latter can tell us what an agent is like.

Hume summarised how central these epistemological building blocks are in the *Enquiry*:

when we entertain, therefore, any suspicion that a philosophical term is employed without any meaning or idea...we need but enquire, *from what impression is that supposed idea derived?* And if it be impossible to assign any, this will serve to confirm our suspicion. (1975, p.22)

For Hume, whose interest is largely in human nature, this is one of the principles that shape his philosophical work; he is concerned with how

people come to know what they know, and proposes here that ideas can only be derived through impressions of the outside world – any other method will always be open to sceptical doubt. However this position is also the one that will cause this investigation the most trouble going forward, as it appears to prevent Hume from making dispositional claims, which implicitly perform a lot of work throughout his writings.

Can either impressions or ideas of any kind be what we mean when we talk about virtues and vices? If they can, then Hume's virtue ethics will be less problematic, and this thesis less motivated, as they will no longer have to be thought of as dispositional features and can be included as part of the bundle self. It seems that I can have a complex idea of a virtue such as honesty; I can read about it, perceive examples of it, and am able to classify people and their actions as honest or dishonest. Whether or not I can have an impression of honesty is a different matter. A primary impression of honesty, as something sensed in the outside world, does not seem intelligible. I may have impressions of people acting honestly, and I might approve of their actions, but the description of the action as honest appears to come from secondary impressions. I have to reflect, however briefly, upon an action to determine its potential honesty. If we imagine a virtuous person with an opportunity to steal some unguarded money, Hume would describe the agent as having original impressions of the money on the table, the lack of witnesses, and the proximity of the money to their hand. Nowhere do they sense the dishonesty of stealing the money, but on reflection, when secondary impressions are employed, the agent can conclude that such an act would be unfair, dishonest and so on.

Such an evaluation would require an idea of honesty, a complex idea since there is no corresponding simple impression of honesty for a simple idea to be copied from. A complex idea of a character trait could be formed from various impressions of the character trait in others, or from our own mental behaviour – feeling compelled to act honestly often could cause one to form an idea of honesty. Yet this complex idea is not equivalent with a character trait. Hume notes that ideas may become so vivid as to resemble impressions ‘in sleep, in a fever, in madness, or in very violent emotions of the soul’ (1978, p.2), but that aside from these ‘few instances’ ideas feel faint to us. The experience of our own character traits is a vivacious feeling, motivating us to act and intruding into our selves with great force, and so traits cannot be equivalent to ideas. It is unclear what Hume meant by ‘violent emotions of the soul’, but if ideas that are vivid enough to be confused with impressions are ‘few’, then it is unlikely that the exception covered character traits.

Ultimately, it cannot be that virtues can be identified with primary impressions or ideas alone, as they lack some of the necessary features of virtues. Chiefly, they lack the kind of affirmation of the subject that would be required of a virtue. Ideas simply represent objects and concepts, with no evaluative or qualitative element, and original impressions are the same. Secondary impressions are a more promising lead; it could be that when we refer to a person’s honesty we are referring to their secondary impressions being of the kind that allow them to correctly identify honest acts from dishonest acts, while further secondary impressions constitute their fondness for honesty and aversion of dishonesty. The only apparent

problems are that secondary impressions appear to lack the kind of deliberation usually associated with virtue, that is they say they just occur to an individual rather than being something the individual desires or affirms; where virtue 'belongs to the will' (1978 p.9) as Foot says, secondary impressions do not. Additionally, the Humean picture of the self has many more facets, such as the will and the passions, that need to be clearly defined before a final judgement can be made. To that end, this chapter will now analyse the rest of Hume's psychology, keeping in mind that secondary passions are a good candidate for being the kind of psychological feature that could be equivalent with virtues and vices.

1.3: The role of the passions

Hume does not always talk of the self solely in terms of impressions and ideas, but also in terms of passions, which book II of the *Treatise* is devoted to analysing. This section will consider how passions fit with the Humean psychology explained above, querying throughout whether they might be equivalent with virtues and vices. It will be argued that one cannot reduce virtues to the passions; they are instead only a subset of fleeting secondary impressions, distinguished by their fleeting nature and studied as separate areas of interest by both Hume and Aristotle.

Hume introduces new terms for original and secondary impressions in book II: original he terms impressions of sensation, secondary he refers to as reflective impressions. He writes 'Of the second are the passions, and other emotions resembling them' (1978, p.275). Alanen further describes the passions as 'a subset of impressions, differing from sense impressions in

arising not directly from stimulation of sense organs by external objects...but from other impressions or ideas' (2006, p.184). Hume lists 'grief, hope, fear' (1978, p.276) among the passions, exemplifying how personal secondary impressions are compared to primary – they express how a person feels about any given object, rather than represent what they see or hear.

It might then be reasonable to think of virtues as being yet another subdivision of impressions; specifically secondary impressions which relate to a concept and are virtuously aligned in some way. In classical Aristotelian virtue ethics, feelings like grief, hope and fear are the subject of much ethical consideration. Each must be co-ordinated by the 'doctrine of the mean' (2002, p.116); the ideal middle ground between overpowering grief and cold-bloodedness, blind optimism and pessimism, cowardice and recklessness. If an agent achieves this mean they are virtuous, if they err too much to either side they are vicious. It would not be too out of place for 'bravery, kindness, greed' to appear after 'grief, hope, fear', so perhaps Hume's account of virtues is simply a continuation of his account of the passions, only with the focus on the ethical status of such passions.

However there are two reasons for rejecting this equivalence between virtues and secondary impressions. Firstly there is the fact that Hume describes all internal impressions as 'fleeting' (1978, p.195) while virtues and vices are claimed to be 'durable' (1978, p.575). To be regarded as a virtue or a vice, a trait must possess durability; the ability to influence a range of actions across long periods of time. Conversely Hume says of passions that 'on no occasion is it necessary to suppose, that they have

existed and operated, when they are not perceiv'd' (1978, p.195). Precisely how Hume's epistemology permits him knowledge of the existence of traits that are not always efficacious is one of the central problems this thesis will address; for now it is enough to say that Hume conceived of secondary impressions and virtues in significantly different ways.

The second reason is that Hume never bundled virtues and passions together. The fact that he provides plenty of examples for each category while never including cruelty, for example, amongst the passions implies that their separation was not an oversight, but an indication of how differently Hume viewed the two. Aristotle similarly distinguished virtues from feelings, which he listed as 'joy, friendly feeling, hatred, longing, envy, pity' (2002, p.115) etc., which are moderated by the virtuous person. The grounding for this distinction will be analysed further in chapter 3. These aspects are all alike in that they relate to one's character, however the fact that traits are described as 'durable' makes them problematic for the Humean in a way that secondary impressions are not.

In conclusion, none of the building blocks of the Humean psychological framework; ideas, primary impressions or passions, are the answer to the question 'what is a virtue for Hume?' None have the requisite durability, nor the rational deliberation that marks virtues as 'of the will' in a way that impressions are not. Thus character traits must be something more than perceptions, and so must exist within Hume's psychological framework but outside of the bundle self.

1.4: Hume's theory of belief

This section will continue with the account of Hume's epistemology as presented in the *Treatise*, covering what the term 'belief' means for Hume, how a belief is formed, and what is needed for a Humean belief to be legitimate. The aim of this section is to clarify the definition of the term 'justified belief' so that the later arguments for the compatibility of Humean epistemology and real causal powers have a solid foundation. It will be shown that Hume defined belief as a particular force or vivacity, which is justified by the presence of empirical evidence and the lack of constraints such as bias. It will also be highlighted that the source of belief is left unexplained by Hume, in much the same way that virtues and vices are left unexplained.

1.4.1: What is a belief?

What is belief for Hume? Throughout the *Treatise* Hume himself struggles with the question. In the main body of the text a belief is defined as "A LIVELY IDEA RELATED TO OR ASSOCIATED WITH A PRESENT IMPRESSION" (1978, p.96). However the subject is analysed in greater depth in the appendix, where Hume claims that a belief is '*a peculiar feeling...or sentiment*' (1978, p.624), before concluding that it is in fact 'a superior force, or *vivacity*'. Throughout both sections Hume also refers to a belief as 'a particular manner of forming an idea' (1978, p.97). This section will examine the reasons for Hume's confusion in order to clarify what is meant by 'Humean belief'.

Hume gives two accounts of what a belief is in the *Treatise*; one has its own section in part III of book I, then another, more detailed, study in the appendix, intended to clarify his position. He prefaces both by claiming that belief must be either an impression or an idea, most clearly in the appendix: 'Either the belief is some new idea...or it is merely a peculiar *feeling* or *sentiment*' (1978, p.623). His primary argument in favour of belief being an idea is that 'there is no distinct or separate impression' (1978, p.625) annexed to primary impressions that we believe in; no impression of 'reality' or 'existence' that we are aware of when we think of a real castle compared to an imaginary castle. If belief is an impression, it must be a secondary impression, since belief is not an external object to be detected. However, Hume argues, it does not strike us or move us in the same way that other secondary impressions, such as grief or joy, do.

Given that Hume framed the debate as an either/or, one might think the question settled, but Hume has lingering doubts about whether belief really is an idea. (1978 p.629). His primary objection is that the mind has control over its ideas, and yet we cannot willingly apply the idea of belief to a concept we do not believe in. This conflicts with Hume's view that the 'imagination has the command over all its ideas, and can join, and mix and vary them in all the ways possible' (1978, p.629). One who has an idea of a horse, and an idea of pinkness, can combine the two to gain an idea of a pink horse, even without first gaining the impression of a pink horse. If belief is an idea, it is singularly resistant to an agent rearranging it in this way. One cannot convince oneself to believe in a pink horse by taking the idea of belief from an object one does believe in and combining it with the

idea of a pink horse. This objection invites parallels between belief and emotional reactions, which are also beyond an agents control, motivating the claim that belief belongs to the realm of the passions.

These doubts lead Hume to conclude that belief is ‘a superior *force*, or *vivacity*’ (1978, p.629); a quality of other ideas like the qualities of strength and immediacy that distinguish ideas from impressions. So although ideas that we believe in do feel differently to ideas we do not believe in, it is only by degrees – there is no distinct impression of belief annexed to one but not the other, the one is merely more immediate or vivacious.

1.4.2: How is a belief formed?

An idea will then possess the vivacious quality associated with belief when we believe the idea to be true, typically because we have a primary impression of the idea to convince us of its reality. For example if I focus on the idea of my own front door the idea will be accompanied by belief, because I have plenty of experience of it, even if it is outside my experiential range right now. ‘A wise man, therefore, proportions his belief to the evidence’ (1975, p.110). Ten instances of what looks like a causal chain might incline one to believe in it, a thousand should make the idea of that chain a forceful one in any mind. However the veracity of beliefs formed in this way will always be open to sceptical doubt. This section will question the ways in which Hume believed a belief could be considered justified.

Experience is the root of all knowledge for Hume, but experience has taught us that experience is fallible. Thus when I gain a new belief – when I believe I have seen my friend across the street – I must examine the process

that led me to that belief. I find that my eyes may have deceived me, and that I cannot be as certain as I would ideally like. However that judgement is open to the same kind of sceptical doubt, as will any judgements about that judgement, and so on ad infinitum. Eventually I just need to throw up my hands and accept the experience as the truth, never mind the worry that I could be wrong. Reason might show us all the available options, but passions do the actual deciding, motivating us to confer the quality of belief on certain ideas. In cases where there is a great deal of evidence both for and against a proposition, we feel strongly conflicted, as we vivaciously believe in two contradictory ideas. How are we to settle these disputes and accept one possibility over another? In the *Treatise*, Hume argues that reason does not determine which side we settle on. Instead reason is a tool we use to assess different propositions and the relations between them, with our passions ultimately determining our choices when presented with contradictory evidence – ‘*belief is more properly an act of the sensitive, than of the cogitative part of our natures*’ (1978, p.183).

This emphasis on sentiment is somewhat downplayed in the *Enquiry* compared to the *Treatise*. In the *Enquiry* Hume does not treat belief as a subject on its own; perhaps because of his dissatisfaction with his own conclusions as expressed in the appendix to the *Treatise*. Instead his views on belief must be gleaned from his writings on other topics, such as section X ‘of miracles’ (1975, p.109). Where the *Treatise* focused on the nature of belief, the *Enquiry* sheds light on how belief comes about; this is not to say that the two volumes contradict each other, but the shift in emphasis is worth noting. The *Enquiry* considers, when one is in doubt about a

conclusion, which has ‘the greater number of experiments’ (1975, p.111), and ‘the superior evidence’; language which has the connotations of objectivity and appears to leave the agent with less agency when forming a belief.

However the agent’s power to decide remains crucial in determining which evidence is superior; the matter is not conclusively decided by the objective nature of the evidence. When Hume talks of probability in the *Enquiry* (1975, p.57), he does so in much the same terms as he does in the *Treatise* (1978, p.73). The probability of an events occurring, it is made clear in both volumes, is not an objective feature of reality, but is dependent upon custom and our own individual experience with like events in the past. Experience of past events is still regarded as ‘*proof*’ (1975, p.110) of future events, enlivening the ideas of each agent to greater degrees in greater amounts. The attention to the subject as a believer is still there in the *Enquiry*, underlying what it means for an outcome to be probable and what counts as sufficient evidence, though Hume keeps much of this concealed to keep his discussion simpler, with the result that he speaks of evidence in more objective terms.

The kind of beliefs presented above concern probabilities, which Hume contrasts with knowledge and proofs. For Hume, knowledge means ‘the assurance arising from the comparison of ideas’ (1978, p.124), though he is aware that people use the term loosely to describe propositions that are almost certain, such as that the sun will rise tomorrow and that all people must die. He refers to this kind of truth-claim, ones that are based on cause and effect and yet are ‘entirely free from doubt and uncertainty’, as proofs,

and those that are based on evidence yet still ‘attended with uncertainty’ as probabilities. Hume later claims that all knowledge ‘degenerates into probability’ (1978, p.180), and that even expert mathematicians must admit that they possess only proofs, not knowledge.

The next section will deal with how beliefs are to be justified in a Humean system, and thus how ‘knowledge’ – the term will be used in the colloquial sense Hume mentions, rather than proofs, for clarity – is to be obtained. One final note on the formation of beliefs is that the mechanism producing and providing consistency for beliefs remains mysterious. The mechanism appears to be a feature of the mind, a topic that Hume did not explore directly, and which will be addressed in chapter two.

1.4.3: How is a belief justified?

The above account of belief has the ring of emotivism to it, inevitably leading to the question of how Humeans can recognise a belief as legitimate. People may be gullible, or biased, or stubborn, leading them to passionately endorsing or not endorsing their beliefs regardless of evidence. The only apparent distinction between an illegitimate and a legitimate belief for Hume is the causal process that resulted in the belief. Thorough experimentation and repeated experiential data would ordinarily lead to a justified belief, even if the belief is untrue. A rushed judgement, or a biased interpretation of experiential data, will result in an illegitimate belief, whether true or not.

With this principle Hume ensures that belief is not merely a passionate matter, we need to have the right kind of causes for our passions for them to be legitimate. This neatly mirrors much of the rest of Hume’s philosophy –

his aesthetics and ethics in particular. In Humean epistemology, as in those disciplines, there is objective data that can only be assessed through subjective passions, reined in by a set of rules that govern legitimacy – consistency, freedom from bias, relevant experience, and so on. It would seem to follow that Hume settles the epistemological debate regarding objectivity or subjectivity in the same way he settled the aesthetic debate: taste determines what we believe, but only some of those beliefs will be legitimate. One's belief in miracles will be illegitimate in the same way that one's belief that McGonagall was a great poet will be illegitimate – not because of the constitution of the object of the belief but because the subject's belief forming process did not (presumably) adhere to the rules. As in aesthetics and ethics, this means that there is a range of epistemic standards that a subject could hold to, i.e. any standard that conforms to Hume's restrictions, and a range of perfect observers.

However Hume himself presents a fairly specific, and restrictive, epistemic standard; or at least he presents one for philosophers to adhere to. Perhaps his text should be read as an example of one kind of epistemic standard, the sceptical kind, as one option among many. Or perhaps Hume felt that his arguments justified legitimising the sceptical method above all others. The former is the more likely option, as Hume appears well aware of the fact that nobody could actually *feel* the way about causation the way he *thinks* about it – it would be a psychological impossibility to truly convince ourselves that one fast-moving ball won't cause another to move when it strikes it, or even that it might not (absent interferers); to contradict the 'natural and necessary' (1978, p.266) inclinations of our minds.

Whether we should actively endorse causal beliefs of this kind is an open question, addressed by Morris (2006, p.79), who concludes that it may be the case that, while we as people inevitably form beliefs and live by them for pragmatic purposes, they are not beliefs that one should rationally endorse on reflection. We may come to see them as unjustified and wish to reject them, though the belief will still remain – any person watching one ball roll toward another cannot help but expect a reaction, even if they consider the feeling contrary to reason.

In conclusion, Hume describes a belief as a superior force that attends the ideas we believe in, which for most people will mean the ideas that they have regular sensory experience of. Probabilistic reasoning can justify belief to a greater or lesser degree. This account will be relevant to this thesis for two reasons: firstly it highlights that a dispositional property – the mechanism that produces beliefs and keeps belief consistent across time – plays a key role in Hume’s psychological account without being explained. The existence of features like this plays a large role in chapter two; belief in particular is returned to in section 8.2.6. Secondly it shows that empirical observation was at the heart of Hume’s theory of belief, and it is this commitment to empiricism that makes it a core part of Humean philosophy. If a belief in real causal powers is to be justified, as will be argued for later, this commitment cannot be abandoned. The following section will highlight the ways in which Hume extended his theory of belief beyond this central epistemic tenet.

1.5: General exceptions and induction

The remainder of this chapter will give an overview of the rest of Hume's epistemology, focusing on the areas in which Hume extends his basic empiricist principles to allow for belief in objects that are not immediately sensible. To achieve this Hume's epistemology as presented in the *Treatise* and the *Enquiry* will be analysed, with particular attention paid to the areas in which Hume's central empirical thesis – that all knowledge comes from experience – is challenged, and how Hume overcomes these challenges. Also highlighted will be instances in which Hume apparently violates his own principles without explicitly justifying his reasoning, particularly his reliance on the laws of nature in the *Enquiry*. Possible justifications for these apparent exceptions to Hume's empiricist foundations will be considered. The aim of this chapter is to show the very limits of Humean epistemology, when the basic notions of impressions, ideas and the copy principle are extended, with an eye to showing that belief in real causal powers is not so far out of the scope of Humean epistemology as might initially be suspected.

Hume offers two counter examples to his basic principle; that all ideas are formed from impressions which they represent. However both can be explained away without sacrificing this crucial factor. The first issue Hume raises is how we have ideas of fictional objects. He uses the example of New Jerusalem (1978, p.3), where the streets are paved with gold and the walls beset with rubies; no such a city exists but one can form an idea of it easily enough. This is because one can combine familiar ideas together to form a more complex idea; in this case the ideas of gold, rubies and cities can be combined to form New Jerusalem, using ideas derived directly from

impressions. Any imaginary person, animal or place can be explained in the same manner.

The second exception is the 'missing shade of blue' thought experiment (1978, p.6), in which we are asked to imagine a person who has never before seen the colour blue being presented with a colour chart showing every shade of blue save one. Hume assumes that the person would be able to form an idea of the missing shade, despite the fact that they have formed no impression of it. Hume does not solve this problem himself, he just proposes it, but there are several solutions. One response is to argue that the shade is not as simple an idea as Hume believes; that each individual shade is composed of a shade, going from light to dark, a hue, which is the colour itself, and density. Hume appears to accept this himself; in a separate section he writes that 'A particular shade of colour may acquire a new degree of liveliness or brightness without any other variation' (1978, p.96). This implies that he did think of colours as complex, rather than simple, impressions, and that they could be rearranged by the mind in novel ways just as in the case of New Jerusalem.

A more general problem for Hume is how he justifies inductive inferences; everyday assumptions we make about how the future will play out based on the present. When one billiard ball strikes another, how does Hume justify my belief that the second will ricochet away? In other words, what idea can I possess that justifies me in making causal inferences? For Hume, it is custom which fixes these beliefs in our minds (1975, p.58), rather than an appeal to external factors like the laws of nature. Hume's epistemological claims are less about what is real, or whether the one ball

causes another to move, and more about how we come to believe what is real. So a person will naturally believe that the ball will cause the other to move if they have experienced that regularly, regardless of whether there is in fact any causal link between the two events (1978, p.102). The regularity is enough to establish a proof that can be relied upon when making inferences.

However this kind of belief is not an all or nothing matter. As explained above, belief for Hume is a vivacious quality of ideas, attached in this case to the relation between the two balls, and to attendant ideas such as motion and velocity. Belief comes in degrees, as does the justification for a belief. Hence inductive inferences can have a different epistemological character to immediate impressions. As a belief formed from induction is based entirely on experience (1978, p.80) it will be open to any sceptical charge, such as hallucination or illusion, that a belief formed from sensory experience is, as well as fresh concerns about the quality and quantity of past experience and the fallibility of memory. Our beliefs about the future are therefore weaker than our beliefs about our past experiences, which are weaker than our beliefs about our present experiences.

Added to these sceptical doubts is the charge that induction is circular; if it is then Hume has even fewer tools for knowledge forming and his epistemology is further limited. In short, the problem is that reliance on induction is circular as in order to prove that induction will continue to be a successful method of inference one must point to its past successes, which presupposes that induction as a method can produce justified knowledge. Hume did not set out this problem out as clearly as later writers, but he was

aware of it. His response is that induction is more a matter of custom than reason (1978, p.103), and that belief in induction and inductive reasoning is inevitable given the psychology of human agents. It is this natural inevitability of humans to make inductive inferences, 'so essential to the subsistence of all human creatures' (1975, p.55), that justifies inductive conclusions; this is Hume's 'sceptical solution'.

If this explanation is unsatisfactory then one of Hume's two methods of justifying knowledge will rest on shaky ground. Fortunately P. F. Strawson has put forward an argument to prove that belief in inductive inferences is more justifiable than Hume himself claimed. The argument runs that asking whether induction is justified is equivalent to asking whether the law is legal (1952, p.257); it has no standards other than its own by which to be judged. He writes that 'to call a particular belief reasonable or unreasonable is to apply inductive standards, just as to call a particular argument valid or invalid is to apply deductive standards' (1952, p.249). The point being that if it is reasonable to proportion your beliefs to the evidence, and if the strength of the evidence just is past experience, then you're already relying on induction, while deduction goes beyond being reasonable to being a definite matter of fact. Strawson anticipates that he will be accused of question begging for making the claim that strength of evidence is past experience; he replies that phrases like 'good evidence', 'good grounds', and 'justification', all refer implicitly to the inductive principle (1952, p.256).

Taking the above argument into account, we can move forward with a stronger grounding for inductive inferences than Hume did himself. While he concluded that induction may not be rational, but is an unavoidable

habit of human psychology, Strawson's argument shows induction to be as reasonable as it could be without becoming deduction, which is more than reasonable, being undeniable. This leaves us with direct sensory evidence and induction as justified epistemological methods for Hume, which opens up his epistemic circle considerably. His claim that we should base our knowledge of unknown objects on known (1975, p.117) is only valid via the inductive principle. For example if I encounter a liquid I have no prior experience of, I can predict how it will react when heated – it will disperse into a gas - because I have experience of other liquids reacting in the same way, and so am justified in inducing that this new substance will do the same.

Justifying induction also permits Hume use of eye-witness testimony, which Hume makes use of in the *Enquiry*. In cases of belief without a primary impression – for example Hume's belief that Rome is a real place – we weigh up the evidence for and against the proposition in question. Hume believes in Rome because reliable eye witnesses have had primary impressions of that place (1978, p.108), and in Hume's experience reliable eye witness testimony, when not outweighed by competing considerations, is good enough reason to believe in their claims. Eye-witness testimony, effectively second-hand empiricism, is a particular method of inductive belief formation which plays a large role in his analysis of miracles in the *Enquiry*. Once you have experience of someone as a truthful person, he argues, you will come to believe that they will need a good reason to lie, and so take their statements as truth without competing considerations – your confidence being based on the constant conjunction between their making a claim and

that claim's being true. As well as being experientially sound, belief in eye-witness testimony is also necessary, more so for Hume than for us. As a historian, Hume would have had to have held beliefs about the past, only possible by legitimising eye-witness testimony. Meanwhile any beliefs about the operation of the world would be sorely limited if they were based only upon the people and places that one has experienced directly. Contemporary technologies; photographs, recordings, brain scans, are legitimate in the same way. If they come from a reliable source¹ and without competing consideration they should be viewed as the next best thing to sensory experience, as necessary a source of information to our lives as newspapers would have been to Hume's contemporaries.

Testimony thus expands our epistemic circle, letting us legitimately believe in objects far removed in space and time. The argument of this thesis will now proceed on the basis that Hume has two main grounds for justifying knowledge claims: induction – including testimony – and direct empiricism, justified by the argument put forward by Strawson. The next section will look at one final relevant aspect of Hume's writings; his reliance on the laws of nature in the *Enquiry*.

1.6: The laws of nature

Induction is then a justified method of belief acquisition, granting us justified belief of what has likely happened in the past, what will likely happen in the future, and what is likely in places removed from us spatially.

¹ And as long as the causal chain from the object to the depiction you experience is not thought to have been tampered with by any deceptive forces

Less clear is how one might use induction to move from observable objects to objects that are by their very nature unobservable – causal laws, powers, character traits and the like. This kind of inference is common in everyday life, particularly when dealing with the psychology, or inner nature, of individuals. For example when one's friend is frowning one can reasonably infer that they are unhappy, due to a constant conjunction between their frowns and their unhappy mood, despite having no direct experience of their internal unhappiness. This inference is valid because each person is familiar with how their own unhappiness operates, and so can analogise that other people go through a similar process. The testimony of one's friend is also good reason to accept this conclusion. However in this example one still has the primary impression of their own unhappiness to work with, whereas the laws of nature have no parallel impression for knowledge of them to be grounded upon. Despite this, Hume relies on the laws of nature in the *Enquiry* to argue against the possibility of miracles. How is he permitted to do this?

In the *Treatise* Hume applies his epistemological method to the laws of nature, concluding that we cannot be sure that the necessity of these laws exists in the external world, only that it exists in our minds as a vivid impression of belief (1978, p.165). He reasons that we can never have an impression of the necessity at play when, for example, a heavy object falls to earth after being thrown in the air. Custom alone, the force of all our previous experiences, determines that we believe in the necessity of the objects falling (1978, p.102). Thus Hume is often interpreted as having a 'regularity theory' of the laws of nature; as Galen Strawson puts it 'his point

is that (regular succession) is all we can know or comprehend of causation' (2002, p.252). This position entails that we cannot be sure that an invisible force like gravity is at play when we see objects falling while it is also conceivable that all objects just happen to fall back to earth when airborne. However Hume uses the laws of nature, and the surety of their constancy, as a key part of his argument against miracles in the *Enquiry* (1975, p.109).

In 'Of Miracles' Hume argues against the possibility of miracles, relying on the constancy of the laws of nature. In short, Hume claims that we should not believe eye-witnesses that purport to have seen a miracle, because a miracle is, by definition, the breaking of the laws of nature, and anything is more likely than the laws of nature being broken (1975, p.114). We must therefore always assume that eye-witnesses, are in error or are being untruthful when they report a miracle. This is an odd claim for Hume to make, as it seems to run contrary to his epistemology. In the case of miracles, Hume suggests, one should ignore the senses of an eye-witness, and trust to one's own senses, which tell us that the laws of nature are regular. Hume writes that 'experience...assures us of the laws of nature' (1975, p.127), but how can Hume differentiate between the strange and the miraculous if he has a regularity theory, or if he is not sure that the laws of nature are constant? We can have no direct experience of a law of nature distinct from the objects the law acts upon, nor can we induce their existence by analogy as they are ontologically unique, apparently leaving Hume no method by which to justify this belief.

One solution is simply that Hume changed his mind between publications, or that the philosophy he presents in the *Enquiry* has been

modified to appeal to his audience. However Hume shows implicit support for constant laws in the *Treatise* as well, particularly in his arguing for the copy principle, which holds that all ideas are copied from corresponding impressions by the mind. This principle, which ties together Hume's theory of impressions and ideas, is invaluable to Hume's philosophy. If the principle was to be proved false, much of his philosophy would need reshaping, yet Hume seems sure that it will not be proved false. Writing on the conjunction between simple impressions and ideas, Hume notes that 'Such a constant conjunction, in such an infinite number of instances, can never arise from chance; but clearly proves a dependence' (1978, p.4). This is a startling claim for Hume to make, given his famous reluctance to proclaim the same for the laws of nature or causal powers.

Hume's belief in the stability of the copy principle looks out of place when compared to his disbelief in other unobservables such as causal powers. The laws of association, which determine how and why we associate certain ideas in our minds, operate in much the same way. Belief in both of these facets of Hume's philosophy can only be justified through induction, as both extrapolate from occasional sensory data to the existence of an underlying causal factor. It is induction that justifies Hume in prioritising the laws of nature above the senses of eye-witnesses, as there is greater empirical evidence for the regularity of nature than there is for the honesty of eye-witnesses, giving the inductive principle more weight.

All of this goes to show that Hume relied heavily on induction as well as direct empiricism, and that this is necessary for gaining knowledge outside of one's limited sensory experiences. At its core Humeanism should be about

empiricism, so should always adhere to that doctrine, while induction – in the sense detailed here – is an invaluable addition to any Humean theory. The following chapters will argue that similar inductive methods of reasoning can be used when discussing our knowledge of causal powers, which can solve many of the problems with dispositions Hume appears to have.

1.7: Conclusion

This chapter had two main aims: to show the tension between Humean metaphysics and ethics, and to examine Humean epistemology for the root cause, and possible solutions, of that tension. The tension has been shown by recounting the bundle theory of the self alongside the Humean conception of character traits, with particular attention drawn to the fact that traits are not merely impressions or ideas and so must exist outside of the bundle yet still in the Humean psychological framework.

To fulfil the latter aim this chapter has recounted the fundamental aspects of Humean epistemology; impressions, ideas, passions and the copy principle. The nature of belief and justifications for belief have been analysed, along with exceptions to the general epistemic rules to show the limits of this epistemological framework. This chapter has also shown that Humean epistemology is not entirely confined to impressions and ideas, but that it relies heavily on inference to justify its claims. Hume's non-committal scepticism on many topics allows us to use these justifications in novel ways, when supported by new empirical data.

However, none of this has satisfactorily answered the question: how does Hume's epistemology allow him access to stable character traits, which play a significant role in his wider philosophy? The problem remains for anyone committed to the basic Humean principles of empiricism and virtue ethics. The next chapter will look at a case that is analogous to dispositions in this regard, the distinction between mind and self.

Chapter 2: The mind-self distinction in Hume

This section will analyse Galen Strawson's argument for the distinction between self and mind in Hume's ontology, with two aims: first, to reiterate the argument itself to justify the use of the distinction in this thesis, and to go slightly further than Strawson in sketching what the particular roles of self and mind could be. Fundamental aspects of Humean philosophy, including the copy principle, will be used as evidence that the mind plays an important role in Humean philosophy.

The second, broader point is to present a reading of Hume that is quite different than that found in much of the contemporary powers debates – Strawson's Hume is a cautious, yet optimistic, sceptic, unwilling to affirm or deny any objects outside the realm of impressions and ideas, as opposed to 'the greater denier of necessary connections' (Lewis, 1986, p.ix), who would actively reject their existence. Hume's particular brand of scepticism will be analysed to make this point.

The conclusion of this chapter, that a mind separate from the bundle of impressions and ideas plays a crucial role in the framework of Humean epistemology, gives the Humean a new basis for grounding truth claims about durable character traits.

2.1: The mind's tacit role in Humean philosophy

Proving that Hume did not openly reject the idea of a mind beyond the self is a fairly easy task; more difficult is proving that Hume did all but openly affirm his belief in such an idea. Analysing quotes from Hume and working with Strawson's reading, this section will show that an investigation into the

mind is not as anti-Humean a notion as one might assume, by highlighting aspects of Hume's writings; the copy principle, the association of ideas, memory and imagination, which appear to rely on the existence of a mind outside of the bundle self.

Hume's analogy of the mind being a theatre in which perceptions appear to play their parts is a telling one (1978, p.253). Commentators tend to focus on the perceptions, the self, which was of course the topic at hand, but neglect to comment on the theatre around the players, a vast, shadowy construct with a backstage, busy crewmembers and directors, without which the actors on stage could not function. These backstage workings, mostly mysterious in Hume's time, have been much illuminated by advances in neuroscience and psychology, and indeed have to be at least partly understood to understand aspects of Hume's theory.

For example one of the most important aspects of Hume's writings, and one that is unintelligible without a theory of mind beyond the bundle, is the copy principle. The principle states that '*all our simple ideas in their first appearance are deriv'd from simple impressions, which are correspondent to them, and which they exactly represent*' (1978, p.4). The principle occupies an important place in Hume's philosophy; without it Hume would need a different account of the origin of ideas, and of how we categorise our impressions. As Galen Strawson points out (1989, p.130), the mechanism behind the copy principle is unavailable to introspection, operates entirely unconsciously, does not exist outside of our heads, and yet does not exist within the self. However, as detailed above, Hume described and relied upon the principle, reasoning that the empirically sensible effects of the

unconscious mechanism – the constant conjunction of impressions and their temporal priority over ideas – were regular enough to convince any sceptic (1978, p.5).

The mental operations of association and combination of ideas work in similarly mysterious fashion while playing an important role in Hume's epistemology. The principles of association dictate how one comes to group ideas together, which is essential to Hume's story of causation; we only believe that one object is the cause of the other when we experience the constant conjunction of the two in sequence (1975, p.70). Our minds must form an association between the ideas of fire and heat for us to believe in the causal connection between the two. Combination meanwhile solves the problem of 'New Jerusalem' for Hume (1978, p.3). We can have ideas of objects of which we have no impression, such as the fabled city, by combining existing ideas in novel ways. Like the mechanism behind the copy principle, both facilities are innate and present in every working mind, yet their operation may differ from person to person, or with time. Neither ability is available to introspection, yet the principles of association in particular must grip and shape our selves almost constantly. These difficulties are analogous to the difficulties Hume has with character traits, in that they all appear to rely on long term dispositions to both explain and predict phenomena, yet it is unclear from which impression these ideas are supposed to derive, and which place they are supposed to occupy in Hume's moral psychology.

Hume also discusses memory and imagination in his work, claiming that the sole distinction between the two is the vivacity of ideas that they provide

(1978, p.9); there is also the fact that imagination is free to rearrange its ideas while memory must repeat them exactly, but Hume claims that this difference is no help in distinguishing the two as ‘this difference is not sufficient to distinguish them in their operation’ (1978, p.85). Hume treats ideas of memory and imagination as different in only these regards, implying that they arise from the same source (1978, p.628).

To illustrate their similarity, Hume uses the example of two people discussing a past event which both perceived, but that which only one remembers (1978, p.627). The less forgetful one may tell the story of the past event to the other, who will picture the story as though they are imagining it, until a certain detail catches their memory and they realise that their mental images accurately represent their own past experiences. Hume concludes that ‘those faculties (memory and imagination) are only distinguish’d by the different *feeling* of the ideas they present...the ideas of memory are more strong and lively than those of the fancy’ (1978, p.628). This different feeling – which is for Hume the only difference between ideas of memory and imagination – must be belief, which, as recounted in section 1.4.1, is the quality of strength or vivacity that is attached to certain ideas. If this difference is the only perceivable difference, then memory and imagination look like a single faculty for Hume; a faculty that allows one to summon ideas which one may or may not believe in. The above example – which is similar to cases of cryptomnesia, when a subject believes a memory is a new and original idea (Gingerich and Sullivan, 2013) – demonstrate that whether ideas are of the memory or the imagination is not empirically

obvious, meaning that there is no empirical basis for thinking that memory and imagination are separate faculties.

The tacit existence of mechanisms that allow agents to order their ideas, and to create new ideas from new impressions, and their important role in Hume's moral psychology, proves that there is more to the picture than a bundle self comprised of impressions and ideas.

2.2: Hume's scepticism

This section will examine why Hume did not comment further on the subject of the mind, despite the fact that it plays an important role in his philosophy, as outlined in the previous section. The cause is identified as his scepticism, another fundamental aspect of Humean epistemology. This brand of scepticism will be analysed to show that it is not so limiting as might be assumed, with the wider point being that it is not so restrictive as to immediately and definitively rule out the existence of real causal powers.

The chief difficulty for a Humean trying to support the distinction between mind and self will inevitably be Hume's strict empiricist methodology, which rendered him unable to affirm belief in an idea without sensory data or irreproachable logic in support of it. Thus Hume, and other sceptics, firmly believed in very little, a position far from the one held by caricatures of sceptics, who declare much to be untrue or false. Instead Hume in particular left many concepts in an unknowable grey area; he seems content to utilise them in his writings but does not actively affirm that their referents exist. The mechanism that underwrites the copy principle is a prime example of this. The belief that the copy principle will

continue to hold – that each future idea will need to be grounded in a corresponding impression – Hume is justified in holding, as there are no exceptions to the rule while there is a wealth of evidence for it, enough to count as a ‘proof’; which is to say that the belief has enough evidence in support of it to approach certainty without yet being irrefutable (see section 1.4.2 for Hume’s definition of a ‘proof’). However, as we have no impression of the concealed mechanism underwriting this principle, that forms ideas from their corresponding impressions, its nature must remain mysterious. As such Hume does not discuss or openly believe in the existence of the mechanism, while treating its effects as the cornerstone of his epistemology. Strawson describes the notion that Hume did not think about this issue as ‘ludicrous’ (1989, p.130), claiming that Hume must have believed that there was more to the mind than occurrent perceptions even if he did not reference them directly.

To quote the epigraph, taken from Hume, of Strawson’s *Secret Connexion* ‘nature...conceals from us those powers and principles on which the influence of these objects entirely depends. [W]e are ignorant of these powers and forces’ (1989). This perfectly captures Hume’s position on causal powers, and fits with how he may have viewed the mind. ‘Ignorance’ implies unfamiliarity with a real thing, while ‘conceals’ plainly suggests that there is something to be concealed. Strawson later gives a list of Hume’s stance on other issues (1989, p.128); reason is ‘unintelligible’, imagination ‘inexplicable’, and we can have ‘no conception’ of the senses of other animals. This demonstrates not only the breadth of meaningful, important concepts that Hume would not engage with, but also how frequently these

concepts relate to the mind. In each case, Hume refuses to actively affirm an idea because there is no corresponding impression of it, though he treats several of them as valid ideas throughout his work.

Hume later makes his stance on the mind more explicit when he writes 'had we any idea of the substance of our minds, we must also have an impression of it; which is very difficult, if not impossible, to be conceiv'd' (1978, p.232). This reiterates Hume's commitment to his basic epistemology, that all ideas must be preceded by a corresponding impression, which prohibited Hume from forming an idea of a substantial self. That Hume regarded the mind as separate to the self is evidenced when he writes: 'to me it seems evident, that the essence of the mind [is] equally unknown to us as that of external bodies.' (1978, p.xvii) Given that bundle theory was an effort to making the self known, or rather was a clarification of topics all people already know intimately, the essence of the mind in the quote must refer to something else. The mind, reason and imagination at least are all exiled from the realm of the self, which, one could reasonably postulate, actually plays a fairly trivial role in the grand scheme of mental features, being no more than introspectable data; the tip of the iceberg. Hume's relative silence on the range of functions performed by the mind is, Strawson suggests (1989, p.129), not due to a lack of belief in the concept, but to intellectual modesty and commitment to his own empiricist principles.

2.3: The role of self and mind

This thesis has so far sought a way that the Humean, with their restrictive epistemology, can ground the idea of durable character traits that play a

large role in Humean virtue ethics. Having argued that Hume's philosophy depends on an implicit distinction between the self and the mind, this section will now turn to how the relationship between the two is to be understood, and what roles each is to play, paying particular attention to the role of construal.

Hume was clear in what he considered the self to be, a bundle of perceptions and nothing more. He was also clear that 'the faculties of the mind are supposed to be naturally alike in every individual' (1975, p.80), unlike the constituents of the self, which are almost unique for every individual. We could therefore move all other 'neutral' mental features he mentions; reason, imagination, memory, the mechanism underwriting the copy principle, into the domain of the mind. The self would then be left with such a diminished role that, contrary to the attention it receives in Hume's work, it would become the subject less worthy of philosophical inquiry while the mind requires more. If Hume was correct about the self then there is no need to discuss it further; the self is the passive sum total of all our perceptions, which are generated, interpreted or deliberated over within the mind.

The mind has another crucial function that will be of particular importance to this thesis, but which received little specific philosophical attention in the time of Hume; the role the mind plays in construing the world around us. To construe something, in a Humean ontology, means to experience secondary impressions about it; to like, dislike, fear, idolise the object and so on. This subject and its relation to character will be analysed in more depth in section 4.6, suffice to say for now that the source of

secondary impressions, for Hume, must be in the mind. The source is not in the self, as we cannot introspect on the source of our secondary impressions, just the impressions themselves – if it were otherwise, we could simply introspect on our own character or dispositions and these tensions would not arise for Hume. Nor is it in the external world, since secondary impressions do not correlate to external data. Yet nothing else could explain the stability of secondary impressions for each person across time.

Additionally, Hume implies that the origin of secondary impressions are of more interest to the philosopher than the origin of primary impressions. He writes that primary arise ‘from the constitution of the body, from the animal spirits’ (1975, p.275). If he is using the term in the same way as Descartes, to mean physical processes in the brain, then the source of primary impressions will be of little importance to philosophers. Hume does not say the same of secondary impressions, which instead ‘proceed from some of those original ones, either immediately or by the interposition of its idea’ (1978, p.275). That Hume did not appear to consider secondary impressions to be of the body in the same way he did primary leaves questions open about the metaphysics of Humean psychology.

Thus we delineate primary impressions, which represent the external world and do not rely on our characters, with secondary impressions, which represent how we feel about the sense data, exhibiting our individual characters. Hume writes of ‘gilding and staining all natural objects with the colours, borrowed from internal sentiment’ (1975, p.294), in reference to taste, which ‘gives the sentiment of beauty and deformity, vice and virtue’. This distinguishes it from reason, which ‘conveys the knowledge of truth

and falsehood'. For Hume it is passions that are excited by moral differences, just as they are by aesthetic differences (1978, 457). We can extrapolate from this that our mind construes the world around us; guiding us to notice certain aspects of a situation above others, including noticing distinct moral factors or giving them different weight, as well as shaping the way we form and hold associations between ideas. We are directly aware of the consequences of these functions, in that we are aware of stability across time of how our passions respond to stimuli; the causes of this stability Hume leaves mysterious.

Mischel and Shoda (1995) give a contemporary psychological account of construal, stressing the importance of the concept to resist situationist accounts of behaviour. They argue that personality types can be revealed not only through broad notions of character (who is kind, cruel, greedy, etc.) but through how people process and interpret their situations, and the associations between ideas that are formed as a result. They write that 'a personality type consists of people who share a common organisation of relations among mediating units in the processing of certain situation features' (1995, p.257). They conclude that their research problem is to understand why people act in certain ways in response to 'different external and internal features of situations' (1995, p.258). It is these 'internal features of situations' that I take to be a more precise, refined account of the 'gilding and staining' posited by Hume, one that benefits from a new wealth of empirical psychology. With interpretation playing such a large role in Hume's ethics, given that any moral judgement is a matter of our passionate interpretation of a situation, Humean philosophy would benefit greatly by

addressing questions about construal and the dispositional tendencies that appear to underline it.

This section has addressed the important, if implicit, role that construal plays in Hume's philosophy. The concept will be particularly relevant when defending Hume against the situationist challenge in chapter 4.

2.4: The peculiarity of character traits

As the next chapter will deal with virtue ethics, calling particular attention to Humean conceptions of virtues and vices, this section will make a distinction between character traits and other faculties of the mind – their malleability and diversity across persons – to further display why traits make an especially interesting example of the tension between Hume's epistemology and his ethics.

Character traits should be approached differently to the other features of the mind discussed here; the mechanism behind the copy principle, reason, memory, and the will, because they vary across persons. The others, in contrast, are universal properties and operate almost identically across people according to Hume's philosophy; the mechanism underlying the copy principle forms ideas from corresponding impressions, reason discovers causal relations, and so on. Their operations may differ in particulars – two people might reason to different conclusions from the same evidence, and people may differ in the kind of ideas they derive from impressions, but these differences can be attributed to differences in personality and character, rather than differences in the faculties themselves. The effects of each faculty are sensible to the agent that possess them, who can provide

testimony of their experiences. As detailed in chapter one, Hume regards testimony as a legitimate source of justified belief, and so other people's self-reports of the copy principle, reason and so on count as some justification for belief in these faculties of the mind. Thus they are well within the reach of Hume's epistemology; the wealth of testimony regarding their operations constituting something near a 'proof' in Hume's terminology.

Personality traits, which will be examined in more depth in section 4.5, are similar to character traits in their diversity, in that they differ between people, but dissimilar in that they are more resistant to situational cues and in that they contain no normative or evaluative aspect (Banicki, 2017). Personality traits include sexuality, neuroses and the extent to which one is an extrovert or introvert, and so clearly do not operate with the uniformity of the copy principle. These traits differ from person to person, but are generally much more resistant to outside factors than character traits; they appear to be hardwired in a way that virtues and vices are not. This resistance makes them far more easy to verify from a first- and third-person perspective; they do not fall foul of the kind of situationist challenge presented in different forms by Doris (2002) and Harman (1999) that claims to provide evidence for the ineffectiveness of character traits on behaviour. Thus character traits present a unique epistemological problem for any Humean wishing to include them in their moral psychology as their variation across situations makes finding evidence of them more difficult.

2.5: Conclusion

This chapter has shown that Hume did not reject the notion of a mind as distinct from the self, and that in fact much of his work, including fundamental features like the copy principle and association of ideas, implicitly relies on the division. Establishing this distinction removes some of the tension between Hume's epistemology and his ethics, as long-term stable traits can be more sensibly understood as properties of the mind rather than of the bundle of perceptions that is the self. However, the relative lack of stability that character traits possess make them a more interesting case than the mechanism behind the copy principle, for example, as these factors cause them to contrast more extremely with fundamental Humean epistemology. The reason for Hume's reticence on the subject, his scepticism, has been identified and its import highlighted. Now that one side of the tension, Humean metaphysics of the self, has been explored, the next section will analyse the other side; Humean virtue ethics.

Chapter 3: Hume and Virtue Ethics

This chapter will analyse the specific brand of virtue ethics presented by Hume, and the various moral concepts employed to support this account. The aim is to show that many of these concepts are incongruous with Hume's empiricist epistemology, chiefly due to the fact that they appear to rely on a metaphysics of dispositions that Hume does not provide. It will be argued that virtue concepts that are fundamental to the Humean picture such as taste, sympathy and virtues and vices themselves all appear to be long-term stable dispositions to behave in a certain way, but that this dispositional nature is at odds with basic Humean epistemology, leading to the metaphysical gaps. It will be concluded that there are significant tensions between these two areas of Humean philosophy, and that either the empiricist epistemology or the virtue ethics must be amended to make the entire account coherent. Chapter six will then present one way in which the Humean could go about this.

Hume's intention seems to have been for his ethical account to mirror the rest of his work in terms of tone, style and content; that it be as near a scientific description of natural facts as possible, limiting its use as a normative theory. In fact he even gently mocks Plato and Hobbes for their descriptions of worlds far removed from our own (1978, p.402); Hume knew that people have never behaved like those in Plato's *Republic* and believes that they never will. James Baillie makes a comparison between Hume and his contemporary Newton (2000, p.12), claiming that both were searching for widespread truths on which to base systematic laws; Hume on morals, Newton on physics. Hume's explanatory phenomena would be psychological

rather than physical, but nonetheless he believed that one could derive rules from them in much the same way. Hume also explicitly links his ethical theory to his aesthetical work (1995, p.204), thus in this chapter the essay 'Of the Standard of Taste' will be utilised to make claims about Humean ethics by way of analogy.

3.1: The Virtues

One of the areas in which Hume's epistemology seems most at odds with the rest of his philosophy is his commitment to moral virtues. As argued in sections 1.2 and 1.3, virtues and vices cannot be understood as equivalent to impressions or ideas, which entirely constitute the bundle model of the self, which is itself entailed by the basic principles of Humean epistemology. Character traits should instead be understood as long-term durable dispositions, defined by the pleasure or pain they cause their bearer and their relations and which explain and predict our ethical behaviour. The question of whether these traits are relative to society will be raised, with the conclusion being that they are not relativistic, but fixed. This section will then draw attention to the metaphysical problem of Humean virtues and vices, namely that they are best understood as dispositional, and that dispositional features are difficult to explain through the lens of Humean epistemology. It will be argued that, as it stands, Hume's epistemology leaves one unable to make justified truth claims about the existence of such traits.

3.1.1: What is a Humean virtue?

First some remarks about what it means for a trait to be a virtue according to Hume's account. Hume is fairly quiet on the subject throughout his work; preferring to examine each virtue and their merits in turn rather than virtue in general. However he does state in the *Treatise* that 'the very essence of virtue...is to produce pleasure, and that of vice to give pain' (1978, p.296). Virtue may produce pleasure in others, such as when the courage of a defender keeps others safe at personal risk (1978, p.295), or it will produce pleasure in the bearer, both instrumentally through the beneficial consequences of being virtuous and intrinsically through pride in possessing the virtue itself – virtues 'have a tendency to the good of the person possess'd of them' (1978, p.618). Hume concludes in the *Treatise* that virtues are a 'means to an end' (1978, p.619), with that end being human happiness. In the *Enquiry* Hume concludes more specifically that 'Personal Merit consists altogether in the possession of mental qualities, *useful* or *agreeable* to the *person himself* or to *others*' (1975, p.268). So a virtue for Hume will achieve at least one, and likely more, of four goals: utility for the bearer, utility for others, pleasure for the bearer or pleasure for others. They consist in mental qualities, rather than patterns of behaviour.

Hume implies that these mental qualities must have longevity if they are to be considered a virtue in a person. He writes that we are not to be praised or blamed for actions that did not emanate from durable features of our characters (1978, p.411), and since virtues and vices are intrinsically praise- and blame-laden (1995, p.205) they must be long-lasting and more than fleeting thoughts or desires, for which we are not to be blamed. Aristotle

made a similar claim, writing that ‘a single swallow does not make spring’ (2002, p.102), just as a single day of happiness would not make a person ‘truly’ happy, or flourishing. Modern virtue ethicists apply the same principle to virtues, with one virtuous action not being enough to constitute a person’s virtue, which is in fact constituted by a durable disposition towards virtuous action. Alfano utilises this criteria, as he includes ‘stability’ in his hard core of virtue ethics, where the stability requirement is defined thus: ‘If someone possesses a virtue at time t , then *ceteris paribus* she will possess that virtue at a later time, t' ’ (2013, p.24). So not only is a person not courageous if they show courage only once, but that courage was not even virtuous when they did display it, instead it was likely a result of certain situational factors that just happened to be present at the time. What is relevant is that a person is not flourishing after one happy day, nor is a person virtuous after one virtuous moment. A virtue for Hume must then be a durable mental quality that achieves at least one of the four goals listed above.

For a trait to be a virtue it must belong to our character; that is it must endure over time, even if it need not be exemplified regularly. In this sense possessing a virtue is analogous to possessing a skill; a person could conceivably pick up a dart, close their eyes, spin around, and throw the dart into a bull’s eye a hundred times in a row without proving themselves skilled at darts. Each instance would be a mere fluke, and if we were to explain how the player achieved their feat we would not appeal to their training, their dedication, their poise, their knowledge of aerodynamics, or to any consideration that would constitute their having a skill. An act of virtue

must possess similar features, full understanding of the situation, intentionality, and durability.

Any modern Humean account of ethics will need to retain these three characteristics of a virtue: they produce pleasure, they possess longevity and they must be part of our moral character. The rejection of any one of those characteristics will have to significantly reconceptualise the idea of virtue, to the extent that the resulting picture could hardly be called Humean at all. However the latter two characteristics are difficult to square with Hume's epistemology, creating a problem for any modern Humean.

Worth noting is Hume's distinction between natural and artificial virtues (1978, p.477). Natural virtues are approved of by all people at all times, while artificial virtues rest upon 'artifice or contrivance', such as the way some aspects of justice rest upon the particular laws of one country – that one should drive on the left side of the road, for example. The artificial can be understood without a metaphysics of dispositions; on Hume's account to be 'just' is simply to conform with the laws of a given society, such as repaying your debts (1978, p.479), with no regard for the motivation producing the just action. We can honestly say that a person is just, no matter how malevolent their motivation, as long as they abide by the law. Since Hume, and this thesis, are primarily concerned with durable motivating states, only the natural virtues will be referenced from this point on.

3.1.2: The nature of Humean virtues

This section will further examine Hume's account of the natural virtues, paying particular attention to the question of what makes a trait a virtue, to elucidate the implicit metaphysics in his ethical theory. The question of whether virtues are relative to social context will be addressed, with the conclusion being that Hume did not view virtues as relative to social context. Thus any modern Humean is adopting an ethical theory with a fixed list of virtues, meaning that this thesis has a fixed list of explananda.

On Hume's account, the principal conditions that elevate some traits to virtues is the pleasure the traits cause, both to the bearer and to their society. This is a break from the Aristotelian tradition, which emphasised solely the flourishing that virtues enable the possessor of virtues to achieve, rather than the approval garnered from others. Cohon (2006, p.257) points out that Hume, contrary to the Aristotelians, focused primarily on the question of how much pleasure a trait brought to others, when questioning whether that trait could be considered a virtue. She states that Hume believed the approval a virtue garners from others will be 'near-universal' (2006, p.256); giving a solid empirical basis for determining which traits are virtues and which are not. Whether the virtue actually brings its possessor any happiness, or contributes to their 'flourishing', is not a universal fact for Hume, making that criterion the weaker of the two. By depending on social approval to determine the value of a trait, Hume comes very close to holding a relativistic thesis. If this was all that there was to determining a trait the thesis would be relativistic; a trait such as honesty would be virtuous only as long as the bearer's society approved of it. If the society were to

collectively approve of deceitfulness instead, that trait would become the virtue, as could any that garnered widespread appeal.

However Hume staves off charges of relativism against his account of justice by making the claim that the sense of the morality of justice, a sense which is a natural tendency despite justice being an artificial virtue, 'extends to all times and places' (1978, p.620). With this claim Hume displays an unusual certainty about people's moral sensibilities. He writes that justice is 'stedfast and immutable; 'at least, as immutable as human nature' (1978, p.620). He writes in the same section that virtues in general 'are universally attended with esteem and approbation'.

This is how Hume seeks to stave off charges of relativism; he argues that people will always value the benefits that true virtue brings and so will always recognise the goodness of virtue. The fact that Hume believed the virtues to be universal shows that Hume believed that at least these aspects of his moral account were not relative to context. By making the claim that morals are inherent in the soul, Hume suggests that there will be an account of virtue that holds true regardless of social context and so applies across all times and places. The opposite view, relativism, would hold that traits are not objectively or universally virtuous or vicious, but that these factors would be determined by context.

Thus the virtues are a fixed set of psychological traits for Hume. When analysing the concept of virtue in Hume modern Humeans do not have to worry about changing social norms, or that the concept of virtue might shift to cover a class of traits very different in nature to those Hume discussed. Hume's virtues are united by common features, necessitating that we

understood the metaphysics underlying them rather than the social context surrounding them.

3.1.3: The metaphysical problem

A major difficulty for a Humean virtue ethicist will be how virtues exist over time, even when empirically unavailable. The very term 'person' implies longevity, continuous existence and a solid notion of personal identity, none of which belong to Hume's notion of the self, which encompasses all of the mental aspects of a person that are available to Hume due to his strict empiricist methodology. Alfano can quote Aristotle as saying that 'the virtuous person is still virtuous when asleep' (2013, p.26), but Hume did not even believe that the self existed when asleep (1978, p.252). These contrary conclusions arise again from contrary methodologies. If we think of natural virtues purely in terms of the self, then Hume's system faces problems with truth claims for character traits; if we can feel no courage, then we have no grounds on which to claim to be courageous, but courage is felt only fleetingly when the virtue is supposed to be durable.

Historically, in virtue theory from Aristotle onwards, substantialism, the theory that the self is one constant substance which bears one's character traits across a life time (Warrington, 1956, p.167), would have supported claims about durable traits, but this theory is unavailable to the Humean. Rejecting that there is an essence to the self, as Hume does after failing to introspect upon a direct impression of such an essence, entails rejecting that there is one constant substance to do this metaphysical work, leaving

the mystery, addressed in chapter 2, of how we are to conceive of character traits on a Humean model.

One explanation, that allows one to account for virtues without ascribing them to the mind, is that the Humean can hold a summative view about virtues. One could argue that describing a person as 'generous' is simply to reference the fact that the person has often acted generously in the past. The ascription has no explanatory or predictive power, it is purely descriptive and retrospective. It has the same force that calling an unloaded die 'six-prone' would have if it the dice happened to land six up numerous times in a row. Hume does not espouse this view himself, so we cannot attribute to him, but one following his ethics might reason this way to fill the metaphysical gap. It has the advantage of looking congruous with the rest of his philosophy; in the same way that there is only a constant conjunction between fire and heat in our experience, with any actual causal power hidden from us, so is there only a constant conjunction between a miser and their greed. It is not that an underlying causal factor is being rejected, only that our experience is limited in such a way that we can never know if one exists.

Another explanation, which will be argued for in this thesis, is that virtues in Hume's ethics should be understood as dispositions of the mind, rather than the self. That there is a theory of the mind implicit in Hume's work was argued for in the previous chapter; if it is a part of Hume's moral psychology then it is a far more likely candidate for bearing long-term dispositions than the fluctuating bundle self is. The justification for character traits would then not be direct empiricism – one cannot introspect

on their mind, and so could not introspect their traits directly. Instead character traits would be justified through induction, in much the same way that Hume appears to justify belief in the laws of nature (1975, p.109). Just as constant patterns of natural phenomena serve as the evidence of the laws of nature, so can constant patterns of thought serve as the evidence for character traits. Adopting this view fleshes out the metaphysics of Hume's position, and in doing so allows for talk of virtues and vices as genuine causal properties with explanatory and predicative power, which the summative account prohibits.

3.2: The role of the passions in virtue ethics

It is here worth looking again at the passions, which share many of the same properties as character traits and so present several parallel problems. It will be argued that the two are not synonymous, but that passions arise in the self as impressions of our own virtues and vices, which are of the mind, in much the same way that hunger is an impression of some other internal physical state.

To recap Hume's description of the contents of the self:

Original impressions or impressions of sensation are such as without any antecedent perception arise in the soul, from the constitution of the body, from the animal spirits, or from the application of objects to the external organs. Secondary, or reflective impressions are such as proceed from some of these original ones, either immediately or by the interposition of its idea. Of the first kind are all the impressions of the

senses, and all bodily pains and pleasures: of the second are the passions, and other emotions resembling them. (1978, p.275)

Secondary impressions are then divided into calm and violent, with passions belonging to the latter kind. Hume gives no indication either way as to whether all violent secondary impressions are passions or whether passions are a subset of them, but since it is difficult to think of a strong feeling about an idea that could not come under the heading of one of the passions, it seems safe to say that they are one and the same.

Hume confesses that this distinction may look a little rough around the edges; noting that the calm passions are concerned with 'beauty and deformity in action, composition, and external objects' (1978, p.276), and as such are rather calculating and methodical, while the violent passions are broadly equivalent to emotions, and constitute such extremes as ecstasy, misery and fury. The distinction blurs because beautiful art can move us, while emotions can dwindle to contentment, boredom or irritability. Perhaps Hume makes the distinction murkier than it needs to be; it could be claimed that recognition of beauty is always so unfeeling as to be mathematical, but that this calm passion causes within us the separate violent passions of love and joy. Described in this way, the calm passions cause us to be elated, rather than constitute our elation themselves.

The passions are then subjected to another sub-division, into direct and indirect. Hume does not go into this distinction in much depth, but writes that the former are the immediate result of pleasure and pain, while the latter are concerned with more complex relations between ideas. The direct passions are listed as 'desire, aversion, grief, joy, hope, fear, despair and

security' and the indirect as 'pride, humility, ambition, vanity, love, hatred, envy, pity, malice, generosity' (1978, p.276). Later the appetites of hunger and lust are added to the direct passions (1978, p.439).

Several of these passions – pride, humility, envy - share names we might give to virtues or vices, but they are distinct in Hume's moral psychology because they are strictly defined as occurrent feelings, which lack the durability or link to deep character that virtue theorists would want from a virtue. For Hume, secondary passions come and go as circumstance dictates; he makes no claims about their source or deep nature. They play a theoretical role for Hume in being constituents of character, and can be used to explain and predict behaviour. Being passing sensations rather than durable traits, passions do not require more metaphysical analysis as dispositional features do.

3.3: Moral Taste

This section will examine Hume's concept of moral taste, the faculty that allows one to make accurate moral judgements. One of the key questions here will be whether moral taste is a distinct faculty of the mind or a matter of reason. If it is the former, and is a feature of human minds rather than an appeal to objective guidelines, then taste appears to be, like the virtues, a long-term, stable mental quality possessed by moral agents, which again Hume may not be entitled to due to his epistemology. Hume's essay 'Of the standard of taste' from his *Four Dissertations* will be drawn upon; it is a work that focuses primarily on aesthetic taste, but which also highlights how close an analogy Hume drew between aesthetic and moral judgement. It

will be concluded that Hume's theory paints moral taste as a durable disposition of the mind which requires further investigation.

3.3.1: Hume's theory of taste

Hume gives a full description of how an agent develops moral taste, the starting point for which Baillie describes as a 'steady and general point of view' (2000, p.189), defined by the pursuit of the four goals of virtue and avoidance of the three moral errors. Within these boundaries we can then evaluate character traits, praising them for the pleasure and utility we perceive them to cause and condemning them for the opposite. Our pre-moral responses, the feeling of sympathy striking us and the vivacity with which it does, seem to be a matter of upbringing and circumstance; as with all passions the feeling assails us whether we will it or not with no element of personal control. However we are then free to adjust this sympathy using our faculties of reason and imagination, which act as a kind of artificial extension of our emotional responses. Once these factors are taken into account, one has a valid moral stance and can assess character through traits exhibited by actions.

Errors in both aesthetic and moral judgement arise from the same sources, all of which can be corrected by education, they are: lack of delicacy, lack of good sense, lack of practice, lack of comparison and prejudice (1995, p.228). These five are separate from but overlap with the errors of sympathy, to be discussed in section 3.4. A lack of delicacy can be overcome by imagination; delicacy being the ability to focus in on particular qualities of a situation and explain how they should affect one's judgement.

An indelicate observer might like a picture for the bright colours, or dislike Raskalnikov because 'he is a murderer'. These judgements might be entirely accurate, but without a full account of why these things are good or bad the judge cannot be said to have given a fully informed judgement, and so the judgement is questionable. Good sense is the ability to contextualise; the recognition of an agent's purposes in acting and the end to which they aim, which are of primary importance in determining their character. In aesthetic terms, good sense is knowing that Picasso and Holbein had very different intentions when painting portraits, and to not judge one by the standards of the other. Good sense also reins in prejudice, which is present when an agent judges a character, or piece, on factors unrelated to the subject in hand, such as the artists background or some passing mood of their own. Ignorance is simply unawareness of all the relevant facts, which may lead to error in judgements, as in not knowing that a person had a very good reason to lie to you, or our sympathy, as when we learn of someone's troubled past and re-evaluate their acts in light of that.

Lack of practice and lack of comparison may be more obviously aesthetic errors, but they have their counterparts in ethics as well. The two are closely linked; one who has had more practice in listening to classical music will be better able to compare various pieces, but they are distinct – one could listen exclusively to Wagner and so be free from lack of practice when judging his music, but if they have heard no other composer then their judgement will remain hindered by lack of comparison. Ethically, practice is necessary to familiarise ourselves with certain acts and motives; for example one might assume that all people act for the most virtuous reasons possible

when practice suggests that some may act out of vice. Comparison is needed to properly judge the actions of others in the context of usual human action. One might think it of highest virtue to feed and house a cat, until they realise that this is fairly unremarkable, though kind, behaviour. These errors overlap in myriad ways – ignorance of details will form a lack of delicacy, good sense will guide one to realise one’s own prejudices – but all skew a clear view of the subject.

Moral taste thus plays a similar, though more limited, role to Aristotle’s virtue of practical wisdom. Both are necessary for recognising virtue and vice; Aristotelian conceptions of wisdom focus on the virtue’s ability to bring about flourishing, or the good life, and the ability of agents to recognise that benefit, while Hume presents a similar picture in the *Enquiry Concerning the Principles of Morals*, arguing that any ‘successful knave’ that maliciously faked possessing virtues would be denied ‘the unbought satisfaction of conversation, society, study...above all the peaceful reflection on one’s own conduct’ (1975, p.283).

Here again the disjunction between Hume’s basic empiricism and his ethical claims is noticeable. The primary difference between Humean taste and Aristotelian wisdom is their metaphysical grounding; the Aristotelian picture depicts wisdom as a power of the agent, the power to recognise what it is best to do. As with sympathy and wisdom, Hume has no explanation for why the moral taste of an agent should be so consistent over time. This problem is different in kind to the similar problems raised above about virtues and passions, and it does not seem that the summative account can be employed to deal with it. The claim of the summative account is that to

call someone virtuous is just to summarise past instances of their virtuous behaviour, with no positing of an underlying causal factor to explain that behaviour, but moral taste is more a matter of perception than behaviour. It is counterintuitive to think of perceptions in a summative fashion; for example it does not seem that describing somebody as 'having sight' is just a summary of all the past times that they have exhibited having sight. Instead it seems that sight is a consistent feature of their being, always active while the agent is awake and in need of no stimulus for activation. It is better thought of on a powers model than a dispositional model, a distinction to be explored fully in chapter 6. Whether Hume's epistemology commits Humeans to the summative view about character traits, or whether they can be thought of as durable powers as we think of sight, is one of the key questions of this thesis.

3.3.2: Moral taste as a distinct faculty

This section will argue that moral taste is a distinct faculty in Hume's philosophy, one that is yet another example of a long-term disposition in need of explanation. If moral taste is a durable quality of persons, that explains and predicts their moral judgement, then more questions will need to be asked about how a modern Humean can gain a sensible idea of it, just as Hume's epistemology problematizes ideas of virtue. That taste does operate like this is not however explicitly stated in Hume. Commentators such as Korsgaard (1999) and Pauer-Studer (2009) read Hume's standards of aesthetic and moral taste as external rules, and all moral judgements as just appeals to those rules. If this is the case then people do not need to be

attributed a disposition of taste, as the faculty of reason can perform the required function. Thus the idea that taste is a distinct disposition in need of explanation will need to be argued for.

Evidence that Hume regarded moral taste as a distinct faculty comes from his analysis of the value of virtues (1975, p.169). Here he claims that the virtues are marked by four factors; the utility they have for the bearer and for others, and the pleasure they bring the bearer and others. What this highlights is that it is people's appreciation of virtue that makes a trait virtuous, when Hume makes it clear that reason can never provide a subject with pleasure on its own (1978, p.415). Thus it appears that another faculty must be providing subjects with this special kind of appreciation.

Many read Hume as a subjectivist across his philosophy, due perhaps in part to the quote 'reason is the slave of the passions' (1978, p.415). Brown (2008, p.219) notes that for Hume even second order desires are passionate. One might think it natural to suppose that a smoker has a passion to smoke, yet is held back by their reason, yet Hume attributes this second order desire, or desire not to desire, to passions as well; the smoker simply has conflicting passions, one for the relief smoking brings, the other to preserve their health. Neither wish is any more passionate or reasonable than the other, as long as the smoker is not deluded about the cost-benefit analysis at work. If this is the case then one will only ever avoid Hume's errors in judgements because of some desire or sentiment, such as the wish to be ethical or praise-worthy, rather than out of pure rationality. Following Hume's rules or not, an ethical judge is still simply following their sentiments; Hume simply assumes that all judges will desire consistency, to

achieve 'good taste'. Thus there appears to be more at work in moral judgements than applying general criteria to particular instances, which would be the work of reason alone.

Further subjectivism can be read in the quote 'if there be no relation in life, in which I could not wish to stand to a particular person, his character must be so far allowed to be perfect. If he be as little wanting to himself as to others, his character is entirely perfect' (1978, p.606). Even unanimous adoration of a character does not amount to an objective valuation of them, just a collection of subjective valuations, so here Hume is again relying on subjective responses to determine what is virtuous. *Virtue in rags* (1978, p.584) poses the same problem in a contrary manner; the problem being that while vile people can gain adoration, great people could never be recognised for their qualities.

Abramson (2008, p.252) has a way to solve some of these difficulties raised by writers such as Korsgaard (1999) who try to prove the objectivity of moral taste. Abramson argues that Hume's four marks of virtues are only identifying features of virtue; they do not in themselves determine what a virtue is. The 'ultimate test' of virtue is whether anyone in a character's inner circle is left wanting, a view which seems to ignore the idea discussed above that one can stand in 'any relation' (2008, p.236) to the subject. Specifically, Abramson would have us ask "what difference would this trait make to people involved in interpersonal interactions with someone who has that trait?" (2008, p.254) In the example given, celibacy could never be a virtue as the bearer would always be wanting, meaning lacking rather than desiring, sexual intimacy, while licentiousness also fails the test as the

bearer, or their circle, would want for emotional intimacy. While this approach does narrow the field of possible virtues, it also seems unduly normative in making claims about how much sexual intimacy is the right amount, while leaving open the criticism that pleasurable company is not necessarily ethically sound company.

In the essay 'The Sceptic' Hume seems to disavow objectivity in judgements completely, claiming that 'those qualities (beauty and amiability) are not really in the objects, but belong entirely to the sentiment of that mind which blames or praises' (1854, p.178), strongly implying that aesthetic, and by analogy moral, judgements have any truth-value at all. He made this point to counter rationalism in aesthetics, a school of thought which might conclude that merely knowing the features of an artwork could provide one the information necessary to evaluate it. Evidently he thought the same of ethical judgements, which is why he advocates putting oneself in another's shoes before judging them. Hume never claims that one should forget oneself entirely, or enter into an objective god's eye view of the situation, proof perhaps that he did not see ethical judgements as anything beyond subjective responses to character.

3.3.3: The role of reason in ethical judgements

The primary argument for Hume's account relying on objective rules more than thus far presented is his use of what Blackburn terms the four refining principles, derived from opposing the errors in moral judgement listed above, they are: delicacy, good sense, practice (particularly in making comparisons) and freedom from prejudice are all necessary to make ones

subjective judgement a valid one, in both the ethical and aesthetic spheres (2008, p.100). It could be argued that, given the rigidity of these principles, there is in fact only one correct judgement that every person should make, yet that human fallibility prevents them from making. This is contrary to the subjectivist claim, which would hold that all judgements are equal, or a less rigidly objectivist view, which would hold that a range of judgement might be appropriate. Accepting this view entails that moral judgements are simply a matter of reason; of applying fixed general criteria to particular instances, and thus that there is no special faculty of moral taste.

Hume's treatment of 'virtue in rags' (1978, p.584) lends credence to the objectivist interpretation. Hume suggests that virtuous people that cannot act on their virtue, poor people that cannot donate despite their generosity for example, should be praised nonetheless. This suggests that the descriptive account Hume gave is one that needs to be overcome by those with proper ethical taste, who will live up to certain objective standards to ensure that their subjectivity does not cloud their judgement. Even entering into the close circle of a person whose virtues are frustrated will not help, as their inner circle will receive no pleasure from unfulfilled generosity, except perhaps for a kind of admiration that would not be roused by a poor miserly person. Hume is effectively suggesting that to properly judge virtue in rags we need to abandon our subjective feelings about them and follow rationalistic principles; generosity is generally a good trait so it will be a good trait here, despite the trait failing to cause the appropriate empirical feedback.

Pauer-Studer (2009), following from Korsgaard (1999), believes that Hume's 'general point of view' is a source of objectivity in his ethical study. Blackburn (2008, p.62) appears to agree, lining Hume up with other virtue ethicists for whom the proper alignment of the passions was the chief concern. This is further supported by Hume's claim that whether one has good taste or not is a matter of fact (1987, p.241), implying that there is an objective criteria for judging one's judgements. Hume does claim explicitly that judgements of taste themselves are not matters of fact, but he also claims that some are 'absurd and ridiculous' (1995, p.210); comparing them to a judgement that 'a molehill be as high as *Teneriffe*', an undoubtedly objective claim. The fact remains; one who has good taste will make good judgements, judgements that one will either value or respect if one possesses good taste, or which one will never be able to properly judge due to a lack of taste.

3.3.4: The Mixed View

Korsgaard's main argument that Hume's ethics are based on objective principles is based on the 'general point of view' (1999), the idea that we should all adopt one identical position, ideally within the inner circle of the person being judged, from which we will all, as long as we are properly refined, come to the same objectively correct conclusion about an ethical character. Exactly what Hume meant by 'entering into the inner circle' is unclear. It seems obvious now that we should not simply consider the people that surrounded moral monsters, such as Genghis Khan, to learn how they reacted to the monster and thus draw our conclusions, since

history is full of people that benefitted their nearest and dearest while being clearly unethical. One could argue that the utility Genghis Khan brought to his community is balanced out by the disutility to the people he disserved, but then we can conceive of one of his contemporaries, just as bloodthirsty but more cowardly, who supported everything Khan did in spirit without ever actually harming a soul. Such a person might seem even less virtuous than Khan, despite bringing about pleasure and utility without directly diminishing pleasure or utility.

Neither does Hume want to advocate a god's eye view position from which we can all make the same rationally delivered judgements, as he notes several times that, without the sentiments, the passions that part constitute us as a subject, we could not make moral judgements at all. Thus the general point of view does not advocate any one objective point of view as the best, neither an imagined god's eye view, nor a sympathetic study of other people's opinions. Instead what Hume is asking us to do is to imagine that we ourselves, with all our subjectivity, passions and taste, were closer to the person in question:

Our servant, if diligent and faithful, may excite stronger feelings of love and kindness than *Marcus Brutus*, as represented in history; but we say not upon that account, that the former character is more laudable than the latter. We know, that were we to approach equally near to that renown'd patriot, he wou'd command a much higher degree of affection and admiration. (1978, p.582)

This approach does not lead to one objective, correct final view, but to multiple subjective views that are better informed. No judgement will be

revealed to be untrue, but some may be revealed to be absurd. For example if one believes that Genghis was a monster that never did anyone any good, one need only enter into the position of one of his beneficiaries to see the factual error in that position. The Khan critic might then argue that they would reject any of Khan's spoils of war out of good conscience, but at least they will have gained a better understanding of the world the character had to live in, of the motives that drove him and the society that praised him. Once we have entered into the general point of view we might, and probably will, each come to our own judgements about the character, none of which can ever be wrong, but which can be better or worse informed.

This mixed approach allows for a smoother comparison between aesthetic taste and ethical. Growing up in a 21st century city, physical courage is not a virtue that is often called upon, or that brings much pleasure or utility, thus a modern city-dweller might naturally struggle to relate to Genghis, Roman soldiers or medieval knights. Similarly realism in art fell out of fashion when the camera was invented, the utility of that new device made realistic paintings obsolete, but that is not to pass judgement on the earlier artwork that was created before the camera. An expert judge of Mozart need not be an expert judge of Persian architecture, just as an Elizabethan courtier will not know how to properly judge a Celtic warrior. These are the kind of 'blameless' differences Hume identifies which can only be overcome with his rules for refining taste – 'A man of learning and reflection can make allowances for these peculiarities' (1995, p.234). This approach also re-attaches Hume to virtue ethics as a study of intentions, as well as helping to forge a humane theory of judgement that allows for a

range of upbringings and societies. The judgements being made will then not be of the objective form 'x is a bad person', but of the subjective form 'I feel this way about x', where feelings include praise, blame, understanding, and acceptance.

We should seek only to become *a* refined commentator, not *the* refined commentator as Korsgaard (1999) makes out. It is not that my subjectivity blinds me to the objective reality of ethical facts, only that some of my subjective responses will be irrelevant to the matter at hand. Prejudice is the best example; it can be as difficult to condemn a friend for a horrible deed as it is to dislike their artwork, anyone who has seen a friend act in a play will know that it can be downright impossible to judge their performance because it is impossible to remove the, aesthetically irrelevant, fact that you know them well. These are the wrong kind of passions to have bearing on our ethical judgements, which is not to say that no passions should have an effect.

3.3.5: Conclusion

Hume does have a few normative rules that he regards as objectively necessary in forming the best ethical and aesthetic taste, but this does not lead to there being one correct position to take, or one best judgement of a certain character. A refined taste will not be the best taste of all; rather it will just meet the minimal standards of good taste. Fortunately this minimally good kind of taste is far easier in the ethical realm than the aesthetic, since everyone has some natural knowledge of what is good and bad, while expertise in the myriad fields of art is harder to come by. The

kind of thorough study of art Hume believes to be necessary to achieve aesthetic expertise may not be needed for ethical judgements, which a far wider circle of adults can be expected to pass. To compare Hume's theory of ethical judgements with judgements on colour, two people can disagree about whether a hue is green or blue without contradiction, but one who claims it is black will be roundly ignored.

That moral judgements are not simply a matter of reason, that they are not just appeals to existing sets of rules, entails that a distinct faculty of taste must be at work. This implies that there is a durable feature of the mind which, like character traits, calls for more metaphysical explanation than Hume provides.

3.4: Sympathy

One of the central claims of Humean ethics is that sympathy, one of the secondary impressions or passions, is essential for an agent's sense of morality. This has the ring of subjectivity, perhaps making it appear like a departure from the Aristotelian virtue tradition, though the similarities between Aristotelian wisdom and Humean sympathy are many. Hume's definition of sympathy is similar to modern definitions of empathy; the feeling of being in someone else's shoes so that your emotions imitate the object of your sympathy. As Hume puts it sympathy 'renders all their sentiments intimately present to us' (1978, p.320), referring to other people. Without sympathy an agent would make no moral judgements of their own because they would have no motivation to do so. Such an agent asked to describe a robbery could relate every fact of the crime without calling it

‘wrong’ because the physical facts will never amount to an ethical fact. Like virtues and taste, sympathy appears to be another major feature of Hume’s ethics that rests on a metaphysics of dispositions that Hume appears to reject.

Hume argues that all human agents possess sympathy to some extent, the small minority that do not are clearly malfunctioning in his view. The extent to which one’s sympathy extends is marked by the rational associative principles; contiguity, resemblance and causation (1978, p.318). That these three principles determine the reach of sympathy is as true now as it was in Hume’s day; people still feel more for a family member than they do a stranger, atrocities at home still attract more attention than those abroad, even contiguity of age can make a person feel more for an agent’s plight by making it easier for one to put oneself in their place.

However Hume avoids pure subjectivism by introducing reason as an arbiter to refine our sense of sympathy. Reason acts in the moral arena much as it does in other areas of perception; by giving us a more objectified stance from which to view our situation. Thus we conclude that the distant atrocity is no less atrocious for being distant in the same way that we know faraway objects are not actually miniature, our initial impression is tempered by what we know to be true, not just what we feel. Hume does not fall back into rationalism to escape subjectivism however, as his claim that an agent without sympathy will never begin ethical deliberation in the first place still stands.

These concepts, sympathy and reason, would benefit from a metaphysics of dispositions that Hume does not provide. It seems intuitive to say that

there is one underlying feature of our psyches that causes occurrent feelings of sympathy, and another that inclines us to use our reasoning faculty and to which degree. As in the search for the self, anyone trying to introspect on their own dispositions could discover past examples of their own sympathy and reasoning, and could summon new feelings and deliberations with the imagination. However they could find no impression of the source of these occurrences; a cause to explain these scattered but related effects.

A summative theory of dispositions might be thought to solve these difficulties. It could be held that each pang of sympathy is its own entity, generated once when my self encounters the right circumstance and then gone forever. This position will be argued against more fully in chapter 5; suffice to say for now that an explanation of the consistency of these faculties would be beneficial to Hume's philosophy without compromising any significant element of it. Some people are more sympathetic or comply with reason more than others, and a good psychological theory should at least attempt to explain how such consistency not only possible, but seemingly abundant.

3.5: Conclusion

The aim of the above chapter was to highlight the fact that several fundamental concepts of Hume's ethical theory seem to be long-term dispositions. Virtues and vices, understood as durable character traits that explain and predict behaviour, are dispositions to behave in certain ways, as are the faculties of moral taste, sympathy and reason. Given Hume's restrictive epistemology, which entails that the self is understood as only a

bundle of perceptions, it is unclear where these dispositions are supposed to fit in a Humean psychology. This thesis will now focus on resolving this tension, with two main solutions in mind: either Humean virtues should be conceived as summative dispositions rather than actual stable traits, or they should be conceived of as real, causally efficacious powers, which would require a re-evaluation of Hume's epistemology.

Chapter 4: Character

Chapter one of this thesis laid out the details of Hume's bundle theory of the self, a theory which depicts the self as no more than a collection of impressions and ideas, maintaining its form only fleetingly before the constituents change. However in other sections of the *Treatise* and *Enquiries* Hume evaluates people in terms of their moral character, with little reference to their selves, or an explicit explanation of how the two are related. Moral character is intrinsically linked to the ethical features analysed in the previous chapter; it is constituted by virtues and vices, and finds expression through factors including the passions, taste and sympathy. This chapter will attempt to clarify Hume's position on the subject of character, hoping to provide a more holistic picture of his metaphysics and ethics.

This chapter begins by drawing out what Hume did claim about moral character, proposing that Humean moral character has three main features: it determines behaviour, grounds moral responsibility and relates to the passions. The distinction between the bundle self and character will then be drawn, before theories of character proposed by various other philosophers and commentators including Harman, Doris and Kupperman are analysed in an attempt to discern which sits best with Hume's writings on the topic, in order to flesh out the concept. Baier – who offers her own interpretation of Hume's work – will also be assessed. Drawing on these resources will clarify what Hume's position on character was, more importantly it will clarify modern Humean interpretations of this concept. General considerations

about moral character will be taken into account alongside the restrictions that Humean epistemology places on any Humean theory.

To present a picture of moral character it is useful to begin by demonstrating what it is not, contrasting the notion with the concepts of reputation, skills, personality and intellectual character. From this analysis it will be concluded that character is generally understood to have two primary functions: it affects how we express ourselves publicly and determines how we construe the world around us privately. This will allow us to clarify what Hume meant when he used the term 'general character' (1978, p.418), and will highlight the tensions that arise when one tries to hold both a bundle theory of the self and a theory of durable character. The problem is metaphysical, in that there appears to be nothing for Hume to attribute character traits *to*, though they play a major role in his ethics, but the root cause is epistemological; Hume's sceptical empiricism makes it difficult for him to affirm the existence of long-lasting dispositions, which form the basis of one's character. As this scepticism is fundamental to Humean philosophy, it will inevitably restrict any theory of character that places itself as part of a broader Humean ontology.

This chapter will conclude that Hume conceived of character as a feature of our minds, similar to memory or the mechanism underwriting the copy principle, which influences our selves without ever being immediately available to introspection itself. Our secondary impressions, the way we feel about raw sense data, are determined by our characters, as are what Frankfurt would term second-order desires (1971, pp.5-20), or endorsements of our first-order desires; the manner in which we associate

ideas is also intimately linked with character. Actions reflecting that character follow naturally, though not necessarily. Character is the function of the mind that provides Humean selves with the stability and consistency otherwise lacking in his account. In later chapters it will be argued that introducing a theory of mind and powers theories can diffuse these tensions without sacrificing any crucial point of Hume's philosophy.

4.1: Hume's position on character

Pitson claims that 'Hume fails to provide a systematic account of the concept of character' (2002, p.83). If he is correct then Humean philosophy cannot be complete without one being engineered, one that is coherent with bundle theory, or Hume's ethics will have to be re-evaluated. Given the significance of character to virtue theory, this re-evaluation would have to be fairly major. Since Hume at no point gives a thorough account of his conception of character, this section will outline the three crucial factors that constitute his unwritten theory of character. These are factors that Hume addresses time and again when writing on the role of character in ethics, and should not be excluded or subverted when trying to flesh out the theory.

Hume gives character an important place in book III of the *Treatise*, and comments in several places on his confidence in character as a predictor of behaviour. He writes that 'Almost everyone has a predominant inclination, to which his other affections and desires submit, and which governs him, though, perhaps, with some intervals, through the whole course of his life' (1987, p.160), which sits uncomfortably with the 'perpetual flux' of a

Humean self. At another point he writes that 'No union can be more constant and certain, than that of some actions with some motives and characters' (1978, p.404), showing an almost startlingly un-Humean certainty about cause and effect that calls for examination. This ability of character to explain and predict behaviour, its apparently causal nature, is the first essential element of Hume's theory; the epistemological difficulties in justifying such certainty its first hurdle.

The second crucial role Hume's conception of character has to play concerns grounding moral responsibility. In his writings on necessity, while defending his position from the claim that it eliminates moral responsibility, Hume argues that we are only truly morally responsible for actions that are caused by our character. When an immoral action is not generated by one's character but by 'accidental' causes, Hume states that:

The action itself may be blameable...But the person is not responsible for it; and as it proceeded from nothing in him, that is durable or constant, and leaves nothing of that nature behind it, 'tis impossible he can, upon its account, become the object of punishment or vengeance. (1978, p.411)

The above quote again highlights Hume's commitment to character traits being long-lasting aspects of an agent, of the kind not addressed in his examination of the self, and the important role that character plays in his theory of moral responsibility – character here being distinct from personality, a division that will be elucidated in section 4.5. Since Hume intended his account to be descriptive, avoiding empirically problematic metaphysics, we can infer that Hume believed that people can distinguish

between which actions proceed from character and which do not, and apply blame or praise appropriately. How this distinction is to be made will be an important theme of this chapter.

These are the two roles that any Humean picture of character must play, the third important element is its metaphysical constitution. For Hume, character was intimately linked with the passions, just as for Aristotle character was linked with emotions. Hume lists ‘meekness, beneficence, charity, generosity, clemency, moderation’ (1978, p.578) and more among the virtues, and while these would look out of place in a list of direct passions (grief, joy, fear, any passion that arises as an immediate reaction to primary impressions) they may be thought equivalent to indirect passions (humility, pride, love, any passion that relies on other, associated direct passions). In both Hume and Aristotle proper management of one’s feelings is the foundation of good character; the virtuous person will be inclined to act virtuously when it might be tempting to act viciously. One reading is that our occurrent emotional states compile to make virtues and vices, which in turn constitute our character. This chapter will argue instead that emotions are determined by our character traits, which insensibly underwrite all of the emotional activity in the self.

These then are the three primary features of Hume’s theory of character: firstly character determines our behaviour, and in doing so makes it possible for others to predict our behaviour, apparently with a high degree of certainty. Secondly character is the foundation of an agents’ moral responsibility, as an agent acting from motivations not ‘of character’ cannot be held morally blameworthy. Thirdly character is related to the passions,

which overlap with but are not identical to virtues and vices, and may either determine or be constituted by our emotional states. The readings of Hume proposed in the following sections must therefore meet these three criteria to be considered a valid extension of Hume's bare-bones theory of character, as well as not contradicting other areas of his philosophy in any significant way.

4.2: Characters and selves

This section will show why it is that Hume's notion of character cannot be reduced to his notion of self, nor can it be a part of the self, and will draw attention to the difficulties this raises.

Hume clearly believed in a self as distinct from one's character, and that both were important subjects in philosophy. The self he delved into at length; as explained previously, the self can only be accessed through introspection on your mental states, it is also constituted by those mental states. I could never, by definition, have access to another's self, because that would immediately entail them becoming part of myself, one more perception to add to the bundle, as detailed by Ayer (1954, p.191-214). I can gain impressions of another's behaviour and have the behaviour remain distinct from my self, but the impression it gives rise to will constitute my self, however briefly.

One's character, on the other hand, may be a different story. A definition of character from Kupperman reads 'X's character is X's normal pattern of thought and action, especially with respect to concerns and commitments in matters affecting the happiness of others or of X, and most especially in

relation to moral choices' (1991, p.17). Kupperman's account of character is explicitly set against any 'snapshot' view of character, focusing instead on long-term projects and the persistence of traits over time. It will be questioned whether character is best described in this way later in this chapter, but for now the definition will serve as an approximation of what Hume and other virtue ethicists mean when they use the term 'character'; if elaborated it could meet the three crucial criteria for a Humean theory of character, and certainly does not contradict them.

Kupperman's account clashes with Hume's theory of the self in several ways. Firstly it defines character as something that must be discerned through study of a person's behaviour, while Hume's self can only be known through introspection. One has direct experiences of one's self at all times, while to judge someone's character one must study their behaviour at some length. This means that the epistemological foundations for one's self and one's character are completely different. Nothing in Hume's writings is inimical to this idea; he never makes the claim that character is solely, or even mostly, available through introspection.

Secondly the priority of 'concerns and commitments' which Williams (1981, p.5), also takes to be central to questions of character. These commitments, including professional commitments such as a doctor's commitment to saving lives, can only appear in a Humean self as occurrent mental states, when Kupperman and Williams clearly intend for them to be more than that – in fact Williams uses these factors as the basis for his rejection of the 'successive selves' model advanced by Parfit, which Williams argues does not reflect the true unity of one's character. Both concentrate

on how these aspects of character are long-term; Kupperman points out that a concern must necessarily be understood as a durable feature of the mind to count as a genuine concern at all (1991, p.17). Elements like this are clearly not fully captured by any view of the self that references only occurrent perceptions and require deeper explanation.

Lecaldano supports the distinction between self and character on the grounds that one self may have several characters over a lifetime (2002, p.190); even as an adult one may spend years with a pessimistic character only to undergo a dramatic conversion and adopt an optimistic character. Pitson (2002, p.84) makes the same point to distinguish character from personal identity, with which it should not be confused. All of this indicates that there is a widespread agreement that one's character and one's self are distinct in Hume's ontology.

4.3: Character as behaviour or reputation

There is then a clear distinction to be drawn between one's character and one's self in Hume's philosophy. But does Kupperman's claim that character is just a 'pattern of thought and action' match Hume's conception of character? If it does, then our character will be in part external to the mental realm, and aspects of it could be known through third-person observation.

Baier interprets Hume as believing that character is something like reputation, and that we should follow Hume in focusing entirely on factors external to one's mind or self. This section will argue first against Baier's view that character is equivalent to reputation, then against Kupperman's

view that character is in part constituted by patterns of action, concluding that a Humean conception of character must be constituted entirely by mental factors. Given Hume's distinction between one's self and one's character, the mind must therefore be considered the bearer of character traits. An argument against the second half of Kupperman's definition, that character is partly constituted by patterns of thought, will be presented later in this chapter.

4.3.1: Baier

This section will scrutinise Baier's analysis of Hume from *Death and Character* (2008), supplemented by her earlier work *A Progress of Sentiments* (1991). Two different readings of Hume emerge, with contradictions that Baier seems unaware of. The first reading will here be referred to as the 'reputational approach', and is put forward by Baier when she argues that character is entirely an external matter. The second will be referred to as the 'mixed approach', which Baier puts forward more implicitly when she discusses the import of mental features as though they were of equal weight to non-mental features. The aim here is to show that Hume's theory of character is not concerned with what others think of us or how we act, but how we think, contrary to both of Baier's approaches. The term reputation will be used to mean our outward facing behaviour and what people think of it, without necessarily meaning a famous or widely held reputation.

When presenting the reputational reading of Hume Baier prioritises the outward-facing aspects of character, stating that 'Character is not some hidden inner constitution of a person; it is the outward expressive face of

that inner nature' (2008, p.4). Baier draws much attention to the ways in which Hume treats reputation as character, with it being the intersubjective opinion of all those in my circle and traits needing to be 'given' rather than simply possessed (1991, p.188). Pitson echoes this by calling the self a private thing and character a public (2002, p.91). Baier seems to draw attention to these aspects of character because Hume did; whenever Hume wrote on the character of a historical figure he used the term interchangeably with reputation, but this does not present good grounds for thinking that Hume wished to focus on those aspects in his ethical philosophy. It may simply be that Hume, when writing about long dead kings and queens, did not choose to speculate on their inner states, preferring the concrete facts of what contemporaries thought those states might be. Hence his using the two words interchangeably could be a quirk of his approach to his subject, rather than any metaphysical commitment.

The reputational approach is not coherent with other parts of Hume's writings on the matter, such as when he writes of 'durable principles of the mind, which extend over the whole conduct, and enter into the personal character' (1978, p.575), and in his examination of free will makes a point of how predictable people are because of their mental qualities. He presents examples in which he exhibits near certainty that mental features lead predictably to behaviour, including one in which he imagines leaving a purse of gold at Charing-Cross station (1975, p.91). He claims that it will as likely 'fly away like a feather' as remain untouched, a confidence that is grounded in his belief about the durable nature of people's characters. It is this nature to which Hume devotes most of his ethical philosophy, and

much of his 'science of man'. If this inner constitution does exist, as Hume apparently believed it did, then it certainly plays at least some role in character, making the purely reputational reading untenable.

However other parts of Baier's writings suggest another reading of Hume, one which gives more attention to the significance of our mental features. She comments that a good reputation may be false, while a good character is good regardless of the opinions of others, meaning that the truth maker for statements about character will ultimately rest in one's inner constitution. Baier writes that it is mental aspects that will determine one's actions, as per Hume's theory of necessity, which are observed by the public at large and one's close friends in particular, who can then form reasonably good judgements of one's character traits. She also writes that 'How a person manages her own conflicting passions...how she uses her intelligence to serve her prevailing passions, as well as which passions those are, will be a matter of character' (2008, p.3). Passion management of this kind seems to only contingently impact on one's reputation and behaviour, but acknowledging the link between character and passion does succeed in meeting one of the three criteria for a Humean theory of character where the reputational approach failed. It could be possible that Baier meant that one's character, understood as reputation, is based on how observers believe a subject can manage their own passions, rather than how one actually does so, but her comments about the possibility of reputation being false, while character cannot be false, contradict this. She also claims that character is 'a person's nature, as in principle perceivable by fellow-persons' (2008, p.4); where nature seems to mean some inner constitution, if not a hidden one.

This will be referred to as the mixed theory as, since the distinction between approaches is mine, Baier does not give a thorough account of it. Thus mental factors vary in importance even within this account, though the focal point of this theory of character remains one's reputation.

Baier works with various texts from various times in Hume's philosophical development, many obscure, and so does not present Hume as being coherent or substantive when it comes to questions of character. Of the two approaches detailed here the mixed is a more convincing interpretation of Hume; it is clear that Hume could not have held a rigid reputational theory, from evidence in his own work and from the obvious objections to such a position. However it is also clear that Hume did not hold a mixed approach; instead it will be argued below that his theory treats character as an exclusively mental phenomenon with behaviour only contingently related.

4.3.2: Kupperman and patterns of action

This section will make the point that Hume's conception of character does not match Kupperman's, which holds that character is partly a matter of patterns of behaviour. This can be seen in a passage from the *Treatise* in which Hume discusses actions signalling one's character (1978, p.575). This shows that, while Hume clearly believed character and action to be linked, he did not believe that one's actions are constitutive of one's character as Kupperman does. Rather character, and its constituent traits, determines behaviour in a Humean model, as Hume details when explaining the importance of cause and effect in human action (1978, p.411).

The differences between the two accounts are in part caused by Hume's scepticism. The character of one's own thoughts are experienced via first-person perspective and so are not open to sceptical doubt, however the actions of others may or may not be a good guide to their character. Thus Kupperman's position will not be attractive to the sceptic, since actions cannot be constitutive of character if they do not accurately represent one's character. Traits, like any disposition, can be finked, meaning an object can react to stimuli in a manner consistent with its having a certain trait when the object does not in fact hold the trait consistently. In terms of physical dispositions a wire might be thought to be electrified because each time an observer touches it they receive a shock, when in fact an electrical current is rigged to run through the wire whenever sensors pick up an object, like a hand, coming close to it. The observer's conclusion will be that the wire has the property of being electrified, when in fact this only happens when certain conditions, which can be hard to detect, are met.

Obviously it is possible for people to misrepresent themselves in the same way, such as to carry out patterns of behaviour they know to be estimable but for nefarious purposes. In the case of a wealthy tycoon giving to charity to look appealing to the public, or to officialdom for merits, it seems inappropriate to state that they have a truly charitable character – like the wire that is not really electrified, the tycoon simply acts a certain way under relevant conditions with the result that observers come to believe that the behaviour is constant. It might be reasonable to speak of their actions having the appearance of generosity, given that this is how generosity typically expresses itself, but they as people lack that trait – in

Humean terms they lack truly virtuous passions and should be regarded as 'successful knaves' (1975, p.283). Kupperman's account can account for this by claiming that the tycoon's actions in fact represent self-interest, since they do not give when there is nothing to gain from giving, but the sceptic will always be inclined to worry about the existence of *truly* successful knaves capable of fooling any observer.

Other examples show the converse problem; that traits can really be present without causing any relevant behaviour. In the world of physical dispositions, the trait of a vase to be fragile could be masked by wrapping it in cotton wool. The standard test for the disposition of fragility, dropping the object onto a hard surface from a height, will not then reveal the disposition the object possesses. This is analogous to Hume's example of 'virtue in rags' (1978, p.584), in which generous people may be unable to act generously owing to their impoverished circumstances.

Cases of masking and finking show that action is a fallible guide to character; masking in particular highlights how relevant behaviour is not necessary for the existence of a trait because interferers may block the manifestation of the trait. Hume's insistence that virtue in rags is still virtue aligns him with this position. For these reasons it seems right to say that actions are motivated by character, and as such are reasonably good indicators of character, but that they do not themselves constitute character.

Evidently both public and private data plays a role in general conceptions of character, with Pitson preferring to focus on the private, mental qualities of a person (2002, p.83) while Kupperman focuses almost

exclusively on the public patterns of behaviour. Hume himself, as Baier notes, focused largely on the public aspects; his discussion of virtues and vices centre on the utility of such traits to the public, and a summative view of character, with no need to appeal to hidden powers or lofty metaphysical notions, sits well with his empiricist sensibilities. The next section will put forward an alternative view: that Humeans should understand character as a purely mental feature, with only contingent links to reputation or patterns of behaviour.

4.4: Character as mental properties

This chapter has so far analysed the ideas that character is in part constituted by patterns of behaviour, that it is equivalent with reputation and that it is equivalent with one's self. This section will continue to argue that these factors, while related to the issue of character, do not even in part constitute it. The same can be said of the self, which can be distinguished from character for the reasons stated above. The internal aspects of character will be highlighted, with the intention of showing that Hume's claims about character are not founded in one's behaviour, or one's self, but entirely in the constitution of one's mind.

Pitson highlights the mental aspects of character in his analysis of Hume, writing that character may 'be understood to refer to the mental qualities or principles which, collectively, make someone the kind of person he is and establish him as a moral agent' (2002, p.86). This definition of character stands in contrast to Kupperman's, which focuses on the patterns of behaviour that follow from such principles. That character traits such as

kindness or cruelty are mental qualities sounds like a reasonable claim, but a commitment to bundle theory, and the empiricist epistemology that influenced bundle theory, make this view problematic.

It seems clear that these traits are not of the self, and that Hume did not believe they were. We cannot introspect upon our character traits whenever we wish; they appear only as fleeting occurrent impressions in the self, if they appear at all. Where exactly character traits are located is thus a problem for the Humean. To solve this dilemma, Pitson uses textual support from Hume to argue convincingly that there is more to Hume's psychology than the self. He wonders where Hume's 'durable principles of the mind' (1978, p.575) might fit in the Humean picture, and writes off the self as the basis for their existence. The self can only be fleeting occurrent impressions and, relevantly, secondary impressions or passions. We know the basis for our impressions; our sensory organs feed us information about the world which observers receive in roughly similar ways, where there is a difference in sense perception a difference in the causal chain of sense organ to self will also occur.

But what is the basis for our secondary impressions, our passions, that make up the way we feel about primary impressions? Plainly not in the self, as such a basis is not available for introspection. If one ponders their love of rock music they will find only examples of their fondness for it, not the disposition underlying those feelings. Pitson instead argues that these feelings are the result of dispositions in our minds, dispositions to think and so react in particular and regular ways. Character is all about how we construe the world around us, for Pitson, with our patterns of behaviour

only following naturally from how we feel (2002, p.90). Certain kinds of dispositions are equivocal with character traits, altogether they constitute our character.

This chapter has so far focused on providing an account of character that can make sense of Hume's virtue ethics, which relies heavily on the concept. Since Hume did not write directly on the subject himself, other sources and inference have been used to present a view that does not conflict with Hume's commitment to empiricism. The idea that character is equivocal to the self has been rejected as character traits, or 'durable principles of the mind', are not available to introspection as the self is. It has also been argued that Hume did not believe character to be constituted by factors external to our mind, such as our behaviour or the opinions of others, because the vast majority of Hume's work on character focuses on mental traits, which are not susceptible to sceptical doubt as external factors are. The position of this thesis moving forward is that Pitson is right to focus on secondary impressions in matters of character, and that persons possess features of the mind that determine these impressions, which constitute how we construe the world around us. This thesis will now proceed on the premise that character is as Pitson describes, a bundle of dispositions in the mind that give rise to our secondary impressions. This conception of character demonstrates the three essential properties of character derived from Hume; it explains and predicts behaviour, grounds moral responsibility and is linked to the management of the passions. The following sections will further investigate the concept of character, with the

aim of showing that some mental dispositions are not relevant to questions of Humean character.

4.5: Character and personality

There is a popular distinction in social psychology between character and personality (Banicki, 2017), and it will be necessary to explore the distinction here to avoid confusion later on. Personality traits are listed as sexuality, mental illnesses and the degree to which one is an extrovert or introvert; essentially any dispositional trait that one cannot be held morally responsible for. Personality traits can be said to shape our characters but are not appropriate targets of moral praise or blame. It is perfectly appropriate to speak loosely of someone possessing an introverted character, yet we do not standardly think someone morally laudable or reprehensible by the absence or presence of that trait. Moral traits are virtues and vices, and are made distinct from non-moral traits by the fact that one is held morally responsible for possessing them, or for actions that emanate from them.

The fact that Hume did not explicitly define his conception of character makes it difficult to say how character and personality traits are linked in Hume's ontology; it is unclear how broad Hume intended his conception of character to be. It seems intuitive to say that shyness is more intimately linked with character than sexuality, and that one is more responsible for the former than the latter, but Hume kept his analysis limited to the more obviously morally evaluable virtues and vices, which he takes each person to be morally responsible for. The crucial point is that all types of trait will

need to rely on a similar metaphysics of dispositions, since all are understood as a tendency to think or behave in a certain way. Hume provides no such metaphysics but, if Humeanism is to be consistent with modern psychology, one may be necessary.

4.6: Character and skills

The question of distinguishing skills, or natural abilities as Hume terms them, from virtues is a troubling one for Humean virtue ethics. The four criteria for determining which traits are virtues, usefulness to the bearer and society and agreeableness to the bearer and society (1975, p.268), cannot explain why a talent for running, which could in the right circumstances meet all four criteria, is not to be considered a virtue. Other similarities muddy the waters further; both virtues and skills are developed through practice in the Aristotelian tradition of virtue ethics, meaning that both are acquired in similar ways. They can also overlap when it comes to performing a virtuous action, for example the more a doctor knows about medicine the more they will be able to virtuously help the sick.

Hume proposed a distinguishing feature of virtues and natural abilities; that virtues and vices can be changed by education while talents cannot be. Terence Penelhum writes this explanation off, mentioning 'long-suffering piano teachers' (1993, p.138) to support the idea that talents can indeed be taught. However one might think that Hume here meant something different by natural ability, something more like a disposition or capacity to learn than an actualised ability. We might say of a new and promising student that they have a talent for something, meaning that they show promise

rather than they are already skilled. This reading aids this thesis somewhat, as it highlights another area in which Hume exhibited a belief in dormant dispositions, but it is again beset by difficulties arising from conflicts with Hume's basic epistemology. The unique problem here is that Hume cannot be justified in saying that abilities cannot be altered by education – that Hume believed in abilities as dispositions at all is difficult to square with his philosophy, this added certainty of their permanent natures is an even greater leap, and is not a position that will be supported here.

Rees and Webber (2014) present an alternative solution to the question of how to distinguish virtues from skills. They argue that motivation is the distinguishing feature; that virtuous action requires the right kind of motivation while skilful action does not. To highlight this, they reference two examples of virtuous action, one from Foot and one from Williams, in which one person must rescue another from drowning in a river. In the example from Foot, a tracker, who has acquired skill in woodland survival tactics, hears a boy crying from a river and rescues him in a competent manner. In the example from Williams, a man, who is faced with saving his wife or a stranger from a river, virtuously but clumsily attempts to save his wife. They conclude that the two people are equally virtuous (2014, p.3), as their motivations are equally estimable, even if the actions of the tracker displayed more excellence in that the rescue was conducted more skilfully.

The tracker's skill in rescuing children from rivers certainly brought pleasure and utility into the world, and so may appear to be a virtue in itself, but it brings these things only contingently. Rees and Webber claim that 'In the absence of the right motivations, any skill can form part of a vice

rather than a virtue' (2014, p.3). It could be that a skilled tracker uses their powers to nefarious ends; perhaps they have set up shop near a particularly dangerous river to extort money from the people they fish out. Virtues cannot be disassociated from their motivation in this way; part of a person's being kind is that they are motivated to perform kind actions, while a person's being skilled does not necessarily entail any motivation at all. Thus skills can be neatly distinguished from virtues; Hume's conditions that a trait must bring pleasure and utility to be a virtue remain sound with the caveat that a virtue must be motivated in the appropriate way.

4.7: Moral and intellectual traits

Though intellectual character is not as central an issue to this thesis as moral character is, it resembles it in some respects. It too is constituted by durable traits; inquisitiveness, adaptability, imagination and so on, and it too can be posited in explanations and predictions of behaviour. It also plays some role in construing the world around us – it will determine whether we view a problem as simple or complex for example. However Hume did not analyse intellectual character to nearly the same degree as moral, making the tensions in his writings easier to identify with reference to moral character.

The divide between moral and intellectual traits goes back to Aristotle, and has an intuitive basis. We would not want to praise ability in mathematics as a moral achievement, just as it seems counter intuitive to praise generosity as an intellectual feat. Some traits blur the distinction between intellectual and moral and are worth examining. Kupperman cites

strength as one of these, defining 'strength of character' as being a person's ability to originate their own actions, rather than having those actions or thoughts determined for them by external forces, such as social pressures (1991, p.14). Hume similarly writes about strength of mind (1978, p.418), which is the ability to pursue one's calm passions over the violent.

Effectively one with strength of character or strength of mind will be the kind of person that confounds situationist studies, following their own goals and principles heedless of situational cues. Strength does not immediately seem to be a moral virtue, Kupperman argues, as one's character could be cruel, greedy or vain despite external influences, while the same is not obviously true of a trait such as kindness. Nor does it seem intellectual as it does not need to be learned in the way one learns technical expertise, though it will shape one's intellectual character – a strong mind will resist situational cues in intellectual as well as moral exercises. However strength is plainly a matter of character, it certainly seems like the kind of trait that 'goes all the way down' in anyone's book; it goes deeper than an affinity for tea and might be the first trait that comes to mind when describing certain people. Other traits, such as one's sense of humour or taste in art, are not standardly open to moral evaluation and yet are not obviously intellectual features, while remaining 'of character'.

Hume does not make specific reference to the divide between intellectual and moral traits, although his work on aesthetics could be seen as an argument for making taste an intellectual virtue, given the way one must rise above one's own prejudices to observe art from a more objective viewpoint. What is more relevant here is the distinction between those

mental features which are counted as constituents of character and those that are not. Fleeting sense data would not be included in a description of one's character and nor would knowledge. Understanding why will shed light on the nature of character, which should be understood as a feature of the mind that influences one's self in some way. Knowledge and sensory data are not thought constituents of character or personality because they are fundamentally impersonal, while character is all about individuality. My knowledge of London being the capital of England is not in some personal sense *my* knowledge; it is not distinguished in any way from your knowledge of the same fact. A piece of art will, *ceteris paribus*, look the same to one person as it does to another; if a third party were somehow to steal into their minds they would be unable to immediately tell one person's perception from the other's. Their reaction to the piece of art will vary, however, as determined by each person's character. Their consideration, enjoyment and opinion of the art varies in accordance with taste in a way that simple impressions, on Hume's view, do not. The extent to which one avoids Hume's errors in judgement will also shape how one reacts to a piece of art but, as noted in the previous chapter, this too will be determined by the passions, which are a matter of character. Determination to avoid prejudice, or ignorance of one's own lack of context, are personal factors.

Intellectual traits are themselves divided into intellectual personality traits and intellectual character traits. The distinguishing features of the two will resemble the features that distinguish character and personality more generally: intellectual character traits are flexible and one can be held responsible for their presence or absence, while intellectual personality

traits are rigid and not appropriate sources of praise or blame. The distinction between intellectual personality and intellectual character traits is helpful, though it is not an exact one; our rationality, numeracy, literacy, knowledge, imagination and memory are constitutive of what kind of person we are, so seem linked to notions of character, but their relation to responsibility varies across the list. Memory, for example, might be thought to have a moral, as well as intellectual, aspect. Remembering to feed a neighbour's pet while they are away is the kind of thing an ethical person does, while forgetting to fulfil one's duties is characteristic of vice. The focus of our memory differs across characters, and is the appropriate target of moral judgement. Numeracy, on the other hand, does not exhibit this moral dimension, nor does it seem intuitive to refer to it as the kind of trait that shapes our personality.

4.8: Conclusion

This chapter has analysed readings of Hume, Hume's work itself, and general conceptions of character to establish what a Humean conception of character could look like, given the limitations of Humean scepticism. Character has been shown to be a matter of mental properties, as opposed to behaviour or reputation, and it has three main roles: to determine behaviour, to ground moral responsibility, and to influence other mental operations such as the passions. Character has been distinguished from personality, while moral traits have been distinguished from skills. Finally intellectual traits were considered, with some belonging to Humean notions of character and some belonging to personality. The next chapter will ask

whether this conception of character is possible, given the sceptical challenge to ideas of character based on situationism.

Chapter 5: Scepticism about character

This chapter will examine one possible solution to the tension between characters and selves: the idea that character is a fiction and Hume is mistaken in employing it. Harman has argued that empirical studies, chiefly Milgram's experiments in obedience and the 'Good Samaritan' experiment by Darley and Batson, have given us good reason to be sceptical about the existence of character, while Doris uses similar data to argue that character traits are more thinly sliced than typical character talk allows for (2002, p.62). Both are situationists; they believe that the situation is a better indicator of behaviour than character. It is clear that Hume did not share this position, as he puts a great deal of weight on character in his ethical writings, yet it does seem congruent with the rest of his work, particularly as Harman accuses the concept of character of being empirically unverified. If his claims are correct then any Humean will need to take them on board, jettisoning a cornerstone of Humean ethics in doing so, but resolving the tension between selves and characters. The claim here is not that a Humean, believing in a rapidly fluctuating self, should naturally align with the situationists, only that this scepticism means that the Humean cannot simply assert that character exists; the concept needs at least some arguing for, or a defence against this attack.

Hume's odd certainty about the predictive power of character certainly seems in question; here two forms of situationism will be analysed to test that certainty. The stronger, presented by Harman (1999, p.315-331), is that we should be extremely sceptical about character traits, and that positing them as an explanation for behaviour could be not only wrong but harmful.

The weaker, presented by Doris, holds that character traits are very thinly individuated, so thinly that the situation is a better predictor of behaviour than character. This chapter will argue that Harman's claims are unsubstantiated; his data does not entail the degree of scepticism about character traits that he claims it does. It will also be argued that Doris' work is convincing, inasmuch as character traits are often thinly sliced; however it will be concluded that Doris pays too much attention to the situation itself and too little to the construal of the situation, when the latter should be the focal point of ethical debate.

5.1: Harman's scepticism

First to establish Harman's position on character. Harman argues for scepticism about long-term, stable dispositional qualities of the mind, or character traits, on the basis that they lack empirical evidence (1999, p.315-331). While he does not firmly state their non-existence, he is extremely sceptical of them, claiming that acting as though they do exist can cause great harm. He summarises 'our attributions tend to be wildly incorrect...there is no evidence that people differ in character traits. They differ in their situations and their perceptions of situations. They differ in their goals, strategies, neuroses, optimism etc. But character traits do not explain what differences there are' (1999, p.329). That people could not possess character traits may seem an extraordinary claim; we all feel we know our contemporaries well and are able to predict their behaviour based on their character with a good degree of success, as Hume did. However Harman has a good deal of empirical evidence to motivate his doubts, which

will be analysed here. The chief contention will be that Harman's conclusions, such as the summary quoted above, go far beyond what the evidence suggests, and that everyday observational empiricism should be enough to convince us of the existence of character traits to some degree.

Harman's approach may seem particularly congruous with Humean philosophy because it mirrors the kind of sceptical, empiricist, eliminativist projects that Hume undertook in his explanations of the self, the laws of nature, and especially causal powers. In each case there is an explanans that is purported to be the root cause of disparate but related phenomena which is subjected to sceptical doubt and the charge of being explanatorily unnecessary. Lacking direct empirical evidence of the explanans, Hume and Harman declare that there is no justification for belief in their target. The analogy with causal powers is particularly strong, as character traits can be understood as a kind of causal power. The following section will show why Harman's scepticism about character traits is unwarranted, just as it will later be shown that Humean scepticism about causal powers is, in the case of character traits, unwarranted.

5.2: Harman's data

Harman focuses his analysis around two studies: the Milgram experiments from the 1960's - 1980's and the 'Good Samaritan' experiment by Darley and Batson. The former is popular in social psychology and situationist critiques; for this reason it will be explained briefly here and analysed less briefly, with an eye to using it to illuminate some general faults in situationist reasoning.

The setup of the experiment is that Yale University advertised for a test subject in an experiment concerning memory, a cover for Milgram's real experiment on obedience. The subjects, who came from a range of socioeconomic backgrounds, were met by an experimenter in a lab coat and an actor portraying another subject in the experiment. The real subject was designated the teacher while the fake got the role of learner. The two were placed in separate rooms with an intercom so that the teacher could ask questions of the ersatz learner and hear their responses. For every wrong answer the teacher was told to administer an electric shock from a dial that ranged from 'slight shock' to 'XXX'. The 'learner' in the other room faked reactions to the shock from a script, starting with grunts of pain and proceeding to cries of pain, begging to be released and finally falling silent. If the subject asked to end the experiment five times in a row the experiment would be ended; the first four requests were to be met politely but firmly with non-coercive prompts. Ultimately 65% of the subjects were fully obedient, giving shocks labelled 'danger: severe shock' to no response from the 'learner' and then giving another shock. From this Harman concludes, rather vaguely, that while most people intuitively blame the obedience of the subjects situational cues are actually the determining factor in this surprising behaviour (1999, p.323).

Harman's argument is further supported by the variations on Milgram's original experiment. These variations differ from the original experiment only in apparently minor situational cues; for example in some iterations the experimenter, playing the role of scientist and authority figure, would not wear the traditional white lab coat. In others the university supposed to be

carrying out the experiment was claimed to be a less prestigious one than Yale, or would have the subject sit in the same room as the 'learner' rather than having them communicate through an intercom. These changes in situation had predictable outcomes; the lack of a white lab coat or the support of a renowned university saw obedience drop, as did proximity to the alleged victim, supporting the theory that situational cues, however minor, are a better predictor of action than character (2005).

This section will not intend to attack the methodology of the experiment; although there do seem to be significant problems they are addressed by Doris (2002) and are not within the scope of this paper. Instead it is Harman's interpretation of the results that will be critiqued, as they are based on a faulty understanding of character and thus miss their mark. The most obvious response to the conclusions Harman draws, which he himself draws attention to (2002, p.71), is that the experiments really tell us very little about character as a concept – they may give us some idea as to the character of the subjects in the experiment, but not to the long-term, stable dispositions Harman is trying to cast doubt upon. A test that succeeds in doing so, though difficult to devise and carry out, would need to show that the same people, and so presumably the same characters, display contradictory traits across scenarios in which we expect stability, with the differences in character correlated to changes in situation.

It is difficult to know what those changes in situation would need to be, as it is not clear what the situationist would want to prove with such an experiment. If the changes are drastic, then the virtue ethicist will reply that changes in behaviour are only natural, while if the changes are small, as

with the Good Samaritan experiment discussed below, then we would have to draw conclusions not only about people's character, but about their rationality. If the experiment shows that my wants and desires do not map onto my behaviour, but random situational clues do, then it is not just virtue ethics that will need revising; any account of autonomous rationality will need the same. Crucially it will also be difficult to determine whether the situation or the agent's construal of that situation is the key causal factor. If construal is an integral part of Humean character as I argue here, then situationism is far less of a threat.

Harman describes the kind of experiment that would satisfy him as to the existence of character traits; 'To have different character traits they must be disposed to act differently in the same circumstances (as they perceive these circumstances)' (1999, p.317). Here is why situationists and virtue ethicists read Milgram's results differently and why a clearer conception of character is necessary, one that puts the focus on an agent's construal of the situation. Reading Harman, one might think that an identical situation is going to be more or less identical with how an agent perceives their situation. However the possibility still exists that each and every one of Milgram's subjects construed their situation in an entirely different way, that construal constituting their character, and that it was this that caused their behaviour. Their similar behaviour is no indicator that similar construal took place. The subjects could see the experimenter as a learned scientist who knows best, an employer, a figure of authority to be obeyed, a way of escaping moral responsibility for their actions, or an accomplice. They could view the 'learner' as a willing volunteer, a victim, a

cry-baby, or a means to a pay cheque. The variations on Milgram's experiment do little to oppose this view; white lab coats and prestigious names facilitate the construal of a situation in a certain way and thus foster less obedience. The paper 'Happy to have been of service' (2014), which offers readings of Milgram that contradict Harman's to focus on character, note that Milgram took care 'to ensure that participants construed his project as a worthy scientific enterprise' (2014, p.7). The virtue ethicist will want to focus on this construal, unfortunately not recorded in Milgram's experiment, while the situationist will want to focus on the situation, as Harman does and Doris largely does. Which is the best explanation is certainly not decided by this experiment, meaning it does not do the work Harman wants it to do.

One might think that Harman's proposed criteria for the test of the existence of character traits, quoted above, can be satisfied by empirical data without the need for a full-fledged experiment. The other night I introduced four close friends to a new board game and each acted not only differently, but also predictably; similarly I can predict with a fair amount of accuracy when a friend will arrive late, and who will enjoy which music. This is founded upon what Hume called the constant conjunction between certain characters and certain actions, and it appears to work well enough to base all kinds of social interactions on. Harman could respond to these unscientific 'tests' in two ways: either that subtle situational factors determined the differences or that they perceived the situation differently. The second response should suit the Humean fine, as they will be more inclined to say that perception is more a matter of character than of

situation, unlike the situationist. Harman is right to say that a cross section of people will act in similar ways if they all perceive the situation in the same way – yet it could be responded that peoples’ character will prevent this from uniformity from occurring. The theory that subtle situational factors caused the differences, apparently supported by masses of empirical data, still needs to be debunked.

Doris makes a comeback to any kind of reliance on folk psychology, comparing it to folk physics (2002, p.13). Our folk intuitions about physics are often wrong, Doris argues, in much the same way our folk intuitions about one another’s characters are wrong. To highlight our folk ignorance he uses the example of a wooden door, which folk would say is solid but which physicists know is mostly empty space. This analogy could be seen as support for the virtue ethicists wanting to reject situationism. Wooden doors still exist, there are no eliminativists about wooden doors based on these findings; if there were personality psychologist Epstein’s criticism of ‘the strange position of denying our everyday experience and of accepting the laboratory as reality’ (1979, p.649) would certainly apply. Instead we have simply learnt that the concept wasn’t what we thought it was. If the studies Harman relies upon do show exactly what he wants them to show – that situations are a greater driver of behaviour than character – then further work will need to be done to explain the relation between this diminished concept of character and the situations surrounding them. Harman’s claim that situations ‘swamp’ (2002, p.46) character requires more empirical investigation.

Harman does provide more challenging examples. He points out that, when questioned about how an object dropped from a moving plane will act, folk intuition tends to be that it will fall straight down, when in fact the object will fall in an arc (1999, p.315). Harman reasons that this kind of error is motivation for questioning whether our moral intuitions are similarly erroneous. However this move seems suspect, largely because of the amount of chances people encounter to disprove their hypotheses. For most people the assumption that the object will fall straight down from a moving plane is just that – an assumption. People believe it because it seems obvious until you see it being tested. The belief can then be corrected; Harman mentions that bombardiers need to overcome this assumption in order to be accurate. It is difficult to find an analogy here with character traits. We are constantly exposed to other people and their consistent behaviours, which is the cause of our beliefs about character traits. We can, after learning about scepticism of character traits, continue to perform folk experiments on people we interact with regularly, with the same convincing results that we found before. For reasons explained above, Milgram and similar experiments do not disprove our folk ethics as Harman's examples disprove folk physics.

If Milgram does not lead one to think that traits do not exist, what conclusions could we draw from it? It is difficult to tell, largely because of the limited pre and post experiment interviews with the subjects. The virtue ethicist analysing these findings is going to want to know why some subjects stopped and why some didn't, expecting a plethora of reasons founded in character traits; a question that the few interviews Milgram did conduct did not address. Doris reasons convincingly that it is highly unlikely any

thought the experiment was a fake (2002, p.45), backed up by an experiment in which used real shocks were administered to a live puppy which found similar degrees of obedience. But we can still suppose that some followed out of blind obedience, some fear, some may have been genuinely uncompassionate or even sadistic, some may have done it for the money, and likely more than some out of a sense of conformity - a trait one might think quite prevalent in Americans in the early sixties, not a decade after the height of McCarthyism. Harman makes a point of how confused the subjects were by the apparently pointless and cruel experiment – while he argues that these situational features caused the agents’ behaviour, I would wish to focus on how the agent perceived those changes. These character driven, psychological accounts of the subjects’ behaviour do not seem a worse reading of the results than Harman’s, and have the added benefit of fitting in with our common sense intuitions about characters.

The paper ‘Personality predicts obedience in a Milgram paradigm’ (2014) makes the above case against Harman; its authors claim that two of the ‘big five’ personality elements (Saucier, 1994, pp.506-516), conscientiousness and agreeableness, do in fact correspond to how high a shock the subject would administer – the more prevalent the trait, the less likely they were to obey the experimenter up to the highest voltage. The ‘big five’ are referred to as personality traits rather than character traits in psychology, but they are closer to the definition of character traits presented in section 4.5 – they are changeable and therefore the appropriate targets of praise or blame.

‘Conscientiousness’ is a trait with a clear moral element, as we would hold a conscientious person worthy or moral praise. ‘Agreeableness’ may appear to

lack that moral dimension, though it may come close to Hume's conception of 'wit', which he does include among the moral virtues (1978, p.590). We can therefore conclude that the two personality elements referenced as motivators by Bègue et al are equivalent to character traits as defined earlier. Thus obedience in the Milgram experiments can be explained and predicted by character traits, giving the Humean virtue ethicist solid empirical evidence for their suppositions about the causal efficacy of character traits.

Harman summarises his theory: 'our attributions tend to be wildly incorrect...there is no evidence that people differ in character traits. They differ in their situations and their perceptions of situations. They differ in their goals, strategies, neuroses, optimism etc. But character traits do not explain what differences there are' (1999, p.329). These conclusions seem to go far beyond what can be concluded from Harman's data. Our attributions *may* have been shown to be wildly incorrect only about the content of people's character, rather than the structure – Milgram's findings surprise most not because they show what a poor concept character is, but because they show that people have worse characters than we might expect; worse in the paucity of saints the experiment revealed. The rest of the quote reveals how far my conception of character is from any conception that Harman is attacking. Perceptions of situations are crucial to one's character, while optimism – which Harman claims is a personality trait, rather than a character trait – seems similar enough to a durable disposition to at least make one wonder why it is not subjected to the same kind of sceptical doubt Harman here applies to character traits.

Harman also relies heavily on the ‘Good Samaritan’ experiment, which found that students who were told they were running late for a talk were less compassionate to a person slumped in a doorway, apparently in distress, than those who were told they were on time. This experiment suffers from all of the failings of the Milgram experiment, with the additional defect that stopping to help a slumped figure is hardly a moral requirement in the way that not shocking a person is. Being on time for talks is important for students, and since this experiment again provides only a snapshot view without any (admittedly fallible) self-reporting, it fails to give Harman the conclusive evidence he needs for such sweeping claims.

Harman (1999, p.326) discusses Flanagan briefly, examining the kind of localised traits argued for by Doris. He points out that these narrow traits are not the sort of trait that virtue ethics needs to rely upon, which I will argue here needn’t be the case. He also wonders how such narrow traits would be distinguished from skills or strategies that usually work and so are usually followed. If I can show here that such a distinction is possible, or irrelevant, then the Humean can continue relying on character, having learnt lessons from Doris’ more sophisticated situationist challenge.

5.3: Doris’ localised traits

Doris, like Hume and Aristotle, paints a character as a collection of traits, both behavioural and psychological. On his account, character traits should explain, describe, and predict behaviour, which is motivated by cognitive aspects. However he claims that these traits are extraordinarily sensitive to circumstance, so much so that traits can only ever be conceived as ‘local’

traits, not universal. For a trait like generosity this means that a generous person is not simply generous whenever the opportunity arises. Instead it is the impact of exterior forces, often totally arbitrary, that determine one's generosity in any situation. Doris does maintain that traits are stable enough to react predictably to identical situations, however, giving him the grounding for local traits. The trait of generosity might become 'generosity in sunshine' or 'generosity to colleagues', with enough empirical support for the stability of these traits to confidently use them as predictors of behaviour, but no evidence that we can predict one local trait based on the existence of another, even very similar, trait. So 'courage under rifle fire' is no clue as to whether the bearer possesses 'courage under artillery fire' (2002, p.62). This is directly contrary to the thesis of the unity of the virtues, and to traditional broad based ideas of virtue, which state that virtues tend to come in groups, possibly due to one underlying principle of, say, selflessness. One holding such a position might think that evidence of bravery is reasonable evidence of kindness, since good people will aim toward the good in all situations. Hume did not support such a thesis explicitly, so it is not a point of contrast between the two. However Hume does take global traits as a given, so this section will analyse how much impact Doris' brand of situationism is to have on Humean virtue ethics.

Doris uses a wide range of studies to support his claim, including an experiment by Isen and Levin (1972, pp.384-388) in which a group of people were led to find some change in a phone booth before being presented with a minor moral choice; whether or not to assist another person who had dropped their files. Another group were tested without finding the petty

cash. The study found that those who found the coin were more likely to help than their unlucky fellows, implying that the discovery of the coin determined the subjects' ethical behaviour. He concludes that most people lack the consistency and unity of character that virtue ethics takes almost as a given (2002, p.6).

Doris also makes much of a more challenging variant of that experiment, which may pose a more serious threat to virtue ethics than Milgram or Zimbardo. This study recorded the behaviour of 41 subjects who entered a telephone booth. When they left a stooge dropped a folder full of papers near them, and the number that helped or did not help was recorded. The variable was that some subjects were left a dime in the coin slot by the experimenters, while some found nothing. Out of the fifteen that did help, only one did not find a dime, while only two of the 26 that did not help found the dime. This is more problematic, for my account at least, because it purports to show that the way we construe situations – which I take to be central to character – is easily changed by trivial factors. In the Good Samaritan experiment at least the subjects reacted rationally – there were in too much of a hurry to help the slumped figure, whether they wanted to or not was unrecorded. In Isen and Levin nothing was to stop the subjects that did not find the dime from helping, other than that they did not want to. Here, unlike in Harman's analysis of Milgram, the arbitrary factor genuinely does look like a better explanation than character traits. From this one could reasonably conclude that traits still exist, but that some people don't possess the kind of traits we thought they would.

The first response to this is that, like the Good Samaritan, these tests do not present deep tests of character, and so are unreflective of any subject's traits. Whether the subject stopped to help or not says little about what kind of person they are – it is only when faced with significant ethical dilemmas that people rely on, and are guided by, their virtues or vices. The second is that what Kupperman calls 'weakness of character' (1991, p.7) could be more widespread than some would assume, meaning that in small-time tests of character like this our true beliefs and desires are not expressed, instead only a short term reaction to good fortune determines behaviour. Weakness of character being unexpectedly widespread could be a challenge to virtue ethics if the theory is shown to be completely incompatible with actual human psychology. Though the experiment does not show that it is, it does not prove the opposite, allowing the Humean to be sceptical but hopeful about widespread strength of character.

Doris' stated goal is to prove that global character traits do not exist, but he incidentally eliminates another cornerstone of virtue ethics – wisdom. For Aristotle, wisdom was the characteristic of an agent that has a good sense of the right thing to do in any situation. The wise person will fully understand their ethical options and will be able to divine the mean between virtue and vice; to know when they are being brave and not just rash. They also know the value of the good, preventing the possibility of any kind of devious wise person that knows what is good and does not care. As detailed in an earlier chapter, the concept plays a similar role to moral taste in Hume's ethics. Doris does not attack this concept directly, but it isn't hard to see why it vanishes along with other global character traits; any study that casts doubt

upon the existence of one global virtue must cast doubt on wisdom, so understood, as well. The fragmented characters Doris presents could not possess wisdom, or they would overcome their apparently inconsistent behaviour and move towards unified virtue.

But the elimination of wisdom might be seen as leaving an explanatory gap for Doris, as mentioned above in the discussion of Harman. What is it that explains one's 'courage before rifle fire' and my 'cowardice before artillery fire'? It cannot be the situation alone, as people react in different ways to rifle or artillery fire. Presumably it will be past situations, perhaps an agent's first experience before rifle fire was somehow less traumatic than their first time under artillery fire. However it seems intuitive to explain a lack of bravery before artillery fire by appealing to a trait, however local, and not the past situation. The situation may have caused the trait, but Doris seems happy to let these localised traits do explanatory work. If this is the case then it seems possible that most people's behaviour can be explained in this way; not by reference to the current situation, but by reference to one's local traits, which together go to make a fragmented character. There is thus some evidence for Doris' claim that characters are fragmented in this way, but it is far from clear that he is able to move from that fact to the declaration that situations 'swamp' (2002, p.46) individual differences, or that situation is a better indicator of behaviour than our –admittedly fragmented – characters.

5.4: Conclusion

In conclusion, this chapter has shown that Harman is not justified in his total scepticism about character traits, but that Doris' reading of the same material offers some interesting challenges to the Hume's virtue ethics. How local or global traits are certainly is an interesting topic; however studies in virtue ethics need not grind to a halt while the problem is being solved, nor are our common sense intuitions about the consistency of character traits necessarily flawed. Focusing on construal as essential to character meets Doris halfway – situations are indeed important in determining behaviour, but it is the construal of these situations, supplied by character, that deserves our attention. Doris is sceptical of self-reporting, but more empirical research into how agents construe situations could settle the matter more clearly.

Chapter 6: Powers and Humean metaphysics

Powers theorists often assert that their ontology stands in direct opposition to Hume's own, with their neo-Aristotelian account lauded as a way to overcome perceived difficulties in the neo-Humean account (Groff, 2013, p.1), and using powers to underpin dispositional properties where Hume has only regularity (Ellis, 2010, p.136). Generally speaking, such theorists aim to eliminate the broadly Humean notions of regularity and summative dispositions in favour of Aristotelian realism about active, causal powers as actual entities existing in objects independently of any observer. This chapter will argue that something of a false dichotomy has arisen in the literature, and that in fact the tension between Hume's metaphysics and his ethics can be somewhat dispelled by adopting talk of real, causal powers when discussing character traits. In short, it will be claimed that the truth-makers for character traits, which are an essential part of Hume's virtue ethics, are a certain kind of mental agential power, explaining how rapidly changing Humean selves maintain durable moral characters. However it is not claimed that the arguments given in support of this view apply to the powers of physical objects in the external world, or that a Humean could justifiably make a claim for the existence of powers in the physical world without jettisoning large and fruitful areas of Hume's epistemology. The view that will be developed is one in which Humean character traits are understood as a subset of mental powers.

This chapter will also analyse the Humean metaphysical stance regarding causal properties; largely a negative thesis that rejects knowledge of real causal powers in favour of subject oriented theories of regularity and

summative dispositions, dealing mostly with accounts of dispositions spawned from Hume's philosophy, in which they were only outlined. It will lay out accounts of powers advanced by modern powers theorists, addressing the relevant differences, both terminological and metaphysical, to present as clear a picture of the metaphysics of powers as possible. It will also highlight the main points of conflict between such accounts and the Humean, presenting the argument that none of them makes a Humean ontology incompatible with certain kinds of mental powers. Also emphasised are the points that motivate the inclusion of mental powers in a Humean ontology.

Finally an account of how Humean mental powers might look will be put forward and the challenges it would face highlighted, with the aim of concluding that the Humean conception of character traits as dispositions is compatible with the Aristotelian notion of powers – in terms of mental powers at least – without contradicting any fundamental aspects of Humean philosophy but by expanding on the principles laid out in the *Treatise* and *Enquiry*. This will be a similar form of argument that was presented in chapter 2 when the distinction between the mind and the self was argued for; essentially that Hume's scepticism left him silent on issues that today's Humean may wish to address – will in fact have to address to support a Humean theory of character.

6.1: Hume's account of dispositions

This section will analyse Hume's account of dispositions, which was developed in light of his rejection of causal powers. The account is largely

negative, in that Hume is more concerned with rejecting existing accounts of dispositions than with advancing a positive thesis. As Hume did not spend as much time detailing the positive features of his account several of its features remain unclear.

This section will focus on identifying two of these features, with the aim of better placing the Humean account among competing theories of dispositions. The first feature to be examined, in section 6.1.1, is the causal basis for dispositions: are they to be understood as purely summative or are there powers underlying them? Hume famously rejected the second option, but that causes problems with his ethical theory, particularly with his reliance on long-term, stable character traits.

The second feature is whether Humean dispositions can be captured by any kind of conditional analysis, analysed in section 6.1.2. If they can, then they can be thought of in terms of stimulus and response; the kind of model that is rejected by certain powers theorists. If they cannot, then the Humean, like those powers theorists, will need a new model to describe their operations.

Ultimately the goal here is to show that character traits cannot be understood in the same way that the Humean would understand physical dispositions. In particular, the Humean needs to conceive of character traits as being underwritten by real causal powers and that they cannot be captured by any kind of conditional analysis. The Humean should thus understand character traits in much the same way that some powers theorists understand powers. The following chapters will then make the case

that this can be done without contradicting any key point of Hume's philosophy.

6.1.1: The causal basis of dispositions

Hume is taken to have rejected objective casual powers during his analysis of cause and effect, during which he stated that 'we can never have reason to believe that any object exists, of which we cannot form an idea' (1978, p.172). This was in contrast with his contemporary Locke, who, like most thinkers following Aristotle, argued for the existence of real casual powers whose effect was to confer properties upon objects. Hume's broader point was about the 'necessary connection' of cause and effect, which he took to be 'nearly synonymous' (1978, p.157) with the idea of powers; the point being that no person can sense such a connection or the objects underlying it and therefore that belief in either could not be justified.

Without the notion of necessary connection, powers had no work to do in a Humean ontology and were otherwise insensible, so they were also rejected. Hume explained away the philosopher's tendency to attribute natural objects with 'secret powers' by explaining away necessary connections; by arguing that the common philosophical belief in powers is based upon reason rather than experience, and that reason cannot provide us with a new sensible idea (1978, p.162). As regularity replaced necessary connections, summative dispositions replaced secret powers. Rather than claiming that a window has the power to break inert in its constitution, the Humean position entails that we could only assume, with a high degree of probability, that a window will break when, say, struck by a brick. This

assumption is based on observing the previous behaviour of windows when struck by bricks, and the principle that similar objects will respond in similar manners to similar stimuli. Attributing a disposition to an object is merely making a summative claim about how objects of that kind have always behaved, not a definite claim about the powers that an object possesses.

The chief problem with the summative account is that it robs dispositions of any kind of explanatory or predictive force. To say that a ball is bouncy or a person is irritable is just a summation of their past behaviour. No root cause of that behaviour is posited, nor is there any reason to believe that the behaviour will continue into the future, other than a general belief in regularity. A powers account does not have this difficulty, as powers are used to both explain past behaviour and to ground predictions of future behaviour – they underwrite causation (Mumford and Anjum, 2011), which Hume had rejected in favour of a regularity theory. This gives it at least one advantage over the summative account. However the Humean seems, on the surface, unable to accept the metaphysics of powers because of their commitment to empiricism, Hume's brand of which prohibits forming an idea of anything for which there is not a corresponding sensory impression. The next chapter will address this issue, and will make the case that a certain kind of power, character traits, does cause a sensory impression and so we can form a sensible idea of it. For now it is enough to say that the summative theory is explanatorily inadequate in comparison to the powers account.

6.1.2: The stimulus-response model

The second feature of Hume's theory of dispositions that was left unclear in his own work is how one can describe their manifestation. If Humean dispositions are to be understood on a stimulus-response model, on which objects passively wait until their disposition meets the appropriate stimulus, then they should be able to be captured by a conditional analysis. This section will ask whether such an analysis can truly be applied to all dispositions in Hume's ontology.

One proposed analysis of what it means to be a disposition is the simple conditional analysis (SCA), endorsed by thinkers such as Quine (1960, p.222) and which reads:

An object is disposed to *M* when *C* iff it would *M* if it were the case that
C

Doris' problem of individuation (2002), which motivates sceptical claims about the global nature of character, appears to rest implicitly on the truth of the SCA. The problem states that character traits are not truly universal, as we might assume a trait such as courage is, but instead are localised to fit particular situations, so that courage-in-the-face-of-tigers is an entirely distinct trait from courage-in-the-face-of-lions, with each trait ultimately sliced so thinly that the list of virtues and vices could be endless. This fits nicely with the SCA, because it seems impossible to specify the relevant circumstances for *C* for supposedly general trait such as honesty, or generosity. 'A person is disposed to *be generous* when...' it is called for? Or it is appropriate? Defined like this universal virtues would be open to the charge of circularity, with the aptness of the situation and the aptness of the

virtue relying on each other. Thinly sliced localised traits have no such problem. Doris focuses his attention on the import of this problem in the field of character traits, though it runs throughout the metaphysics of dispositions. Trying to explain any disposition in terms of stimulus and response will run into similar difficulties – they will need to become so specific, to rule out any possible countervailing factors, that the analysis loses any kind of explanatory power. Doris offers one solution to the problem, which is to accept that traits truly are thinly sliced and to conceive of them as such (2002, p.62).

The problem should also motivate the Humean to offer a solution, and there are three main candidates: firstly they could attempt to refine the SCA, a project which has been fraught with difficulties (Van Miltenburg, 2015), to make it appropriate for universal traits. Secondly they could adopt Doris' scepticism of universal traits and give up on Hume's notion of 'general character', a position which, as has been said, might already look appealing due to Humean epistemology sharing a methodology similar to Doris'. Thirdly, the position that will be argued for in this chapter, they could abandon any attempt at a conditional analysis as Mumford and Anjum have done (2011, p.8) and retain the notion of universal traits using another account of what it means to be a trait, one utilised by powers theorists in their rejection of the SCA.

Two other noticeable problems raised in objection to the SCA but which trouble disposition and powers ontologies are worth raising here. One formulated by Martin (1994, p.2) and first raised in section 4.3.2 has great significance, particularly when dealing with character traits. Dispositions, it

is claimed, might in fact be 'finkish' in that the truth of the reaction might depend too heavily on the truth of the stimulus. His example is that an apparently live wire might not actually be live at all, it may only become operational when some complicated machinery detects an object moving close, thus giving the impression to any experimenter that it is in fact live. As highlighted before this is problematic when speaking of virtuous dispositions such as generosity; when conducting a psychological experiment into the nature of generosity we might find that subjects think more deeply on their own generosity than they would in the run of their normal lives, causing them to genuinely exhibit a trait that they do not exhibit when not being examined. An observer would, understandably, attribute the trait of generosity to them, and the SCA would bear this out, as it would with the wire and its trait of 'being-live'. On the other hand an experiment such as Milgram's might cause traits that actually are there to appear absent; if the subject actively disguises their own compassion to comply with the objective, scientific atmosphere of the experiment.

Another objection to the SCA is the problem of mimicry. In physical terms, the problem could be exemplified by the operation of a doorbell. If I visit a person at an agreed upon time every week, ring the bell, and see that they appear at the door, I might come to believe that the bell has the property of summoning my friend whenever it is pushed; a belief that the SCA would support. In reality of course the bell just has the property of being loud enough for my friend to hear and, due to social convention, respond to, without there being any direct link between that bell being pushed and their answering the door. When speaking of virtues and vices,

mimicry is an obvious problem in one sense, in that people could simply ape good behaviour for all kinds of nefarious ends, without truly possessing the character traits that originally caused goodness in the one being imitated.

The example also shows that the mimicry may cause agents to misunderstand the causal chain. We might describe somebody as having a kind disposition if they volunteer at a homeless shelter; this assumption may disappear if we later learn that they are forced into doing so unwillingly. That is not to say that they are not kind, only that their act has been misattributed to a disposition to be kind, when some other factor was the root cause. Without sophistication, the SCA cannot account for these kinds of cases.

These problems will be returned to in section 7.2.1, after an argument is made for adopting powers in the role of character traits. Suffice to say that these problems, as applied to the dispositions of inanimate objects, do not affect my thesis. They have been raised here to highlight some of the problems that dispositional accounts face, and which powers theorists aim to overcome. Whether the Humean could respond at all, or whether powers ontologies are more satisfactory, does not concern this thesis beyond the question of character traits. Suffice to say that, where character traits are concerned, a non-summative account is preferable because of its explanatory and predictive power and that the stimulus-response model is highly flawed.

6.2: The Neo-Humean account

Modern powers theorists often contrast their view with neo-Humeanism, but it will be argued here that this creates a false dichotomy; it also misrepresents the Humeanism this thesis trying to defend and amend in this paper. As explored later in this chapter Mumford argues that one can justify belief in real causal powers using Humean methodology, so it is here worth examining why it is that the conception of Humeanism as the enemy of powers ontologies is so deeply held that one review of a book on powers by George Molnar referred to Humean orthodoxy as ‘the forces of Mordor’ (O’Connor, 2004).

Anjum, Lie and Mumford sum up the debate and its problems nicely:

Much of modern metaphysics seems to present us with a choice between Humeanism, in which all is purely contingent and, as Hume says, “loose and separate” and anti-Humeanism which urges a necessitarian view of the world. Lewis...is probably the most influential advocate of neo-Humeanism in which all is contingent (Anjum, Lie and Mumford, 2013, p.234)

Hume is presented as believing events in the world to be ‘loose and separate’, meaning there is no necessity to tie them together. Instead, one event, such as the cue ball striking the target ball, occurs before another, the target ball rolling away. There is no causal necessity tying the events together, just as no causal necessity ties together my breakfast in the morning with the weather in the afternoon. Constant conjunction of the two events might bring one to believe that there is a causal link between them, and it may even be practically wise to rely on the belief that one will

definitely follow the other, but the certainty will never be anything more than imagined. Anti-Humeanism here holds the view that some events, like the archetypal cue ball striking the target and the target's moving, are tied together by causal necessity, which exists in the world, independent of any observer, and governs causal relations. Certain effects necessarily follow their causes; the way the cause is constituted will make this inevitable.

The primary issue here is that Hume is presented as developing a positive thesis, pure contingency, and that Lewis is attributed with supporting the view. However Hume never espoused a positive thesis like this, rather he adopted a quietist, sceptic stance, as exhibited by his position on the division between mind and self and indeed on powers. The Hume quote from above is more fully 'all events *seem* loose and separate' (1975, p.74). This is a slight difference, but it is the difference between Hume being the enemy of many contemporary metaphysicians and being amenable to powers ontologies. In the same paper Anjum, Lie and Mumford claim that Hume was 'able to argue persuasively that there was no necessity in nature' (1978, p.234). Again the aim of Hume's argument is to show that we cannot know of necessity in nature, an epistemological point, rather than the metaphysical point that none existed.

The upshot of this is that contemporary Humeanism and Aristotelianism are often unfairly placed at odds, when their differing methodologies can in fact lead to identical conclusions.

6.3: Neo-Aristotelian powers

This section will give an account of different powers based ontologies, usually cast in direct opposition to Hume's disposition based ontology. Since this is a rather recently resurrected field, terms are sometimes used differently between authors; clarifying the terminology that is being used will be one of the primary aims here. There is some unity among prominent powers theorists on the question of what powers actually *are*, with most of the conflict arising from what they should be called or the areas in which they are thought to exist. Since this thesis aims to import powers only into a limited sphere of Hume's ontology, this section will deal with some of the major issues with powers that apply to those specific kinds of powers. Of primary importance will be Mumford and Anjum, whose powers ontology could be thought representative of current theories, and Bird, who could be thought atypical, to show the range of viewpoints in the current debate.

Hume's basic epistemology, that we can only know about objects via observation, is often thought to entail that once we observe a ball bouncing we can state that the ball has the disposition to bounce, this disposition being no more than an acknowledgement of past instances of bounciness with no claim to how the ball will behave in the future. A neo-Aristotelian would go further, arguing for the existence of an object, the power of bounciness, that is the underlying cause of this constant conjunction. As Mumford puts it: 'there is a causal basis of the disposition. There is some persisting state, it is claimed, that is able to cause the manifestation when certain conditions are realised.' (2013, p.12).

These powers, Mumford and Anjum argue, may never be manifested (2011, p.4), but are no less real for that. An example given is the fragility of glass. Fragility, understood as a disposition to break under appropriate circumstances, might never actually come to manifest in an object, if the glass object never breaks. However this does not entail that the object is not disposed to break, as it possesses the power of fragility. It is important to note here that Hume himself, despite the positions of his followers and accusations of his critics, does not claim that causal powers do not exist; only that they are insensible and thus empirically unprovable.

Here is the first point of disagreement, the kind of summative anti-realist versus substantive realist debate that has played out over subjects such as personal identity and causality, which Humean scepticism has also had a great influence on. Modern Humeans may apply the same type of regularity theory used by Hume when discussing causation to character traits (1978, p.73), holding that there is too little evidence to convince anyone that these active powers really exist, and that the best we can do is to talk in summative, retrospective terms about how objects have previously reacted. This leaves dispositions with no explanatory or predictive force. Power theorists could interpret virtues and vices as being powers within us, constituting and empowering their holder even when not manifested, substantively more than just summaries of one's past behaviour. As was argued above, despite the obvious attraction of the summative anti-realist view to one following Humean epistemology, it is a view that cannot support much of the rest of Hume's philosophy, and so should be jettisoned to provide a more holistic Humean ontology.

The chief dissention between powers theorists that will be addressed here is the disagreement between Bird (2013, pp.25-47) and most other powers theorists, whose views Mumford generally represents. As stated earlier, there is no major disagreement about what powers are; they are mind-independent entities that cause their bearers to exhibit properties. Though Bird calls them 'potencies', Groff argues convincingly that the name 'powers' is just as apt (2013, p.222). Bird makes the distinction because the scope, not the nature, of his concept is substantially different to his fellows. The disagreement will be addressed here as Bird's view gives precedence to this thesis' later claims about the distinction between character traits, conceived of as powers, and physical properties, which Humean epistemology should commit one to viewing as summative dispositions.

Bird is unusual in that, while Mumford and Anjum expand the ontology of powers, arguing that both categorical and dispositional properties are powerful, and that many areas of philosophy can be tackled anew with powers as a conceptual tool, Bird wishes to restrict the ambit of powers to only the fundamentally metaphysical (2013, p.33). Bird talks in terms of 'levels' – the fundamentally metaphysical, the non-fundamentally metaphysical, the biological, the physical, the mental, and so on. True powers, or potencies for Bird, operate on this fundamental level alone. They then constitute properties on the macro levels, properties which constitute objects which are not themselves the bearers of powers. While other theorists are happy to discuss tables, planes and people possessing powers, Bird believes this to be an error. Bird is also distinctive in his assertion that powers do not figure into accounts of causation. Since he holds that powers

operate only at the fundamental metaphysical level, he claims that they impact upon higher level causal chains ‘only in a general and indirect way’ (2013, p.36), while he is unconvinced that causation at the fundamental metaphysical level exists at all.

Powers theorist John Heil has a similar view of ‘levels’ of reality (2003). He claims that truths about one level, say the physical, need not impact other levels, such as the biological. What is of note is that there is a strain of powers thought that emphasises distinguishing between levels of reality and levels of truth, in which powers can be relegated to just one level without impinging upon the truth of the other levels. This is a useful way for the Humean to approach questions about powers; as will be argued in chapter six some levels of powers might be more empirically available than others. Locating powers at the fundamental metaphysical level is totally contrary to Hume’s basic empiricism, putting the Humean at odds with Bird, as we can have no direct sensory experience of this fundamental level. The class of powers that are mental however, particularly those which are character traits, we have far greater empirical justification for.

6.4: Capacities and Skills

Before that, a note on some of the other terms often used in this debate, so as not to conflate them with those already mentioned. ‘Capacity’ is regularly used to mean the same as ‘power’, not without reason when discussing physical properties. It seems intuitive to say that a car has the power or capacity to travel at sixty miles per hour, even if the metaphysical status of such an entity is in doubt. However, applied to mental entities the picture

becomes even more cluttered by the inclusion of skills. It seems highly unintuitive to speak of a car as having the skill to go so fast, yet intuitive still to speak of a person having the skill, capacity, or power to drive a car. Just as oddly we can speak of an adult that has not yet learned to drive a car as having the capacity to drive, yet not the skill.

The confusion over the term 'skill' can be cleared up easily enough. In Aristotle, the term 'technê' is often translated as 'craft', or 'skill'. The term therefore does not mean simply 'ability to perform', but comes bundled with ideas of sentience, of foreknowledge and memory, properties unavailable to anything except living beings. The important thing to note is that skills are differentiated from capacities purely by the cognition that grounds them, and that this sets a precedent for other mental concepts to be delineated in the same way. If a robot were to drive, we could not sensibly talk of its skill, but of its creator's skill in designing it, just as we would not naturally talk of a dog's skill in being able to catch a stick. There is an intentional factor involved in the definition of a skill that distinguishes it from actions performed by automatons, or from mere luck, by ensuring that, to exhibit a skill, the agent must be aware of what they are doing.

Can capacities then be made distinct from powers? There certainly seems to be an intuitive distinction that the terminology needs to grasp, between what one 'can do' based on empirical evidence, and what one 'could do' based on potentiality. If we see a person driving a car then we have empirical evidence of what they can do, and when they are not driving we can safely assume that they could drive again if they willed it. The picture seems slightly different in the case of a healthy adult person that has never

learned to drive a car. Clearly the lack the skill to drive, the technical expertise needed to perform the task, yet we could say that they have the capacity to drive, as people just like them have the overwhelming tendency to obtain this skill through instruction.

There is some dissention between powers-theorists when answering this question. Stout (2011) makes the claim that in order to be able to manifest the power to look, one must already be manifesting the power to see. So if I wish to look for Wally in a cluttered illustration, I must first manifest the power to see the picture. Having my back to the picture means that I have the power to see it, since I could turn around, but not the power to look for Wally, since I need to first turn around to be able to manifest that power. An object only possesses powers that can be immediately manifested on this account.

However others, such as Marmodoro (2016), argue instead that objects possess a power if they have the potential to manifest it at all, which multiplies the number of powers an object can possess hugely. Powers can be described in 'stages' – in the above example my power to look for Wally is a two-stage power, consisting first of manifesting the power to see the picture, then by manifesting the power to look for Wally. Potentially, an infant on a planet with no Where's Wally books could possess a multi-stage power to look for Wally; stages would include needing to manifest the power to grow up, to fly to Earth and find a picture of Wally. Although objects can, on this view, possess many counter-intuitive powers, it has the advantage over Stout in that powers are no longer mysteriously popping in and out of existence. A distinction between powers would be helpful for the purposes of

this thesis – perhaps powers as described by Stout should be referred to as abilities, that which objects or subjects can manifest immediately, while the further powers described by Marmodoro could be potentials, powers which objects or subjects can acquire over time through exercising their abilities.

This difference in approaches raises interesting questions about powers in the physical world – does a lump of marble possess the power to amaze because it may become a statue? – but the primary importance of this issues is its relation to mental powers, particularly ethical powers.

Essentially I wish to ask whether a perfectly morally good person possesses the power to perform morally bad actions, a question I wish to examine in relation to Hume’s ethics once I have argued that his ethics are in fact underpinned by a powers ontology. Stout would answer no, as the person needs to acquire the power to do something immoral before committing such actions, while Marmodoro would answer yes, with the ‘power to be immoral’ simply having multiple stages.

This chapter has so far drawn attention to the ‘levels of reality’ concept used by Bird and Heil, and the concept of ‘skills’, which necessarily rely on a mental component, from Aristotle. These two should be kept in mind as this thesis continues to advance a positive thesis of how Hume is able to incorporate some powers without abandoning his basic epistemology.

6.5: The empirical accessibility of powers

Jaworski writes that ‘automobile engines have the power to move automobiles, computers have the power to calculate sums, and humans have the power to think, feel and act’ (2013, p.145). On this conception

powers seem much like capacities, and the difficulties in justifying belief in powers via Humean empiricism becomes apparent. Anyone committed to Hume's strict sceptical methodology will be less willing to accept that past instances can grant knowledge of future instances than one without such a methodology. Another route to establishing the existence of real causal powers is to argue for them on pragmatic grounds. This kind of reasoning is explicit in Molnar, who states that 'the reason to accept real powers into our ontology is the work that they can do' (2003, p.186). This argument from conceptual convenience, which succeeded in allowing a Humean to adopt a theory of the mind as separate from the bundle self, faces certain barriers specific to power theory that render it unconvincing. However Hume can be attributed a quietist stance on the question of powers in the mental realm, and just as his silence on the question of the mind left room for theorising, his silence here can be replaced with claims about the modern powers debate without contradicting his epistemology.

Bhaskar rejected the demand that causal powers be directly sensible (2008, p.46). Other theorists have attempted to meet this demand in a bid to satisfy empirically minded critics. This section will ask whether this supposed empirical availability will satisfy Hume's methodology, providing even greater justification for causal powers than the pragmatic benefits of holding such a position. The conclusion will be that the answer will be different for different kinds of powers, but that character traits, as mental agential powers, are among those that do satisfy Humean epistemology.

The primary barrier to any Humean wishing to argue for belief in causal powers will be their apparent insensibility. Recall that for Hume all

knowledge is derived from impressions, and powers do not seem to be the kind of object that we can have direct sensory experience of. One may be able to sense that a glass is fragile by touching it, but perhaps might never be able to feel the hidden causal power that causes the fragility – as with causation, such an impression could even be inconceivable. If this is the case, and there is no impression of powers from which to derive an idea, then one can have no true knowledge of powers in a Humean framework. This stands in contrast to the summative view, which does not posit dispositions as actual objects in the world but instead as explanatory posits, conferred on an object because of repeated empirical observations.

However Dicker (1998, p.104) raises the point that we can on occasion gain an impression of powers. He uses the example of trying to lift a heavy weight, or struggling with a difficult question, events which might give rise to a feeling of power acting upon power in our muscles and minds. If this is a true impression of power, then we can form an idea of powers in general, and so a Humean could be permitted a powers based ontology. Dicker, like Hume, dismisses the idea, responding that one cannot extrapolate out from that one feeling to all objects, particularly when inanimate objects have no way of feeling the power as subjects do. He writes ‘causal relations between the will and its effects are only a tiny sub-class of all causal relations. So, with the possible exception of that sub-class, an idea of necessary connection derived from human volition cannot advance our understanding of causality’ (1998, p.68). However this thesis is not concerned with causality or powers in general, only whether Humean character traits can be understood as causal powers without contradicting Humean epistemology.

Dicker quotes Hume ‘the resistance which we meet with in bodies, obliging us frequently to exert our force, and call up all our power, this gives the idea of force and power. It is this *nisus* or strong endeavour, of which we are conscious, that is the original impression from which this idea is copied’ (1975, p.67). Hume goes on to claim that this does not entail anything about powers belonging to inanimate objects, but for the purposes of this thesis he has said enough.

Hume’s primary goal in these passages was to deny that an introspectable impression of powers could get us any idea of real causal powers in the external world, treating the question of whether this *nisus* is such a power as an aside. Perhaps in a more critical analysis Hume would have found grounds to deny this kind of power as well, however he seems to be claiming that the *nisus*, the internal striving we feel when exerting ourselves physically or mentally, is the root impression of all ideas about powers. From the *Enquiry*:

the animal *nisus*, which we experience, though it can afford no accurate precise idea of power, enters very much into that vulgar, inaccurate idea, which is formed of it (1975, p.67n)

Hume seems to be accusing powers theorists of ascribing this faculty of the *nisus* to inanimate objects without the correct epistemological justification. The fact that the *nisus* really is a causal power therefore plays an important part in Hume’s philosophy; it is Hume’s explanation for how a misguided powers ontology came to be a dominant philosophical doctrine from Aristotle onwards. If the *nisus* was not an example of a real causal power, according to Hume’s epistemology, we might well have formed no idea of powers at all.

It could be argued that this evidence is phenomenological rather than empirical, but whichever Hume believed it was, it is clear from the above quotes that he regarded the experience as being the right kind of impression to give rise to ideas. Hume does not question the existence of impressions and ideas, though they are not available to our senses in a way, because they were always phenomenologically accessible. Thus it is reasonable to treat phenomenological impressions as similar to sensory impressions in their ability to spawn ideas and legitimate beliefs in Hume's epistemology.

Dicker concludes that Hume did in these passages admit to there being a certain kind of causal power which did not defy his empiricism, owing to our ability to gain an immediate perception of that kind (1998, p.105). This impression of a physical agential power may only afford us a vulgar idea of powers in general, but it seems hard for the Humean to deny that it justifies including at least one in a Humean ontology. Interestingly Mumford uses an almost identical approach in an attempt to 'answer the sceptical claims about powers even within Hume's terms' (2013, p.19). He defines 'proprioception' as being 'the muscular sense of required effort' (2013, p.20) felt when, for example, straining to lift a heavy object. This sense, Mumford argues, should give the Humean the sensory impression of causation required for them to form an idea of causation in general, and so even radical Humeans should accept real causal powers. Mumford's line is noteworthy for highlighting that Hume's epistemology is not so inimical to the idea of powers as many believe; however it does not do all the work Mumford wants it to as Hume, in the quotes above, explicitly rejects generalising the impression of powers felt by a person to ideas of powers

within objects, which have not been sensed – a rejection that is entailed by his fundamental epistemological principles. According to both Dicker and Mumford then, Hume concedes that we do have impressions of the power at work when experiencing muscular straining. Whether we have similar evidence of mental powers, such as character traits, acting upon us will be addressed in the next chapter.

It is worth noting that Mumford was, in the passage referenced above, arguing for the empirical accessibility of causation in general, from which the Humean should derive the idea of powers, rather than for the accessibility of powers directly. Mumford explains that this notion of causation in general, though justifiable via Humean empiricism, is not the notion of cause and effect criticised by Hume, in which causes occur temporally prior to their effects and somehow lead inevitably to them (2013, p.19). This conception of causation will be empirically inaccessible to the Humean, Mumford argues, as it is a flawed conception. Instead it is the mutual manifestation of powers that constitutes causation, and it is this that we feel when we exert ourselves. Mumford further claims that the example of inner strain illustrates his broader point about mutual manifestation nicely, since we do not experience a cause (the willing) and effect (the muscles straining) when we exert ourselves, but one concurrent process. Causation as classically understood can, in the external world, always be denied by the Humean, as they can rely on constant conjunction as justification for our causal beliefs, rather than any fact of causation. Physical strain, however, is experienced as one process of willing and straining, and so highlights the flaws in thinking of causation as Hume did.

Here it will be useful to point out that, despite aiming to adopt elements of powers ontology and using an argument similar to Mumford's to justify powers in a Humean framework, a Humean theory of powers does not need to adopt this mutual manifestation model. Mumford claims that this new model is the only way to truly understand causation, which would prevent Humeans from being able to conceive of real causal powers. The defence against this anticipated criticism is that the Humean does not need a full account of causality to support their causal powers; instead they can use the constant conjunction account to link the *nisus* with its intended effect. Whether there is any real causing of the kind Hume sought to question is not a question that needs to be answered; the important thing is that gaps between Humean ethics and epistemology can be bridged by defining character traits as mental agential powers, constantly tending towards their outcome rather than reacting to stimulus as the dispositional model suggests. Accounts of causation are rife elsewhere in Hume, as in the copy principle, where causation is understood as constant conjunction. Thus deriving an idea of causal powers from our impressions of internal straining does not require the adoption of the mutual manifestation model as Mumford suggests.

Another possible objection is that one cannot move from an impression of the *nisus* to ideas of all the kinds of character traits that Hume relies on in his virtue ethics. A virtue such as bravery, it could be objected, does not feel like a force within us, in the sense that we feel a mental force when solving a difficult problem. Can these specific impressions of power be generalised to ground all character traits? In some cases it seems there are

parallel justificatory impressions; several of these cases are examined by Webber (2013), who argues that character traits should be understood not as dispositions awaiting a stimulus, as we might understand the fragility of a glass, but as ‘a power or tendency towards some outcome’ (2013, p.1093). Traits that are not activated by some feature of the environment but seem to exert a constant force on an individual, such as the force a considerate person feels to pay heed to the feelings of others, can clearly be sensed in the way Hume would want. As for traits that might appear to be more dependent on a stimulus, for example the bravery of one that does not seek out danger but will also not run from it, there might also be an equivalent striving. Either the brave person will feel some kind of compulsion to stay and face the danger, or, if they face danger calmly and without any apparent force acting on them, they will feel a compulsion to not flee. Then the power of bravery will not be sensible each time they are brave, but it will be sensible when they consider not being brave, which is the cause of their brave behaviour. Analogously a generous person might feel no force acting on them when they act generously, but they may feel one if tempted to not act generously. Thus each kind of character trait will be accompanied by its own nisus-like strong endeavour, allowing the Humean to justify each one in turn rather than generalising from one experience. A more detailed argument for this analogy will be provided in section 7.1.

6.6: Powers and Character Traits

This section will address further benefits that arise from importing real causal powers into a Humean ontology; namely that they can support

counterfactuals and that they can enliven the Humean picture, which has been criticised for its perceived lifelessness.

One benefit of conceiving character traits as power rather than summative dispositions is that counterfactuals can be used to support any mental powers that may go unrealised, unlike summative dispositions. For example, if a person lives all their life in a cell, all the while feeling the impulse to escape and injure innocents, we can fairly say that they possess the power to be incredibly cruel, even if this power is never instantiated, and the full extent of their cruelty is unknown even to them. This is again distinct from a physical power, such as a car's power to travel at sixty miles per hour. A counterfactual will prove the truth of this as well, but in the personal case there is something empirically felt to give weight to claims made of it. We can tell that the prisoner and the car possess their power by seeing how their engine or mind is constructed, or by looking at past instances of their behaviour, but only in the one case does the subject feel their power waiting to be instantiated. Such a feeling is unobservable to an outsider, but only as unobservable as causation, personal identity, or the mind, all of which we can grant within a Humean framework because of the undeniable purpose they serve in their ontology.

The distinction between powers in people and dispositions in objects is then clearly based in the fact that people have sentience, the ability to observe powers at work within their own minds without the need to see them instantiated. Here Lowe's description of the will as a 'two-way power' (2013, p.177) will prove useful. If we imagine a person running, or deliberating, then we can say they have the power to run, or deliberate. By

saying this we are implying that they have the power also to stop performing that activity. If they could not, then the action would not truly be in their power, and we could not claim they have the same powers as one who has the power to both continue or cease. There is of course no evidence that physical objects have the same capability, if a rock rolls, it does so purely because of physical conditions, with no power to stop itself. The rock does not have the power to roll in the same way that one possesses the ability to summon a memory; it needs the right kind of stimulus, being on the right kind of incline, to roll at all, whereas an agent's power is accessible, open to introspection and present even when not being manifested.

This introspective evidence might also be good grounds for thinking of mental powers, of the kind mentioned above, as tendencies towards outcomes rather than necessitating their outcomes. A feeling of happiness does not always manifest itself in the same way after all; it might tend to make us smile, but this is also dependent on other aspects of our characters, as well as the situation the agent is in. Mumford and Anjum (2011) argue that all powers are like this; they occupy a 'third modality', distinct from necessity and probability, which they call 'tendency'. This entails that powers only tend towards their outcomes, for example paper's power to burn when lit is only a tendency to burn when lit because interferers could always prevent the power from manifesting. This couples nicely with what has been said here about mental powers, and the dispositional modality will be returned to extensively in chapter eight. Certainly the arguments for tendencies of external powers work just as well as for mental powers, but mental powers do not bolster the tendency claim.

They might do if it were shown that the exact same manifestations sometimes led to different outcomes, despite identical manifestation conditions, but it is far from clear that this is the case.

This may be a semantic point, but the language of the Humean in reference to agents is much enlivened with talk of powers rather than dispositions. 'The dead world of mechanism' (Groff, 2013, p.209) is an unappealing prospect to many, and it is important that the terminology reflects the fact that agents, while still part of the entirely causal world, are a wonderfully unique part of it. Terms such as choosing or deliberating may just become metaphors in this view, but they are still metaphors aptly applied, while the same cannot be said for talk of a chair's choice. The term 'power' in the mental sense could come to represent not just our disposition to act, or our habit of continually reacting to external stimuli, but also our willing to perform such an act, hinting at the character motivating it. An agent's having the power of bravery implies something deeper about their choices past and future, it recognises the will to act behind the action itself.

6.7: The SCA revisited

Now that it has been shown how real causal powers can be justified via Humean epistemology, it is worth returning to the simple conditional analysis and noting how, by adopting the dispositional modality model rather than the stimulus-response model that the SCA attempts to capture, the situationist challenge can be resisted.

As mentioned earlier, the SCA seems to entail that character traits are like Doris' localised conception, owing to the difficulty of substituting *C* with

a manifestation condition appropriate for a general character trait such as generosity. However this may be due to a misunderstanding of how character traits operate, as outlined by Webber (2013). Instead of thinking as a character trait as a disposition dormant until stimulated by some hard to define manifestation conditions, we should think of them as constantly tending toward their outcome. Hume's basic empiricism will prevent the Humean from accepting real causal powers in the external world, but in the case of character traits, whose tendency towards an outcome is immediately sensible to us, this conception fits well. This conception of dispositions requires real causal powers, making them a necessary posit in any theory that rejects the stimulus-response model. Not accepting real causal powers commits one to such a model, the best analysis of which is the flawed SCA.

6.8: Conclusion

Ultimately it is fitting for the Humean to discuss physical traits in terms of summative dispositions, reserving talk of powers for character traits to preserve the unique rational and deliberative powers of agents. The ability people have to predict the future, to intend and command their bodies and minds to operate and their ability to report on uninstantiated mental states all serve to distinguish mental traits, or powers, from physical traits, or dispositions. It has been shown that Humean epistemology is not so limited as to deny the existence of all powers; Hume himself admitted that the *nisus* provided an impression of a real causal power, which will be explored further in the next chapter. However the existence of powers in objects

external to ourselves lacks empirical support and no persuasive argument has shown that such a concept would be beneficial.

Chapter 7: Integrating powers into a Humean metaphysics of character and the self

Thus far this thesis has drawn attention to areas of conflict between Hume's metaphysics and his epistemology, particularly as regards his commitment to long-term durable character traits, which are a crucial feature of his ethical theory. The previous chapter concluded that conceiving of character traits as powers, rather than dispositions, solves many of these problems. It has also been argued that powers, in the limited sphere of certain mental features, are empirically accessible without contradicting Hume's basic epistemology.

This chapter will further explain the arguments motivating this reconceptualisation, particularly the argument from the *nisus* mentioned briefly by Hume, expanded upon by Dicker, and interpreted by Marshall and McBreen. Hume's own objections to the *nisus* argument will be put forward and rejected, as they do not apply to the limited utilisation of powers argued for in this thesis. This will aid in cementing which mental features might benefit from being conceived as real causal powers while others, such as the mechanism underwriting the copy principle, are still best conceived as simple input-output functions. Questions that remain about Hume's belief in character traits; the extent to which they change over time and what their metaphysical and causal natures are, will then be addressed.

7.1: The *nisus*

Section 5.5 relayed a point made by Dicker (1998, p.104); that Hume admitted to an impression of power occurring when one experiences the

nisus, or the 'strong endeavour' we are conscious of when exerting force of our own (1975, p.67n). This section will explore the point made by Hume, and later interpretations of it, in more depth, with the aim of showing that Hume's original point supports the claim that certain kinds of mental power are empirically accessible. Using this impression as the basis of the idea of real causal powers, a Humean can then posit the existence of powers to explain the regularity of character, which the summative account fails to do.

During his analysis of causality in the *Enquiry*, Hume supports his scepticism about powers by arguing that the idea of power does not correspond to any impression. He writes that 'our idea of power is not copied from any sentiment or consciousness of power within ourselves' (1975, p.67), a claim which, coupled with Hume's basic epistemological precept that all ideas must be copied from impressions, make causal powers 'unknown and inconceivable'. This last quote leads to a footnote in which Hume attempts to deal with an anticipated objection: that our idea of power is in fact copied from the impression of exerting our force when we meet with resistance in physical objects, which Hume refers to as the animal nisus (1975, p.67n). This is possibly a response to powers realists such as Locke, who makes the claim that any impression of force upon the human body – such as the feeling of the floor beneath our feet supporting us – can provide good empirical grounds for the idea of power. Even handling solid objects can give us some idea of force, as solid objects 'do by an insurmountable force, hinder the approach of the parts of our hands that press them' (1997, p.124). After Hume, Reid proposed a position similar to the one critiqued in the footnote, stating that 'It is very probably that the

very idea of active power, and of efficient cause, is derived from our voluntary exertions in producing effects' (2010, p.204). More recently Mumford has provided a similar argument; claiming that 'the muscular sense of required effort' (2013, p.20) called proprioception is direct empirical evidence of power at work.

Hume solely discusses physical straining in the objection. Dicker later groups in mental straining, using the example of forcing yourself to 'think about something difficult or unpleasant' (1998, p.104). This may cohere with Reid's position, as the 'voluntary exertions' (2010, p.204) mentioned are not explicitly limited to muscular or bodily exertions. We often do experience something analogous to physical straining when we perform strenuous mental exercise; apart from Dicker's example we strain to solve difficult puzzles, to remember something we've forgotten, and, most relevantly to this thesis, to exert control over our own passions. This ranges from forcing ourselves to view a negative situation positively to reining in impulsive or unhealthy behaviour; any situation in which our passions conflict and we feel the need to strongly endorse one side over another. These cases appear to be direct empirical evidence of force in the same way that Hume believed the *nisus* was.

Given that Hume only concedes the existence of one physical power, the analogy between it and the mental powers that are being posited as the causal basis of character traits should be addressed. The analogy will be a successful one if these mental powers meet the criteria that Hume laid out in the *nisus* footnote, and if they do not fall foul of any fresh objections. In the *nisus* footnote, Hume writes that the sensation of straining is enough to

afford one the vulgar idea of a physical power. The straining of character traits is directly analogous to this physical straining, and so should provide evidence in a parallel fashion. This kind of mental straining is particularly evident when character traits are tested; for example an honest person will feel a strong impression of honesty when lying is a tempting option. This straining to tell the truth may be strong enough to overcome the temptation and motivate the individual's honest behaviour. Other traits, such as generosity, will impress upon the subject at times. A generous person, seeing somebody in need of assistance, will offer assistance because they feel a straining to act in accordance with their generous nature. There appear to be no specific objections to a Humean conception of character traits as mental powers, though Hume does offer objections to the Humean conception of the *nisus* as a physical power, discussed below. As these objections do not impede the argument presented here, this thesis holds that Hume admitted to there being empirical evidence strong enough to justify belief in one physical power, and that empirical evidence for character traits as mental powers is strong enough to provide analogous justification for their existence.

Hume reasons that one cannot base an ontology of real causal powers on the impression of the *nisus*, presenting two objections to support this (1975, p.67n). First he points out that 'we attribute power to a vast number of objects where we can never suppose this resistance to take place' (1975, p.67n), listing the Supreme Being, inanimate objects and 'the mind in its command over ideas and limbs' as examples of objects that feel no such resistance. This objection is aimed at any argument that would try to infer

from the impression of the *nisus* to causal powers in the physical world. Secondly Hume points out that the 'sentiment of an endeavour to overcome resistance has no known connexion with any event', by which he means a necessary causal connection. Dicker interprets Hume as meaning that for one to obtain a sensible idea of causal power, one must be able to deduce the cause from effect and vice versa, when instead exerting ourselves physically might result in failure or success (1998, p.105). Hume concludes the footnote by confessing that 'the animal *nisus*, which we experience, though it can afford no accurate precise idea of power, enters very much into that vulgar, inaccurate idea, which is formed of it'. Dicker notes that Hume is here admitting to there being a genuine power being empirically sensed in the *nisus* example, but agrees with Hume's two objections and so treats the case as nothing more than an interesting exception.

However Hume's objections to the argument from the *nisus* do not impede the theory that character traits can be understood as powers in a Humean framework. Here those objections will be analysed to show that they are only effective against the theory that one can gain an idea of powers from an impression of physical straining. They are not effective against the theory that one can do the same from an impression of mental straining. Hume's first objection is that power is often attributed to objects which cannot or do not experience this kind of resistance in the same way as people. He lists the Supreme Being, inanimate matter and the mind as examples of objects imbued with powers when we have no experience of them having powers. This is a valid criticism; this thesis will not seek to establish powers in objects, or claim that there is empirical evidence of

powers in the objects listed that would satisfy Hume's basic epistemology. Locke, and more recently Marshall, do employ the broader notion of exertion to empirically justify powers in physical objects, arguing that feeling the resistance of the floor beneath our feet is sensory evidence of the power of the floor (2014, p.310). However this is not the notion of exertion that Hume discusses, nor is it one that would necessarily be endorsed by modern Humeans, as the original, narrower notion would have to be. Hume's mention of the mind as one of the objects that cannot be imbued with powers is more relevant to this thesis and will be returned to in section 7.5.2.

This first objection fails to prevent one gaining an idea of certain powers through the *nisus* because one does not need to infer from one's own experience to conclude that other people possess mental powers. Instead one can argue that certain mental powers must be posited in order to explain the consistency of other people's behaviour. As highlighted in chapter one, Hume did believe that people possess stable characters, a belief that is at odds with Hume's conception of the self as a perpetually fluctuating bundle of perceptions. Conceiving of character traits as summative dispositions leaves this mystery open; summative dispositions have no explanatory power, they are merely summaries of past behaviour that do not provide a solid basis for predicting future behaviour. Real causal powers, on the other hand, do solve this mystery; people act consistently because they possess mental powers that manifest in their behaviour. Parallel arguments for powers outside of the limited real suggested here do not have the empirical justification as support – for a Humean to posit a

power in their metaphysics, they need a valid idea of a power, which needs to be derived from an impression. The *nisus* provides that impression, allowing the Humean to bolster their psychological theory by replacing summative dispositions with mental powers to explain character traits. Thus Hume's first objection does not apply to this argument, which does not rely on inferring from the *nisus* to other objects or people.

The second objection is that the 'sentiment of an endeavour to overcome resistance has no known connection with any event: What follows it, we know by experience, but could not know *a priori*' (1975, p.67n). Dicker interprets Hume as meaning that, since we cannot make a valid, deductive inference from the *nisus* to a particular outcome, the example can only give us a rough idea of the concept of power, one less rigorous than Hume would like (1998, p.105). This objection might be thought to hold more strongly against a theory of character traits as powers than it does a theory of physical objects possessing powers. If the trait of impulsiveness is a power that exerts force on the bearer, then we have no way of deducing how that power will manifest. The same power of impulsiveness that encourages a person to take up smoking might encourage them to go for a run the next day, or the power could be suppressed or trumped by another power so that no effect manifests at all.

Both Dicker and Marshall have a way of dealing with this criticism. Dicker questions whether Hume is correct in his assumption that one needs to have an *a priori* understanding of how a power will manifest in order for us to gain a cogent idea of one. He writes that 'Hume is saying that in order even to *feel* a power, one must "know" it, apparently in the sense of being

able to foretell its effect without relying on past experience' (1998, p.103). Conversely Dicker claims that we do feel a causal power when experiencing resistance, and that our lack of *a priori* understanding about its exact function is irrelevant. If Hume is unwarranted in his assumption that one must have an *a priori* understanding of a power in order to know it, as Dicker states, then his second objection to the argument from *nisus* also fails to hit the mark. That the feeling of ambition does not indicate the type of action the feeling will motivate does not prevent one from gaining a sensible idea of the power of ambition, one gained from a corresponding impression as per Hume's fundamental epistemology. Rather there is solid empirical evidence for some powers, those that are sensible to us as character traits for example, but that further empirical work needs to be done to establish the exact causal relation between them and action.

Marshall argues against the idea that one needs perfect certainty about causal relations to properly understand them, given Hume's general scepticism about the topic. He points out that we can never be certain that, in his example, a door will open when pushed rather than pushing back, just as Hume claims that we cannot be certain fire will warm us (1978, p.270). Marshall reasons that we still have a wealth of experience of the causal relation and so, in Hume's framework, lack of perfect certainty should not be a barrier to understanding causal relations (2014, p.312).

Ultimately Dicker agrees with Hume that the argument from the *nisus* is of little importance, while Marshall and McBreen use it as a central point in their arguments that Humeans should rethink causation. Hume's conclusion, that the idea is a vulgar one, is challenged by both. Marshall's

counter-objections to Hume's concerns are intended to show that the idea is less vulgar than Hume proposed, while McBreen simply dismisses the idea that there is a distinction between vulgar and non-vulgar ideas at all – 'There is only one concept of power which everyone shares' (2007, p.434) – concluding that the idea of power derived from the impression of the *nisus* is as good as can be found and should not be dismissed by Hume. That the argument from proprioception is still being used by powers theorists like Mumford shows how durable the argument is, indicating that Hume likely was wrong to treat it in such quick fashion.

If Hume believed that the *nisus* was an impression of a real casual power, then it is worth asking what kind of power it is, and by extension what kind of category of powers character traits fall under. The *nisus*, as first presented by Hume as 'the resistance which we meet with in bodies, obliging us frequently to exert our force' (1975, p.67n), should be understood as the feeling of a physical agential power. The power belongs to the muscles straining, leading to an impression of the power at work in the agent, who has control over the exercise of the power.

The further examples contributed by Dicker, straining to solve a difficult puzzle or thinking of something unpleasant, should be understood as mental agential powers. Boccaccini and Marmodoro write, 'paraphrasing Locke, one can say that mental powers concern all that can fall within the compass of the mind' (2017, p.435), a definition that applies to the examples given. The power is agential in that an agent can deliberately bring about its manifestation, as they could by endeavouring to lift a heavy object in Hume's example, or by performing the mental tasks suggested by Dicker.

Character traits should also be understood as mental agential powers; they are aspects of our psychology and can be deliberately manifested by the agent bearing them. This thesis is only concerned with proving the existence of agential powers, as these alone can provide the *nisus*-like impression necessary to lead to an intelligible idea. Mental powers are the focus, as it is their introduction into Humean metaphysics that will solve the tension between Humean metaphysics and virtue ethics.

In conclusion both of Hume's objections to the argument from the *nisus* are ineffective against this particular theory. His first objection, that one cannot infer from the *nisus* to powers in the physical world, does not apply here as only certain mental, agential powers are being posited. His second, that we could not deductively infer a cause from the *nisus*, does not prove that we do not gain an impression of a power when we experience the *nisus* or analogous strainings, from which we can form an idea of a certain kind of power. The feeling of exertion, whether physical or mental, is the impression of power necessary for a sensible idea of power in a Humean framework.

7.2: The epistemology of powers as character traits

This section will revisit the sceptical theories of character presented in section 4.7, aiming to show how the argument above, that the idea of power can be sensibly derived from the impression of mental straining, entails two new objections to scepticism of character traits.

Harman (1999) and Doris (2002) argue, to different degrees, that our notion of character is faulty, and that situations are the more important factor when explaining and predicting behaviour. To support this they rely

on empirical studies such as the Milgram experiment, in which subjects were asked to administer electrical shocks to a stooge and did so with a surprising degree of obedience, and the 'Good Samaritan' experiment, which found a high correlation between subjects being in a rush and subjects stopping to help someone in need. Harman argues that Milgram, who found a high degree of obedience in all of his subjects, showed that differences in the character of the subjects did not matter; what mattered, as variants on the experiment are supposed to show, were differences in the situation. The 'Good Samaritan' experiment is used to reinforce this idea, as subjects in one kind of situation, being in a hurry for a talk, generally chose not to help the slumped figure while those in another, not being in a hurry, generally chose to help. Both Harman and Doris argues that this shows the inefficiency of character traits, and the efficiency of situations, as motivating factors.

Their argument relies on third person accounts of behaviour; necessarily since the experiments involved set out to test large groups of people and collate the resulting data. In each case, the experimenters observe how their subjects behave from a third person perspective. Any conclusion that Harman or Doris draws from them is based on the observable behaviour of the subject, rather than their mental behaviour. However the argument from the *nisus* is based on a phenomenological, first-person account of mental behaviour; chiefly the straining we feel when beset with an impulse to act in a certain way. Traits are only ever discoverable this way, via an immediate, sensory feeling that cannot be masked or finked as dispositions can be when viewed from a third person perspective, and from inferences drawn from this

impression. As mentioned earlier, Milgram's post-experiment interviews were not designed to collect experiential data from his subjects, which is one of the primary reasons his work does not support scepticism about character traits. If the subjects had been interviewed about their behaviour after the fact, a multiplicity of explanations may have emerged – obedience to authority, fear of men in white lab coats, a desire to see the experiment performed properly, or even blind cruelty. Explaining their behaviour in this way, with reference to character traits, allows us to be more exact and nuanced than assuming that the identical behaviour of the participants was the result of identical mental processes. This is the first new objection that the argument from the *nisus* provides against scepticism of character traits.

The second new objection is that the notion of 'character traits' used in the experiments that support the scepticism of Harman and Doris is an entirely different notion than the one presented here. Both Milgram and the 'Good Samaritan' experiment rely on a classic stimulus-response model in an effort to detect typical causal relations. Both experiments put their subjects in tightly controlled scenarios, introducing minor changes in situation to detect how each change affected behaviour, or how alterations in input – the situation – correlates to differences in output – the behaviour. This approach implicitly assumes that character traits can be described in simple input-output terms, entailing that they could be captured by governing laws such as the SCA. If, however, character traits are not stimulus-response dispositions but powers, they will likely fail to be revealed by tests like these. Traits as powers, such as the power to be disobedient, will not reliably be uncovered in an experiment such as Milgram's because

they are not waiting passively for a specific cause, but are active constituents of their bearer at all times. Competing powers, such as the power to obey authority or participate in scientific research, can then be used to explain why they did not manifest. The relevant trait in the ‘Good Samaritan’ case is considerateness, which governs the behaviour of the agent in directing attention to the figure in distress, and competes against the subject’s inclination to not run late.

7.3: The copy principle

This thesis argues that Humean character traits should be reconceived in two major ways. Firstly they should be thought of as powers rather than summative dispositions. Secondly they should be seen as constantly tending to their manifestation, rather than as passively waiting for the appropriate input to produce the appropriate output. However not all dispositions that play a role in Hume’s psychology need to be reconceptualised like this. The mechanism underwriting the copy principle is one crucial mental disposition that is still best understood as a summative disposition that can be captured by a conditional analysis.

Hume’s copy principle holds that all ideas are formed from impressions, which they represent (1978, p.4). The universality of this principle makes it appear very much like a law of nature for Hume; it ensures that no idea can arise without an impression for it to represent. This disposition of the mind to copy impressions into fainter ideas could be viewed summatively, as Humean character traits have classically been understood, or it can be viewed as a power, as this thesis argues Humean character traits should be

understood. Additionally, any disposition can be understood either as constantly tending towards its manifestation, or as passively waiting for the appropriate input to produce the appropriate output.

The objections raised against character traits being understood as passively waiting for a stimulus to manifest do not apply to the copy principle, meaning that one holding Hume's basic epistemological principles should not think of it as tending towards its outcome in the same way as character traits. One of the primary difficulties facing the conception of traits on the classic input-output model is that it is impossible to specify all the circumstances that make up *C*. A leading analysis of dispositions, the SCA, reads:

An object is disposed to *M* when *C* iff it would *M* if it were the case that *C*

The SCA failed to properly capture character traits no matter how many caveats were added and amendments made, in part because character traits do not have the dispositional structure that the SCA was constructed to describe. There is no 'correct' stimulus for a character trait, nor is there an appropriate way for that trait to manifest, making an analysis of this kind unsuitable.

The copy principle does not suffer from this problem; the mind's ability to instantly copy ideas from corresponding impressions can be fairly easily captured by the SCA: *M* represents copying an impression into an idea, and *C* represents being exposed to an impression. The 'appropriate' inputs and outputs are far more tangible than for character traits, so the motivation to understand the copy principle in terms of powers is not there.

Further, though the copy principle is knowable from a first-person perspective in that we observe how we form ideas after gaining impressions, we do not gain an impression of it acting upon us in the same way that we gain impressions of our character traits. Thus the two factors that motivate character traits being understood as tending towards an outcome rather than waiting for a stimulus, first that they cannot be captured by the SCA and second that phenomenological evidence supports such a claim, are absent in the case of the copy principle.

The second major reconceptualization, that character traits should be understood as powers rather than summative dispositions, was based on the fact that powers have more explanatory and predictive power than summative dispositions. It was then supported by the phenomenological evidence mentioned above; we experience character traits as forces that act upon us and manifest in our behaviour, an experience the summative disposition model cannot explain. Again this evidence is lacking, so although thinking of the ability to create ideas from corresponding impressions as a summative disposition does lack explanatory and predictive power, it is the only theory the Humean can justifiably hold. The empirical evidence for thinking of the mechanism behind the copy principle as a power simply is not there.

Thus some aspects of Humean psychology are still better understood as dispositions than powers. Only dispositions that meet the criteria should be seen as powers; those that both cannot be captured by analyses such as the SCA and are empirically available to us as powers.

7.4: Powers in Humean ethics

If Humean character traits are better understood as causal powers than summative dispositions, questions must be asked about the metaphysical makeup of these powers. The first that will be addressed here is the origin of character traits in Hume's philosophy, and the extent to which he believes traits change over time. Answering these questions will help to fix which of the powers theories presented in chapter 5 best fits Humean character traits. This section will argue that there are aspects of nativism in Hume – in that Hume believes all people have natural tendencies to think in certain ways – but that the origins of specific character traits is not a subject Hume wished to address. This section will conclude by returning to the question, posed in chapter 2, of where in Hume's model of an agent powers could be located.

7.4.1: Nativism of Humean character traits

There is some evidence that Hume was a nativist about character traits, believing them to be near essential aspects of all people. In the *Treatise* he writes that 'in all nations and ages, the same objects still give rise to pride and humility; and that upon the view even of a stranger, we can know pretty nearly, what will either encrease or diminish his passions of this kind.' (1978, p.280) This certainly suggests that Hume believed that all people are naturally inclined to be proud of certain features, which Hume lists as 'power, riches, beauty or personal merit'. In the next book he elaborates: 'The minds of all men are similar in their feelings and operations, nor can any one be actuated by any affection, of which all others are not, in some

degree, susceptible.’ (1978, p.575) This comes in the middle of an analysis of sympathy, which Hume believes all people exhibit to some degree, and further reinforces the view that Hume was a nativist. In another section, Hume mentions that ‘we *naturally* approve of (1978, p.578)’ qualities that are for the good of mankind, proving that he was a nativist to at least some degree.

However Hume, who was aware that some people possess ‘monkish virtues’ (1975, p.270) does recognise that some people will not be inclined to be proud of the same things as others. These differences he attributes to differences ‘in the tempers of complexions of men’, which leads us again to the question of whether these tempers are set from birth, or are malleable, or are entirely formed from experience. Hume addresses these differences in temper, making reference to the first ‘mechanic’ to be proud of the work he had done crafting tables and chairs (1978, p.281). This specific source of pride, writes Hume, could not have been ‘foreseen and provided for by nature’. There would need to be other principles dormant in each person, one for each of all the possible sources of pride in existence, for this to be so. This is good evidence of the belief that Hume thought that every person is born with an innate sense of pride, but that the specific object of that pride is not fixed from birth.

When Hume does directly address the question of whether or not differences in moral taste are natural (1978, p.474), he answers that it depends how nature is defined. If it is defined as it has been used here, contrasted with ‘learnt through experience’, he answers that it is impossible

to give a precise answer one way or the other, but that it may later turn out that 'our sense of some virtues is artificial, and that of others natural'.

This shows that Hume was a nativist about character traits in general; all people are born with a capacity for pride, humility and sympathy. However the specific objects of these passions, and the extent of their influence, will differ across people, and will likely be down to their experiences. His assertion, quoted earlier, that one can predict how to excite the pride of a stranger is not intended to capture any necessary truth about people, but is an observational claim – Hume believes that an overwhelmingly large number of people take pride in riches and power, without any one person being born with any specific propensity for riches or power. Elsewhere Hume notes that passions rise from impressions (1978. P.275); presumably meaning that a particular pride will only be felt, and so can be known to exist, when it is excited by its particular object, and cannot be known before. On how these new passions arise in people, Hume writes that 'As these depend on natural and physical causes, the examination of them wou'd lead me too far from my present subject, into the sciences of anatomy and natural philosophy.'

7.4.2: Where do powers fit in Hume's moral psychology?

In section 6.1 it was noted that Hume makes particular reference to the argument from the nisus not being enough to imbue minds with powers, as they are objects that cannot feel resistance. However in chapter 2 it was argued that powers, if they are to be a part of Hume's metaphysics, must be

a part of either the self or the mind of the agent. This section will clarify where and how mental powers would fit in Hume's metaphysics.

First to expand upon Hume's comment that:

we attribute power to a vast number of objects, where we can never suppose this resistance or exertion of force [the nisus] to take place...to the mind in its command over its ideas and limbs, in common thinking and motion, where the effect follows immediately upon the will, without any exertion or summoning up of force (1975, p.67n)

This may be thought to curtail any attempt to infer from the nisus to the idea of other mental powers. However the latter half of the above quote reveals Hume's meaning to be more specific. His objection is directed at any theory that would try to imbue the mind with powers that do not meet with resistance or require exertion, such as the command of one's limbs. One may be physically unable to move one's limbs, or encounter exertion when trying to decide whether to move one's limbs if one is conflicted, but standardly one is able to move one's limbs without having to exert mental force. The act of trying to be charitable however, the process of working to ensure that one feels charitable impulses and views the world with a charitable outlook, is the kind of mental act that does require mental exertion, as other character traits will act against the attempt. Therefore it is justified to view character traits as powers because we can feel their pressure affecting us, where the same is not true of the ability to move one's arm.

With that objection out of the way, does the mind remain a more viable candidate for bearing powers than the self? As detailed in chapter 2, the

mind is responsible for causing many of the mental processes that are only knowable via their effects on the self, such as the copy principle and the combination of ideas. However character traits, and the phenomenology that grounds them, operate differently. It is not just the effects of character traits that appear in the self, as when the effect of the copy principle appears in the form of a new idea in the self. Instead they are knowable as they operate. We know we are ambitious, for example, because we feel ambitious urges as secondary impressions in the self, acting upon us as a kind of mental force, particularly when those urges are thwarted or challenged. If we have a choice to work at getting a promotion or go to the cinema, and doing the work seems far more appealing, we might be justified in thinking ourselves ambitious. Is the trait constituted entirely by these impressions, or does a feature of the mind underpin it as in the case of the copy principle?

Evidence for the proposition that traits are powers possessed by the self could come from Hume's quote that 'the faculties of the mind are supposed to be naturally alike in every individual' (1975, p.80), whereas character traits differ greatly across individuals. However, as noted in the previous section, 'faculties' here should be taken to simply mean 'abilities', with no reference as to the extent of those abilities. Reason is a faculty of the mind, yet Hume should not be read as thinking that all individuals have identical powers of reason, just that they possess reason to some degree. Similarly every miser has the ability to be generous, just little or no tendency to act generously. This is in contrast with the mechanism behind the copy principle, which is a simple function of the mind that is identical in all persons.

The factor that will settle the question, as discussed in chapter two, is the placing of causes for our secondary impressions. Our primary impressions, Hume notes, come from the external world and the ‘animal spirits’ (1978, p.275) of our bodily constitution. Secondary impressions, which constitute how we feel about the raw sense data of primary impressions, do not rely on the external world in the same way. Primary impressions arise in our selves already gilded and stained because they have been filtered through our characters, they have been subjected to the biases and tastes that Hume speaks of elsewhere in his ethical and aesthetic work. If it is not the external world causing us to feel a certain way about a certain object, and if the cause cannot be found in our selves, then our minds remain the best candidate for being the cause of secondary impressions. Therefore, if Humean character traits are best understood as powers as has been argued here, they are best understood as being powers of the mind.

7.4.3: Powers and Causation

This section will detail two accounts of causation, both Hume’s and that of Mumford and Anjum, whose model of powers comes closest to the one advocated in this thesis. The goal is to show that Humean character traits can be conceived of as real causal powers, as Mumford and Anjum understand them. This can be done without contravening fundamental Humean beliefs about causation, despite the fact that Mumford and Anjum and Hume have different accounts of causation.

Hume’s theory of causation holds that we should believe two events are causally linked if they exhibit three properties: they are contiguous in time

and space, the cause precedes the effect, and there is a constant conjunction between them. If a fourth property, the necessary connexion, were apparent, we could say with certainty that the two events were causally connected. As it is we cannot detect any such connexion with our senses, thus cannot form an idea of the connexion and cannot justifiably believe that any two events in the world are causally connected (1978, pp.73-78). As a result of this scepticism, we cannot be certain that effects will follow causes in familiar ways; Hume notes that there is a chance that fire will not warm us and water not refresh us (1978, p.270).

Mumford and Anjum, in contrast, deny that a necessary connection is necessary to qualify something as a case of causation (2011, p.47). Their argument is that, if a necessitates b, then it must be a sufficient condition for it. Causation does not operate like this however: striking a match may cause it to light, but the striking is not a sufficient condition for it lighting, since one might strike the match and it not light. Mumford and Anjum anticipate the objection that only successful match strikes cause the lighting, but respond that this is a misuse of causation – ‘If someone were to say that, in some cases, being water necessitated being H₂O but in others did not, *prima facie* it would seem as if they did not understand the meaning of necessity’.

Another objection that Mumford and Anjum consider is that the truth of causal necessity would be shown if we could take note of every causal factor present in a causal process (2011, p.66). For example we could amend the statement ‘striking a match causes it to light’ by adding a list of the conditions needed to necessitate the truth of that statement. Their first

response to this objection is that some of those causal factors may need to be absences – such as the absence of a strong gust of wind in the match case – and that it is unclear how absences are supposed to work as causal factors. The second response is that one would need to include every relevant causal factor, a list so detailed and precise that it would become difficult to ever prove necessity. Even laboratory conditions would, Mumford and Anjum seem to suggest, never be truly repeatable. Thus every case would need to be taken as a single instance, and a general pattern can never be established.

Mumford and Anjum also espouse a process based theory of causation, rather than an event based one (2011, p.3). Their claim is that in a standard case of causation such as scales reacting to a weight placed upon them, it is not the event (the placing of the weight) that does the work but instead the properties (powers) that the weight possesses, in union with the powers of the weighing scales. Event ontology, they argue (2011, p.106), is flawed because either there must be a temporal gap between events, which would need to be crossed by some kind of mysterious link, or there is no temporal gap and the causal chain looks more like one long process. Under a process ontology: ‘Something will still be identifiable as cause, and something else as effect, but these will not be two events that must stand in some relation, as the standard account tells us. Rather, the cause will be depicted as merging into and becoming the effect through a natural process. (2011, p.107). Thus Mumford and Anjum adopt the dispositional modality, a theory of causation that rejects both necessitarianism and event ontology.

In summary the first difference between the two theories is that their rejection of any necessary connection is based on different grounds – Humean scepticism entails that one cannot justifiably believe in the necessary connection because no corresponding impression of it can be found, while Mumford and Anjum argue that a necessary connection is not an essential part of causation. The second is that Mumford and Anjum posit the dispositional modality as an explanation of causation while Hume claims that constant conjunction is the closest we can come to being sure of any causal link. The third is that Mumford and Anjum utilise a process ontology rather than the event ontology they ascribe to Hume, or at least to ‘followers of Hume’ (2011, p.2).

None of these differences are fatal to the idea that the Humean should understand character traits as mental powers. The first entails that there will always be a metaphysical difference between Humeans and Mumford and Anjum, but the difference is not material to the subject in hand; character traits as mental powers. The crucial point is that a Humean can incorporate powers into their ontology without needing to adopt necessitarianism, which would constitute a major alteration of Humean philosophy. A theory of powers that does not entail that powers necessitate their manifestations, such as that provided by Mumford and Anjum, can be adopted by the Humean without contradiction. The Humean can remain merely sceptical of the existence of a necessary connection for the reasons provided by Hume, they do not need to conceive of the necessary connection being inessential to causation as Mumford and Anjum do.

On the third difference, The question of event ontology as opposed to process ontology is not one that needs to be addressed by a Humean wanting to adopt aspects of powers theories in the way suggested here. Hume never puts forward a positive account of event ontology, and though his followers may have adopted that theory, there is nothing in process ontology that contradicts any crucial area of Hume's philosophy.

The second difference, between constant conjunction and the dispositional modality, is the most interesting. In general the Humean should be sceptical of necessity and the dispositional modality, for epistemological reasons, and this thesis will not suggest that Humean metaphysics needs to be altered outside of the realm of character traits. Nothing in this thesis has shown that there is an impression to ground belief in these concepts outside of character traits, nor has it suggested that one would not be necessary. However this thesis will argue, in section 8.1.3, that Humeans should accept the dispositional modality as an explanation of how character traits as powers operate. In short, Humean character traits currently rely on a notion of necessity, as Hume presents this as the only alternative to contingency in human action. This belief in necessity is contrary to much of Humean metaphysics, and the dispositional modality is a promising way of abandoning that reliance on necessity without having to adopt contingency.

In conclusion a Humean can understand character traits as mental powers without jettisoning any major philosophical point of Hume's. The advantages of such a position; that the explanatory and predictive uses of

powers are more in line with how Hume uses character traits in his writings, coupled with the empirical availability of powers as character traits, remain.

7.5: Conclusion

This chapter has laid out the argument from the *nisus*, an argument which seeks to prove that character traits in Hume's metaphysics are best understood as powers, based on the impression of mental straining these traits cause. It has been shown that Hume's objections do not impede such a limited view of powers, and that many features of Hume's psychology are still best viewed as summative dispositions, owing to the lack of empirical evidence for their being otherwise. The virtues of this position, that it can aid the Humean in resisting situationism, have been highlighted. Questions regarding the metaphysical nature, location and causal nature of character traits as powers have been addressed, to give a more complete picture of Hume's moral psychology once this reconceptualization is implemented.

Chapter 8: Resolving the tension

This thesis has been motivated by the need to resolve the tension between Hume's basic epistemological precepts and his brand of virtue ethics. It has been argued that conceiving of character traits, which are central to Hume's moral psychology, as powers rather than stimulus-response dispositions can aid greatly in resolving this tension, and that doing so does not subvert Humean empiricism in any significant way.

This chapter will analyse other key concepts in Humean philosophy: first it will be shown that Hume's moral psychology, which included a commitment to psychological determinism, is not strictly compatible with a belief in real causal powers, though the Humean can retain the opposition to the liberty of indifference that motivates the inclusion of determinism. Then central concepts of Humean virtue ethics, analysed throughout this thesis and all integral to character, will be revisited here with the aim of highlighting the benefits of, and solving the difficulties with, introducing powers to Hume's philosophy in this limited way. The ultimate conclusion will be that the introduction of powers, the plausibility of which was argued for in chapter six, does not have any major negative implications for Humean philosophy and has several merits.

8.1: Psychological determinism

The goal of this thesis has been to show that Humean ethics is in tension with Humean metaphysics, and that one notable source of this is that character traits are incongruous with rigid sceptical empiricism. The

solution offered has been that the Humean should conceive of character traits as being an expression of real causal powers. The previous chapter has shown how this does not contradict Hume's basic empiricism; the notions of ideas, impressions and the copy principle that are an essential part of Humean philosophy.

However it may be thought that the introduction of powers into Humean philosophy, even in the limited sense suggested, contradicts other essential elements of Hume's work. In particular the introduction of any kind of real causal powers into the metaphysics of character could be seen to jeopardise the thesis of psychological determinism which, as will be explained below, is an important element of Humean moral psychology. Accounts of powers such as the one offered by Mumford and Anjum, which purports to show that powers do not necessitate their effects, might spell trouble for the central claim of this thesis – there is little point in solving the initial tension identified above if this would cause tension with another important element of Humean moral psychology.

This section will begin by recapitulating the Humean argument against the liberty of indifference, which appears to entail a commitment to psychological determinism. It will be argued that the dichotomy presented by Hume (1978, p.407) – that the workings of the human mind must be governed either by necessity or by chance, is false and that a third way is available. This third way is the dispositional modality presented by Mumford and Anjum, a thesis which states that powers do not necessitate their effects but merely tend towards them. It will be argued that, if the Humean accepts powers as a justifiable addition to aspects of Hume's moral

psychology, then they should also accept the model of dispositional modality. This position has the benefits of cohering with other Humean beliefs about the nature of causation, and in making more sense of the link between character and blame.

The section will conclude that the Humean should accept this new model of causation within moral psychology, which fits well with other Humean ideas of causation, and which allows them to continue to reject pure contingency while accepting powers as the basis of character traits. This chapter will then go on to consider the implications of the use of powers for other aspects of Humean philosophy of mind outlined earlier in this thesis.

8.1.1: Hume's argument against the liberty of indifference

This section shall present Hume's arguments for psychological determinism and against the liberty of indifference. The aim is to show that the rejection of contingency is a central commitment of any Humean theory of moral psychology, and that an alternative must be realised in order to make sense of the theory. What will be identified as problematic is Hume's notion that the only alternative to determinism is blind chance – this offers a false dichotomy that Humeans need not accept.

One of Hume's major contributions to the free will debate was the distinction between liberty of spontaneity and liberty of indifference (1978, p.407). Liberty of spontaneity is equivalent with what is now referred to as liberty of action. This is simply the ability agents have to bring about change in their physical and mental states, to shake one's head or recall a memory.

When powers theorist E. J. Lowe discusses the ability to move one's arm in his analysis of the will as a power (2013, p.172), it is this liberty he is addressing. Hume grants that people have this liberty, within certain confines; evidently our bodies and mind need to be in good working order if they are to respond in the way we would like, and without external forces such as chains that will stifle movement.

It is liberty of indifference that Hume focuses his critique on. This can be defined as the ability to generate one's own thoughts, or the idea that any of one's thoughts or actions might come about without a cause – in Hume's words the liberty of indifference 'means a negation of necessity and causes' (1978, 407). Locke's account of the will as the only truly active power (Rickless, 2013), capable of causing without being caused, stands in direct opposition to Hume's view. Hume wrote that the feeling we have of authoring our own mental states is just a 'false sensation' (1978, p.408), an illusion that we can will more than nature allows us to. The distinction between the two kinds of liberty amounts to being able to act as we will – liberty of spontaneity - and being able to create our own will – liberty of indifference, only the former of which Hume takes to be a realistic proposition.

Hume's argument against the idea of liberty of indifference is based on the idea that our mental features operate in the same way as physical features, that each must have a cause and that like causes will always produce like effects (1978, p.401). Hume wrote that we could identify this by searching for a constant conjunction of mental states and physical actions, one that certainly does seem to exist in a given sphere of human action. In the physical world, the idea of constant conjunction could be expressed by

showing how a vase breaks each time it falls onto a hard surface from a height, because of the physical factors involved in the constitution of the vase, the floor and the height of the drop. Finding a constant conjunction between certain character traits and certain behaviours should be enough to convince us of a necessary link between the two just as we should believe in the necessity of cause and effect – which is to say, from a Humean perspective, that we are very well justified in believing in such a link, given the wealth of evidence in its favour, though speaking in a strictly philosophical sense we cannot know that the link exists.

Hume makes it clear throughout his analysis that he regards the mental world as being as much regulated by ‘necessity’ as the physical (1978, p.409). That is to say that the constant conjunction of cause and effect observed in the natural world can also be observed in the human mind, which justifies belief in the mind as predictable and deterministic, as the physical world appears to be. Hume often uses character traits to exemplify this necessity, writing that:

a spectator may commonly infer our actions from our motives and character; and even where he cannot, he concludes in general, that he might, were he perfectly acquainted with every circumstance of our situation and temper, and the most secret springs of our complexion and disposition (1978, p. 409)

A particular example is given; the obstinate gaoler (1978, p.406) whose obstinacy is as sure a bar to the escape of a prisoner as the walls and bars around them.

Hume further argues that the liberty of indifference has been denied by all serious thinkers, and gives several examples to show that common thought rejects it as well. Hume points to the consideration of what he calls 'moral evidence' (1978, p.404). When we are discussing an action, for instance a crime, Hume claims that people appeal not only to physical evidence, but to moral evidence, either to imply or conclude that an agent did indeed perform the action. Physical evidence might be that I was in Spain when a murder in Britain took place, a fact that at least strongly implies I did not directly murder the victim; otherwise there might be blood on my clothes that implies I did the deed.

Hume reasons that my character may absolve or implicate me in just as certain a way. If I am known to have a greedy character, and my wealthy relative passed away, then the proposition that I murdered them for their inheritance will be taken seriously. At the very least, one who has a grasping, duplicitous nature is more likely to have committed the crime than someone who is kind and generous. Clearly this kind of evidence will not hold in the rigor of the law courts, but Hume is convincing when he states that people tend to act as we expect, unless they have a very good reason for acting otherwise.

Hume's writings on moral evidence and the predictability of character do more than just show the widespread intuitiveness of psychological determinism; they also reveal the extent to which his conception of character relies on the truth of that thesis, despite the fact that this reliance on determinism is in conflict with his general metaphysical stance. Hume concludes that the only alternative to determinism is chance, or as he puts

it: 'necessity makes an essential part of causation; and consequently liberty, by removing necessity, removes also causes, and is the very same thing with chance' (1978, p.407). If one's thoughts and will were not governed by necessity, then their actions and the desires leading to them would be random, and we would be unlikely to display any stable kind of character trait - if they did, it would only be through chance that they did so, and there would be no guarantee that their character tomorrow would resemble their character today. Moral evidence would thus become irrelevant, as the kind and generous person would be as likely to murder their rich relative as the greedy duplicitous one, and the obstinate gaoler as likely to relent and release their prisoner as not. Given that experience tells us that people do not act in this arbitrary way, Hume concludes that there must be a necessary link between character traits and their behavioural expressions.

Thus Hume presents psychological determinism as a key commitment of any moral psychology, which, given the options he presents, is reasonable – psychological determinism is a better explanation of the stability human behaviour than pure chance. This thesis has so far advanced the idea that Humeans, who are necessarily committed to the epistemology of impressions and ideas, can and should accept real causal powers as the basis of stable human behaviour. These two ideas, that Humeans should adopt powers into their moral psychology and that they should retain the concept of psychological determinism, appear to be in conflict, as modern powers-based theories often place themselves against determinism. The rest of this chapter will argue that Hume was wrong to put the choice between necessity

and contingency in such stark terms, that an alternative option is available, and that it can be taken up without jettisoning any core Humean precepts.

8.1.2: The dispositional modality

The possibility of an option between pure necessity and pure contingency has been put forward by Mumford and Anjum (2011, p.175) and will be analysed here. The following sections will argue that the Humean should accept this dispositional modality as a model of how powers cause their effects, and that doing so allows the introduction of powers into Humean moral psychology while not having to turn to pure chance, which Hume proposed was the only alternative.

Mumford and Anjum refer to the dichotomy presented by Hume – between a world of necessity and a world of contingency – as modal dualism (2015, p.2), which offers only two modes of causation. Hume is interpreted as a proponent of contingency (2011, p.49) given his rejection of a necessary connexion (1978, p.165). This may seem at odds with Hume’s rejection of chance in favour of necessity in the passages quoted above (1978, p.407); a reflection of the central tension this thesis aims to solve. When dealing with metaphysics, where the object of study is causation, Hume is clear that we can have no sensible idea of necessity, as we have no impression of it, and that constant conjunction is the only justification we have for believing that future instances of cause and effect will mirror the past. However when dealing with ethics, when Hume is trying to explain human behaviour, he argues that necessity is the only explanation for stable human behaviour.

This may be a loose use of language; Hume is using 'necessity' to mean that we can almost be sure, given the constant conjunction of characters and behaviours, that future instances of character and behaviour will mirror the past. However, given that contingency is an unsatisfactory explanation for stable human behaviour, as explained above, and Humean empiricism cannot allow one to truly commit to necessity, a third option should be attractive to the Humean.

Mumford and Anjum note that several powers theorists accept this duality of Hume's while trying to refute him (2011, p.51); according to them theorists such as Shalkowski provide good arguments against contingency, only to fall into Hume's erroneous dualist way of thinking and advance a necessitarian view as the only alternative. Mumford and Anjum reject the duality offered by Hume and espouse a third way, that causation can be understood in terms of dispositional modality, a middle ground between contingency and necessity. What this means is that a cause does not lead to its effect by pure chance, nor does the cause necessarily entail the effect, instead the cause tends towards its effect. Take a classic example of causation, a white billiard ball striking a red, followed by the red speeding away. Hume and Mumford and Anjum all reject that there is a necessary link between the events, but for different reasons. Hume claims that no such link is perceivable, and therefore we cannot have a proper idea of it. Mumford and Anjum instead argue that, for there to be a necessary link between A and B then the two must always be found together (2015, p.5). But the white ball's striking the red may not always be followed by the red speeding away – someone may hold it still, or the white may be stuck to the

table, or the red might break if hit hard enough. Therefore there is no reason to believe in a necessary connection between the two.

Mumford and Anjum go on to reject contingency because we experience enough stability in the world to convince us that there must be some kind of link between cause and effect (2011, p.193) – a white ball striking a red has never caused it to vanish into a puff of smoke so there is no reason to think that it ever will. With contingency and necessity rejected, a third way is all that remains; Mumford and Anjum call this alternative the dispositional modality. They conclude that causes tend towards their effects; the white ball striking the red does not necessitate the reds speeding away, but the red does tend to speed away when struck in our experience, giving us good reason to think that it will in the next instance. A range of outcomes are possible given the causal powers of the two balls, just as a range of outcomes are impossible for the same reasons (2015, p.6).

8.1.3: Two models of causation

This section will analyse an objection from Mackie (2014) with the aim of clarifying Mumford and Anjum's position and dispelling a potential objection to it. Their response will invoke the idea that there is a distinction to be made between determinism in general and causal determinism, the latter of which they reject. It is this model of causation that will solve the central problem of this chapter, and will allow Humeans to accept powers into their moral psychology without abandoning any important principle of their philosophy.

The account of dispositional modality rests on the principle of interference, which states that a cause can never necessarily lead to its effect because interferers might come into play and prevent the outcome (Mumford and Anjum, 2014, p.22). In the case of a match being struck, its being struck is not a sufficient condition for its manifesting the power to light because a gust might blow, or rain might pour, or some other interferer will prevent the manifestation. Therefore we can only say with certainty that the match will tend to light, rather than it is necessarily going to light.

The central problem Mackie identifies with this argument is that the definition of determinism that Mumford and Anjum are arguing against is not one shared by many determinists. She argues that they believe that ‘a match being struck necessarily causes it to light’ is equivalent with ‘a match being struck necessarily causes it to light *no matter what else happens*’, which is not an entailment many would endorse or defend. Instead when a determinist claims that ‘striking a match will cause it to light’ they are bracketing the absence of interferers in their statement; they are also implying the presence of enablers such as oxygen. The presence or absence of interferers, Mackie continues, is just as determinate as everything else (2014, p.596). When I am about to light a match I expect it to do so, ignorant of whether or not a gust of wind will prevent me from doing so, but there is still a fact to the matter: the wind will either blow or not blow depending on past material circumstances.

Mumford and Anjum can respond to this objection by denying the truth of causal determinism while leaving the question of determinism in general open. Causal determinism is defined as ‘the vehicle by which determinism

does its business (2011, p.75)', as opposed to determinism in general, the 'core idea' of which is 'fixity of the future by the past'. Mumford and Anjum hold that this broader notion of determinism is compatible with their dispositional modality, but do not commit to the truth of it (2015, p.6). Thus Mackie's objection fails to land – returning to Mackie's example of a match being struck, modal dispositionalists can hold that the causes at play, the match, the wind etc., can cause the effect of the match's lighting or not lighting. However they do not need to accept that these factors necessitate the outcome, owing to the fact that the powers involved are themselves indeterministic. Thus two identical instances of striking a match may produce different outcomes, and this indeterminacy does not contradict the thesis of determinism as a whole (2011, p.76).

Hume's claim that the link between character and behaviour is a necessary one can thus be re-examined. The Humean need not believe that causal determinism is at work here, that character traits can lead only and necessarily to one behavioural expression. Hume suggests this is the case in his ethical writings, but only because he rejected contingency and could not see another option. With one available, the Humean is free to reject contingency, as Hume does in his ethical writings, but also to reject necessity, as Hume does in his metaphysical writings, and argue that character traits can tend towards their outcomes without necessitating them. This approach allows the Humean to posit real causal powers as the basis for the stability of character traits, while retaining the denial of contingency that is integral to Humean moral psychology.

8.1.4: Necessity and blame

The above reconceptualization also has an impact on Hume's theory of blame, which is presented as resting upon psychological determinism. This section will show that necessity is not essential to Hume's theory of blame, that it is the rejection of contingency which is crucial, and that the dispositional modality can play this role just as well as necessity. Realising character traits as real causal powers rather than summative dispositions in fact gives one a far more solid foundation for theories of praise and blame.

When writing on psychological determinism, Hume notes that a world of psychological contingency would mean never being truly able to blame anyone for their actions. In a world where virtues could lead to vicious action as easily as virtuous, where the necessary link between good traits and good behaviour is gone, then virtuous action is a roll of the dice. Kindness might lead one to act cruelly, making it impossible to judge a character based on their actions. Hume concludes that 'tis only on the principles of necessity, that a person acquires any merit or demerit from their actions' (1978, p.411).

Certainly the rejection of contingency in this matter is something modern Humeans should want to emulate. The argument that a world of contingency would mean that one is unable to justly praise or blame a person's character based on behaviour is sound, where Hume errs, as explained in the preceding sections, is in assuming that the rejection of contingency leaves one with only necessity. As this is not the case, a more attractive avenue is open for the Humean. Both necessity and contingency

can be rejected in favour of the dispositional modality, which can explain the causal links between traits and action without having to resort to insensible necessity – the ‘constancy’ ‘betwixt motives and actions’ (1978, p.404) is still explicable. Character traits tend towards their behavioural outcomes, which experience bears out. Even the sturdiest of character traits will sometimes fail to affect the behaviour of the bearer; this is a position in fact more in line with Humean epistemology than Hume’s own unwarranted confidence in certain characters to always produce certain actions. Hume’s example of a man that ‘doubts not the obedience of his servants’ (1978, p.405) is, for Hume, only explicable through necessity, though in fact believing that the servants’ characters tend towards obedient behaviour most of the time fits in far better with the general Humean project of scepticism – the man *should* have some doubt about whether the servants’ trait of obedience will be efficacious when tested.

Hume further observes that one is only truly responsible for one’s actions, and therefore is only the appropriate target of blame or praise for their actions, when those actions are the result of one’s character, described by Hume as something ‘durable and constant’ (1978, p.411), rather than of any passing fancy. Otherwise ‘The action itself may be blameable...But the person is not responsible for it’ (1978, p.411). This notion, that character traits are the locus of ethical investigation, is integral to Humean ethics. However, without a coherent theory of character to back this up, the claim is hard to interpret. We know that Hume theorised that character was the source of all moral approbation, and that therefore character was something

more than the self, but the exact nature of character, its durability and constancy in particular, remained mysterious.

Following on from the arguments advanced here, we are at least closer to solving the problem. The durable and constant elements of one's character, character traits, should be understood as real causal powers that are the appropriate source of praise or blame in a way that summative dispositions could not be.

8.1.5: Conclusion

These sections have shown that the Humean does not need to choose between necessity and chance. The dispositional modality offers an attractive alternative that in fact sits very well with Humean empiricism; necessity is insensible but chance cannot explain the stability of character traits and their link to certain behaviours, while our experience of characters is that they tend to exhibit certain behaviours, but may not. Thus the Humean can accept powers, which do a much better job of explaining our certainty of how people will act than the summative model, they are also not resigned to pure contingency. Doing so does not jeopardise Humean notions of moral evaluation, but bolster them, giving praise and blame an appropriate target.

8.2: Humean moral psychology and powers

The introduction of powers into Humean philosophy has ramifications outside of the question of psychological determinism. Many concepts that

are central to Humean ideas of character could and should be analysed anew in light of this change. This section will recap Hume's basic epistemology and the strict limitations on knowledge that the methodology entails. The arguments for importing aspects of powers-based metaphysics will be summarised, to support the alterations to be suggested in the remainder of the chapter. The two primary criteria for the reconceptualization of certain Humean psychological factors will be the empirical – that the power of the factor is sensible, as with the *nisus* – and the analytic – that the factor cannot be captured by a conditional analysis, meaning it cannot be understood as a summative disposition. A related concern will be whether the factor in question can possess the explanatory and predictive power if it is understood as a summative disposition. Character traits cannot, despite their importance in Humean ethics, which is one of the motivators for their reconceptualization. It may be that other Humean elements in moral psychology are also incapable of doing the work they need to do if understood as summative dispositions.

The one rule that is central to Hume's epistemology is that all knowledge must come from experience; it is this rule which establishes Hume as an empiricist, and it grounds all other aspects of Hume's work. His particular presentation of the rule, which distinguishes Hume from his empiricist contemporaries, is that every idea we possess must come from an impression (1978, p.4). Impressions are sensory data, they comprise everything we see, hear and feel, while ideas are their fainter copies, which we can combine, divide, and introspect upon at will. If we possess an idea that does not emanate from a corresponding impression, argues Hume, we

should doubt that the idea represents anything real at all. No philosophy can consider itself Humean without adhering to this principle, as every other part of Hume's philosophy relies upon it; it is intended as 'the foundation of his philosophy' (Noonan, 1999, p.51). The resulting scepticism is thus an integral part of Humean thought, and is laudable for keeping Humean philosophy grounded in that which is concrete and sensible. However it also imposes strict limitations on the kind of concept Humeans can reasonably discuss.

This has, in modern philosophy, led to Humean empiricism being positioned as the enemy of neo-Aristotelian powers theories (Groff, 2014, p.422). Powers, both supporters and detractors of Hume have claimed, are not represented by ideas that can be traced to a sensory impression, and so they must be excluded from Humean philosophy. This methodology seems to lend itself to the belief that dispositions can be understood only summatively, it might even be thought to restrict one to thinking this way about dispositions necessarily. Summative dispositions, as presented in the standard conditional analysis (SCA), are merely conditional propositions rather than objects in themselves, as Aristotelian powers are. To say that an object is disposed to break is simply to say that it will break under the appropriate circumstances, with no comment necessarily made as to the root cause of this behaviour. To say that an object has the power to break, however, is to say that it currently, in its non-broken state, possesses a quality that will cause it to break in the presence of other causes. The object's fragility is underwritten by a metaphysical object in the powers picture, while the dispositional picture involves only general rules learnt

from experience, with no commitments as to the fundamental nature of things.

As a result Hume rejected powers, and the metaphysics of summative dispositions to explain the counterfactual behaviour of objects appears an obvious alternative. This account has difficulties, especially if it is supposed to be the metaphysical support to Humean virtue ethics, a crucial aspect of which are the virtues, which are a kind of character trait. These traits resist being viewed summatively, while appearing closer to Aristotelian powers in some ways. The SCA, the most promising yet still controversial analysis of summative dispositions, clearly struggles to deal with character traits as it is unclear exactly what their appropriate manifestation conditions. If a person is disposed to be ambitious, what is the appropriate stimulus for them to act or think ambitiously? It seems traits such as this do not have to wait for a stimulus to trigger them, rather they constantly exert a pressure on the bearer to behave in a certain way. Hume's empiricism appears to entail a metaphysics of summative dispositions, but this metaphysics does not support his virtue ethics. Conceiving of character traits as powers; as real, durable objects, elements of one's psychological make-up, solves this issues.

In the preceding chapters the argument from the nisus has been proposed as a way to retain Hume's basic epistemology and accept real causal powers without contradiction, at least in a limited sphere – that of character traits. Certain powers are empirically accessible, as Hume concedes (1975, p.67), and so can be integrated into his philosophy without contradiction. Other important aspects of Hume's work, such as his

rejection of contingency in human behaviour, are not necessarily undermined by this approach, as argued earlier in this chapter. There are many benefits to the approach; for one the tension that motivated this thesis is defused, and Hume's epistemology and virtue ethics can now exist more comfortably together. This new integrated system has several advantages over the old, for example the focus on the phenomenological, rather than third-person, accessibility of character traits, which gives the Humean a unique defence against situationist objections to virtue ethics. It can also elucidate much of the rest of Humean philosophy, including the ethical concepts introduced in chapter 3, and the will. The following sections will move from potential problems this thesis has raised to the potential benefits.

8.2.1: The will

Explaining the metaphysics of Humean character traits has been the primary aim of this thesis, but the proposed solution has benefits in many related areas of Humean moral psychology. This section will argue that the will should be understood as a power in a Humean ontology, drawing on the same justifications used in the case of character traits. These are the pragmatic reasons for the change (that summative dispositions cannot do the explanatory and predictive work that powers do, and that character traits defy the SCA) and an empirical justification for it (powers are empirically available to us through *nisus* like straining).

First to explain how the will should be conceived in a Humean framework. Hume writes that the will is 'nothing but *the internal impression*

we feel and are conscious of, when we knowingly give rise to any new motion of our body, or new perception of our mind' (1978, p.399). He adds that it is a simple impression, like pride or love, that cannot be broken down into any smaller constituent parts. The idea that the will is an impression heavily implies that it has a passive role, as does the thought that it is something we 'feel and are conscious of' rather than something we generate or control. As I might perform an action or think a thought angrily or generously, so might I perform or think wilfully; or not in the case of a sneeze or a spasm. My action or thought belongs to the will because, as the action or thought occurs, I sense an impression of intentionality, of wanting the action or thought to occur in the way it does. The will is not necessarily prior in this story, it is merely attendant. This conception of the will follows directly from Hume's basic empiricism, as it is all that can be empirically verified, and so should be part of modern Humean thought.

The important question is whether the argument from the *nisus*, which intended to show that Humean character traits are best thought of as powers rather than dispositions, can be applied to the will. Dicker (1998, p.104) uses the example of straining to work out a difficult puzzle in his discussion of the *nisus*, while struggling to remember something provides an example of the kind of resistance encountered by the faculty of memory. Does the analogy hold in the case of the will?

At first, instances of the will being felt as a force or exertion might be easy to find; we feel our wills acting upon us in the same way that we feel our character traits motivating us to act. However this conception of the will is a far broader one than Hume is working with. He described the will as the

impression accompanying intentional action, rather than an originator or motivator of thoughts. The will is simply the feeling that one means to do what one is doing, we do perform any particular action because we feel the will insisting that we do so. Character traits do this work in a Humean framework, and any apparent example of one's will being conflicted will actually be a conflict between two character traits, or passions, which are motivating the subject to perform incompatible actions. The Humean will is only felt when one is actually performing an action. Therefore it cannot be in conflict as one cannot perform incompatible actions. This means that subjects are not conscious of impressions of the will exerting force in these examples.

However instances of the will's being frustrated do provide evidence of the kind of straining that serves as good empirical evidence of the existence of powers. For example if one very much wants to board a plane, but cannot because they do not have the money, then they will likely feel the strain of the will impressing on them, just as one might strain to work out a puzzle or remember something unpleasant. As Hume relays in the *nisus* footnote, it is this frustration and straining which provide the impression that justifies the idea of real causal powers. Thus the empirical evidence for reconceiving the will as a power is there, just as with character traits – the power of the will is known to us through the impressions we sense when the will is frustrated.

The analytic reason for the change is also present in the case of the will, as it is with character traits. The SCA, realised as 'An object is disposed to *M* when *C* iff it would *M* if it were the case that *C*', where *M* means manifest

and *C* represents suitable conditions, is inapplicable to the will. *C* is difficult to define if *M* represents the will, as the suitable conditions for the wills' manifesting will be incredibly diverse in different cases. Thus the will is difficult to understand as a simple stimulus-response disposition; it instead appears to be something that occurs within an agent and asserts itself as character traits do.

8.2.2: Passions

Other crucial factors of Humean theories of character can be empirically supported in the same way as virtues and vices. The passions, listed by Hume as 'grief, hope, fear' (1978, p.276) and analysed in sections 1.3 and 3.2, are similar enough that the same form of argument can be utilised to support their being viewed as powers rather than dispositions.

Hume writes that 'on no occasion is it necessary to suppose, that they have existed and operated, when they are not perceiv'd' (1978, p.195), though this leaves his moral psychology severely lacking. Supposing that passions do not exist and operate when not being perceived commits one to thinking of them in terms of summative dispositions which, as with character traits, conditional analyses struggle to capture. The appropriate stimulus for grief or hope is difficult to ascertain, suggesting that passions have a more active role rather than a passive, reactive role.

A person can know if they have the power to feel hope via introspection, through the hopeful feelings that we experience when hoping that compel us to act or speak in certain characteristic ways. Similarly a character trait

such as ambition might strike us with great force, and can overwhelm other feelings or desires. As with hope, these ambitious thoughts and feelings are the proof that one bears the power of ambition. If occasional strainings can, via Humean empiricism, provide evidence of the durable power of the nisus, then occasional passionate impressions should be adequate evidence that there is an active power behind them.

8.2.3: Taste and sympathy

Two more important Humean features can be re-evaluated in this light: sympathy and moral taste. Both are intimately linked with ideas of character, and both could be seen as problematic in the Humean account as it is they, like virtues, are difficult to sense when not active. Section 3.3 argued that taste is a distinct faculty in Humean psychology, one that is a long-term stable disposition to act in a certain way that is thus in need of explanation. Section 3.4 made the point that sympathy is similar in this respect. Both are intimately linked with ideas of character, as both can be understood as character traits in their own right. Taste is not morally evaluable in the same way that virtues and vices are, but it remains a stable disposition to behave in a certain way and is expressive of a person's character in much the same way.

Like the virtues and the passions, sympathy and taste can be recast as powers in a Humean ontology without contradiction and with some benefit. Neither fits neatly into the SCA as it is difficult to determine their appropriate stimuli. Both are felt strongly by their possessor, indicating that

they are causal properties rather than dispositions passively waiting to be triggered. Sympathy, as defined by Hume, has the effect of rendering other peoples' 'sentiments intimately present to us' (1978, p.320). Thus it can be sensed as a vivid impression, such as when we feel a pang of emotion in sympathy with another person, in the same way that the nisus of strong physical or mental endeavour can be sensed as a vivid impression. It is this strong impression that provides the empirical justification for thinking of sympathy as a mental power. Conversely there is no such impression associated with the copy principle – we are sensible of the effects of the mechanism causing ideas to be copied from corresponding impressions, in that we are conscious of the results of the process, but there is nothing equivalent to the straining described by Hume that is the basis for the idea of a power.

The faculty of taste causes a similar impression, which gives good justification for thinking of it as a power rather than a summative disposition. In terms of aesthetic taste, the beauty of an object can strike us vividly and with force, and the same is true for moral judgements that are the result of our moral taste at work. Recasting sympathy, moral taste and aesthetic taste as powers that act upon an agent and express themselves in impressions allows modern Humeans to explain the stability of these traits; how the faculties operate in similar ways across a person's lifetime. One's moral or aesthetic taste, or tendency to feel sympathy with others, is stable enough to suggest that a durable property is the root cause. This stability was not explained by Hume, and can not be explained by a purely summative model of dispositions. Realising the faculties as real causal

powers explains the stability, and thus, when justified by the impressions of these powers at work, fleshes out the Humean picture.

8.2.4: Personality traits

Personality traits, as referenced in section 4.5, also benefit from being understood in this new light. Distinguished from character traits by the fact that one is not morally responsible for them but otherwise very similar in nature, personality traits include sexuality, neuroses, and the extent to which one is or is not an extrovert.

Personality traits are empirically available in the same way as virtues and vices, as each is experienced as a strong motivation to act in a particular way. Even more so than virtues and passions, it is difficult to capture personality traits with the SCA if they are taken as stimulus-response dispositions. Sexuality in particular is not experienced as a response waiting to be triggered, but an ever present causal facet of our psychologies. They are therefore more accurately construed as powers rather than summative dispositions, with the same explanatory and predictive benefits.

8.2.5: The copy principle and other mental operations

Several mental operations that play important roles in Hume's philosophy are however still best viewed as dispositions, either because the motivation to reconceive them is absent or because the empirical evidence is unavailable. This is not necessarily to say that they are not underwritten by

real causal powers, only that Humean scepticism suggests that one refrain from making deep metaphysical claims about their nature. The copy principle is ensured by one such mental operation, as argued more fully in section 6.4. As discussed in chapter 2, Hume's strong belief in the copy principle – the principle that all simple ideas are derived from simple impressions which they represent – is an indication of his belief in a mind as distinct from the self. This is due to the fact that only the effects of the function, the ideas arising in our minds, are empirically accessible while the root cause of the effects is not. We feel no mental straining as the result of the copy principle working upon our minds, nor does the principle defy the SCA; a person is subject to the copy principle if their simple ideas are generated by simple impressions. Again, there may well be a real causal power that causes our simple impressions to be copied into simple ideas, but without empirical evidence or pragmatic motivation for a belief in such a power the Humean should reject it.

Similarly, there is little argument to be made for reconceptualising basic mental operations. If it is accepted that there is a distinction between the mind and the self in Hume then the mind will be said to have certain abilities: the ability to create secondary impressions and ideas, the ability to form and examine memories, the ability to reason. Intellectual abilities such as numeracy and literacy fall into the same camp, as do skills, such as the ability to drive a car. Although these might standardly be thought of as powers in neo-Aristotelian accounts, there is no impetus to do the same in the Humean account, in part because an 'ability to x' intrinsically implies a hypothetical, as it refers to a capacity to act in a certain way under certain

circumstances, and so is more amenable to the dispositional account. None of the above are felt in the way that passions and virtues are felt; reason Hume specifically states ‘can never be a motive to any action of the will’ (1978, p.413) and so is unlikely to impress upon us in the same way as them, so the evidence for their existence as powers is absent. The motivation to no longer conceive of them as dispositions does not apply to the ability to create impressions and ideas, as stated above, though it may for memory and reason.

Is there reason to think of reason as a power in Hume’s ontology? The empirical proof, the straining associated with other powers, is absent, but does it defy the SCA in a way that makes it problematic for the Humean? If reason is taken to be a stimulus-response disposition of agents, then it is difficult to say what exactly its appropriate trigger is. One solution is to slice reason into thinner dispositions, as Doris attempts to do with notions of character into localised dispositions. Instead of an agent having one ability to reason, it would be beneficial to conceive of them having a multitude of reasoning abilities, each with their own specific manifestation conditions. This would make reason less problematic for the Humean, and would render any other kind of reconceptualisation unnecessary and unsupported.

8.2.6: Belief

The final case that will be analysed is that of belief, which, due to its odd nature in Hume’s psychology, somewhat defies categorisation. Is the Humean concept of belief best understood as a summative disposition or a

power? It is necessary here to distinguish between belief, which is a quality of ideas, and the ability to believe, which is an ability of individuals. Belief, the vivacity associated with the ideas we believe in, certainly seems to be experienced as a straining, equivalent to the straining one experiences when trying to solve a difficult puzzle. Passionate or motivating beliefs operate similarly to virtues in this respect, and so there is a case to be made for belief being conceptualised as a power as this thesis has suggested character traits including ambition and greed should be. Belief also fails to adhere to the SCA, which indicates that its metaphysical nature is more complex than a stimulus-response disposition – it may instead be a power that is constantly tending towards its manifestation.

On the other hand, the ability to believe appears more like the mechanism that causes the copy principle to hold true. Its effects are empirically accessible, in that beliefs are known to us via introspection, but the faculty itself never enters into the bundle of impressions and ideas in the way that character traits do. The ability to believe does defy being easily captured by the SCA, in that the appropriate stimulus that should entail the manifestation of the ability is hard to define, meaning that Humeans that wish to explain the ability to believe in terms of stimulus-response dispositions might encounter problems. But without the additional empirical justification present in the case of character traits, it cannot be concluded that the ability to believe should be viewed as a power in a Humean framework.

8.3: Solving the problem of character

The problem of character, identified in section 1.1, concerns the relation between Hume's bundle theory of the self and his theory of moral character. The issue is that it is unclear how the self, which is a bundle of impressions and thus subject to constant flux, is related to character, which is composed of stable traits.

The bundle theory of the self is one of the most famous aspects of Hume's work, and it illustrates his empiricist methodology well. When faced with the question 'what is the self?' Hume sought to examine the subject, to introspect upon his own self and observe its constitution. He concluded that the self is no more than a bundle of impressions felt by the subject at any one time; that it is entirely constituted by our sense data and reflections on that sense data. There is no essence of the self beyond this, meaning that selves change rapidly and constantly.

The concept of character might be thought to be intimately linked to notions of the self, but Hume treats them as separate topics of enquiry. His analysis of character is entirely concerned with its role in ethics, in which it predicts behaviour and grounds moral responsibility, with little attention paid to its metaphysical make-up. He does describe character traits as 'durable' (1978, p.575), which puts them at odds with the fleeting nature of impressions in the self. Reading Hume as holding a metaphysics of summative dispositions about character traits entails that these durable, explanatorily powerful aspects of his ethics had no underlying metaphysical

cause, and that his talk of character traits is just talk of the constant conjunction of an agent's past behaviour and future behaviour.

However, if character traits are conceived of as powers, then a more metaphysically robust account can be given. We can turn to powers, to durable, real objects, to explain and predict virtuous or vicious behaviour in agents, rather than relying on uncertain inferences from past to future behaviour. The central problem of how character relates to the self is also elucidated. As proposed in chapter two, Hume's psychology implicitly relies on a mind that is distinct from the self, one which can serve as the locus for mental powers. These powers manifest in the self as secondary impressions, the way we feel about and interpret the world around us. Mental powers also ground the stability that is evident in the self; they explain why one's secondary impressions do tend to resemble one another, and why one does not wake up every morning acting and thinking in radically different ways.

8.4: Conclusion

This chapter aimed to show the benefits that can come from understanding the dispositions that ground Humean character traits as real causal powers. Hume made the point, in his note detailing the argument from the *nisus*, that one could not infer from that to powers in the physical world. The use of powers suggested by this thesis does not contradict this: this Humean view does not postulate real physical powers, only psychological ones. It has been argued that one cannot infer from certain kinds of mental powers to others, if there is no motivation to do so or evidence to warrant such a

change. The mechanism behind the copy principle, the faculties of reason and memory, and certain abilities, are all still best viewed as summative dispositions. Other factors, such as the passions, personality traits and taste, do warrant a reconception. One does not have to hold to an ontology entirely comprised of dispositions or entirely of powers, as Bird has previously argued (2013, p.25). Those aspects of Hume's philosophy that are currently problematic, and which could be improved by being conceived of as powers rather than dispositions, should be, as long as this does not contravene the basic tenets of Humean philosophy, particularly the commitment to empiricism. Ultimately one can hold a recognisably Humean epistemology, metaphysics and ethics, as long as character traits are given a robust metaphysical underpinning that can explain why they remain stable across time.

Bibliography

- Abramson, K., 2008. Sympathy and Hume's Spectator-centered Theory of Virtue. In E. S. Radcliffe, ed. 2006. *A Companion to Hume*. Malden, MA: Blackwell Publishing pp.240-256.
- Anjum, R. L., Lie, S. N. A., Mumford, S., 2013. Dispositions and Ethics. In R. Groff and John Greco, eds. 2006. *Powers and Capacities in Philosophy: The New Aristotelianism*. New York: Routledge pp.231-247.
- Alanen, L., 2006. Powers and Mechanisms of the Passions. In S. Traiger, ed. 2006. *The Blackwell Guide to Hume's Treatise*. Oxford: Blackwell pp.179-198.
- Alfano, M., 2013. *Character as Moral Fiction*. Cambridge: Cambridge University Press.
- Alvarez, M., 2013. Agency and Two-way Powers. *Proceedings of the Aristotelian Society*, 113 (1). pp.101-121.
- Aristotle, 1956. *Aristotle's Metaphysics*. Edited and translated by John Warrington. London: J. M. Dent and Sons.
- Aristotle, 2002. *Nicomachean Ethics*. Translated by Christopher Rowe. Introduced by Sarah Broadie. Oxford: Oxford University Press, 2002.
- Ayer, A. J., 1954. *Philosophical Essays*. London: Macmillan.
- Baier, A., 1991. *A Progress of Sentiments: Reflections on Hume's Treatise*. Cambridge, Mass: Harvard University Press.
- Baier, A., 2008. *Death and Character*. Cambridge, Mass: Harvard University Press.
- Baillie, J., 2000. *Routledge Philosophy Guidebook to Hume on Morality*. London, New York: Routledge.

- Banicki, K., 2017. The Character-Personality Distinction: An Historical, Conceptual and Functional Investigation. *Theory and Psychology*, 27 (1), pp.50-68.
- Bègue, L., Beauvois, J., Courbet, D., Oberlé, D., Lepage, J. and Duke, A. A., 2014. Personality Predicts Obedience in a Milgram Paradigm. *Journal of Personality*, 83 (3), pp.299-306.
- Bhaskar, R., 2008. *A Realist Theory of Science*. London, New York: Routledge.
- Bird, A., 2013. Limitations of Power. In R. Groff and John Greco, eds. 2006. *Powers and Capacities in Philosophy: The New Aristotelianism*. New York: Routledge, pp.25-47.
- Blackburn, S., 2008. *How to Read Hume*. London: Granta Publishing.
- Boccaccini, F. and Marmodoro, A., 2017. Powers, Abilities and Skills in Early Modern Philosophy. *British Journal for the History of Philosophy*, 25, pp.435-442.
- Brown, C. R., 2008. Hume on Moral Rationalism, Sentimentalism, and Sympathy. In E. S. Radcliffe, ed. 2006. *A Companion to Hume*. Malden, MA: Blackwell Publishing pp.219-239.
- Cohon, R., 2006. Hume's Natural and Artificial Virtues. In S. Traiger, ed. 2006. *The Blackwell Guide to Hume's Treatise*. Oxford: Blackwell Publishing ltd. pp.256-275.
- Deleuze, G., 1991. *Empiricism and Subjectivity*. New York: Colombia University Press.
- Dicker, G., 1998. *Hume's Epistemology and Metaphysics*. London: New York: Routledge.

- Doris, J., 2002. *Lack of Character: Personality and Moral Behaviour*.
Cambridge: Cambridge University Press.
- Ellis, B., 2010. Causal Powers and Categorical Properties. In A. Marmodoro,
ed. 2010. *The Metaphysics of Powers*. New York, London: Routledge.
Pp.133-142.
- Epstein, S., 1979. Explorations in Personality Today and Tomorrow: A
Tribute to Henry A. Murray. *American Psychologist*, 34, pp.649-653.
- Foot, P., 1978. *Virtues and Vices*. Oxford: Blackwell.
- Frankfurt, H., 1971. Freedom of the Will and the Concept of a Person. *The
Journal of Philosophy*, 68 (1), pp.5-20.
- Gingerich, A. C. and Sullivan, M. C., 2013. Claiming Hidden Memories as
One's Own: A Review of Inadvertent Plagiarism. *Journal of Cognitive
Psychology*, 25 (8), pp.903-916.
- Groff, R., 2013a. Introduction. In R. Groff and John Greco, eds. 2006.
Powers and Capacities in Philosophy: The New Aristotelianism.
Routledge pp.1-6.
- Groff, R., 2013b. Whose Powers? Which Agency? In R. Groff and John
Greco, eds. 2006. *Powers and Capacities in Philosophy: The New
Aristotelianism*. New York: Routledge pp.207-228.
- Groff, R., 2014. After Hume: Realism About Powers in Contemporary
Analytic Metaphysics. *Journal of Critical Realism*, 13 (4), pp.422-442.
- Harman, G., 1999. Moral Philosophy Meets Social Psychology: Virtue Ethics
and the Fundamental Attribution Error. *Proceedings of the Aristotelian
Society*, 99. pp.315-331.

- Haslam, S. A., Reicher, S. D., Millard, K., McDonald, R., 2014. Happy to Have Been of Service. *British Journal of Social Psychology* 54 (1). pp.55-83.
- Heil, J., 2003. *From an Ontological Point of View*. Oxford: Oxford University Press.
- Hume, D., 1854. The Sceptic. In *The Philosophical Works of David Hume*. Boston: Little, Brown and Company, pp.174-198.
- Hume, D., 1975. *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals*. Edited by L. A. Selby-Bigge. Oxford: Oxford University Press.
- Hume, D., 1978. *A Treatise of Human Nature*. Edited by L. A. Selby-Bigge and P. H. Nidditch. Oxford: Oxford University Press.
- Hume, D., 1987. *Essays: Moral, Political, and Literary*. Edited by Eugene F. Miller. Indianapolis, Ind.: Liberty Classics.
- Hume, D., 1995. *Four Dissertations*. Introduced by John Immerwahr. Bristol: Thoemmes Press.
- Jaworski, W., 2013. Powers, Structures and Minds. In R. Groff and John Greco, eds. 2006. *Powers and Capacities in Philosophy: The New Aristotelianism*. New York: Routledge pp.145-171.
- Korsgaard, C. M., 1999. The General Point of View: Love and Moral Approval in Hume's Ethics. *Hume Studies*, 25 (1-2), pp.1-39.
- Kupperman, J., 1991. *Character*. Oxford: Oxford University Press.
- Lecaldano, E., 2002. The Passions, Character and the Self in Hume. *Hume Studies*, 28 (2), pp.175-194.
- Lewis, D., 1986. *Philosophical Papers Vol. II*. Oxford University Press.

- Locke, J., 1997. *An Essay Concerning Human Understanding*. Edited by Roger Woolhouse. London; New York: Penguin.
- Lowe, E. J., 2013. The Will as a Rational Free Power. In R. Groff and John Greco, eds. 2006. *Powers and Capacities in Philosophy: The New Aristotelianism*. New York: Routledge pp.172-185.
- Mackie, P., 2014. Mumford and Anjum on Incompatibilism, Powers and Determinism. *Analysis*, 74 (1), pp.593-603.
- Marmadoro, A., 2016. Dispositional Modality Vis-à-vis Conditional Necessity. *Philosophical Investigations*, 39 (3), pp.205-214.
- Marshall, C., 2015. Hume Versus the Vulgar on Resistance, Nisus, and the Impression of Power. *Philosophical Studies*, 172 (2), pp.305-319.
- Martin, C. B., 1994. Dispositions and Conditionals. *The Philosophical Quarterly*, 174, pp.1-8.
- McBreen, B., 2007. Realism and Empiricism in Hume's Account of Causality. *Philosophy*, 82 (3), pp.421-436.
- Milgram, S., 2005. *Obedience to Authority: An Experimental View*. London: Printer and Martin.
- Mischel, W., and Shoda, Y., 1995. A Cognitive-affective System of Personality. *Psychological Review*, 102 (2), pp.246-268.
- Molnar, G., 2003. *Powers: A Study in Metaphysics*. Oxford; New York: Oxford University Press.
- Morris, W. E., 2006. Belief, Probability, Normativity. In S. Traiger, ed. 2006. *The Blackwell Guide to Hume's Treatise*. Oxford: Blackwell Publishing ltd. pp.77-94.

- Mumford, S., 2013. The Power of Power. In R. Groff and John Greco, eds. 2006. *Powers and Capacities in Philosophy: The New Aristotelianism*. New York: Routledge pp.9-24.
- Mumford, S., and Anjum, R. L., 2011. *Getting Causes from Powers*. Oxford: Oxford University Press.
- Mumford, S., and Anjum, R. L., 2013. A New Argument Against Compatibilism. *Analysis*, 24 (1), pp.20-25.
- Mumford, S. and Anjum, R. L., 2014. Powers as Causal Truthmakers. In *Philosophical Research Bulletin 4* pp.5-31
- Noonan, H. W., 1999. *Routledge Philosophy Guidebook to Hume on Knowledge*. London; New York: Routledge.
- Pauer-Studer, H., 2009. Humean Sources of Normativity. In C. R. Pigden, ed. 2009. *Hume on Motivation and Virtue*. Basingstoke: Palgrave Macmillan pp.186-207.
- Penelhum, T., 1993. Hume's Moral Psychology. In D. F. Norton and J. Taylor, ed. 1993 *The Cambridge Companion to Hume*. Cambridge: Cambridge University Press pp.117-147.
- Pitson, A. E., 2002. *Hume's Philosophy of the Self*. London, New York: Routledge.
- Quine, W. V., 1960. *Word and Object*. New York, London: The Technology Press of The Massachusetts Institute of Technology, John Wiley and Sons, Inc.
- Rees, C. F., and Webber, J., 2014. Automaticity in Virtuous Action. In N. Snow and F. Trivigno eds. *The Philosophy and Psychology of Character and Happiness*. New York: Routledge pp.75-90.

- Reid, T., 2010. *Essays on the Active Powers of Man*. Edited by Knud Haakonssen and James A. Harris. Edinburgh: Edinburgh University Press.
- Rickless, S. C., 2013. Locke on Active Power, Freedom and Moral Agency. *Locke Studies*, 13, pp.31-52.
- Saucier, G., 1994. Mini-markers: A Brief Version of Goldberg's Unipolar Big-five Markers. *Journal of Personality Assessment*, 63, pp.506-516.
- Steward, H., 2012. , *A Metaphysics for Freedom*. Oxford: Oxford University Press.
- Stout, R., 2011. *Things That Happen Because They Should: A Teleological Approach to Action*. Oxford Scholarship Online.
- Strawson, G., 1989. *The Secret Connexion*. Oxford: Clarendon Press.
- Strawson, G., 1994. The Impossibility of Moral Responsibility. *Philosophical Studies*, 75, pp.5-24.
- Strawson, G., 2002. David Hume: Objects and Power. In Peter Millican ed. 2002. *Reading Hume on Human Understanding*. Oxford Clarendon Press, pp.231-257.
- Strawson, P. F., 1952. *Introduction to Logical Theory*. London: Methuen.
- Taylor, J., 2006. Virtue and the Evaluation of Character. In S. Traiger, ed. 2006. *The Blackwell Guide to Hume's Treatise*. Oxford: Blackwell Publishing ltd. pp.276-296.
- Van Miltenburg, N., 2015. *Freedom in Action*. CPI-Koninklijke Wöhrmann.
- Webber, J., 2013. Character, Attitude and Disposition. *European Journal of Philosophy*, 23 (4), pp.1082-1096.
- Williams, B., 1981. *Moral Luck*. Cambridge: Cambridge University Press.

Wright, J. P., 2006. Composition, Reception and Response. In S. Traiger, ed.
2006. *The Blackwell Guide to Hume's Treatise*. Oxford: Blackwell
Publishing ltd, pp.5-25.